University of the Western Cape

EMS Faculty

Institute for Social Development

An application of synthetic panel data to poverty analysis in South Africa

PhD thesis submitted to the Institute for Social Development, Faculty of Economic Management Sciences, University of the Western Cape

In fulfilment to the requirements for the degree, Doctor of Philosophy in Development Studies

Rejoice Mabhena

Student Number: 3506569

Supervisor: Professor Julian D. May

Date: 16 September 2019

Table of Contents

Table of Contents i

List of Tables v

List of Figures vi

Declaration viii

Dedication ix

Acknowledgements x

Acronyms and abbreviations xi

Abstract xiv

Chapter 1 1

1 An application of synthetic panel data to poverty analysis in South Africa ..... Error! Bookmark not defined.

1.1 Introduction and Rationale 1
1.2 The context: declining global poverty? 3
1.3 Background and Contextualization of the Study 5
1.4 Data for poverty measurement 6
1.4.1 Poverty Analysis using Cross Sectional Data in South Africa 7
1.4.2 Poverty dynamics in South Africa 10
1.5 Comparison of cross-sectional and panel data in poverty analysis 15
1.6 Statement of the Problem 19
1.7 Objectives and Research Questions 20
1.8 Research Design 22
1.8.1 Research Methodology 22
1.9 Structure of the study 23

Chapter 2: 25

2 Analyzing poverty in the absence of appropriate survey data: A literature review of alternative data approaches 25

2.1 Introduction 25
2.2 Household survey availability for poverty measurement 26
2.3 Data inadequacy in poverty measurement 31
2.4 Analysis of poverty trends with incomplete cross-sectional data: A discussion of Survey-to-Survey Imputation approaches 34
2.4.1 Asset Indices 35
2.4.2 Poverty Mapping 37
2.4.3 Regression based SSI approaches 39
2.5 A brief review of synthetic/pseudo panel data approaches applied in the analysis of poverty dynamics 42

2.5.1 Cohort based synthetic panel data approaches 43
2.5.2 Household/individual level approaches to synthetic panel data 44
2.6 Motivation for using pseudo panel data in developing countries 47
2.6.1 A place for pseudo panel data in South Africa 49
2.7 Conclusion 51

Chapter 3 53

3 Conceptualizing dynamic poverty 53
3.1 Introduction 53
3.2 Dynamic poverty 54
3.3 Understanding the distinctions between transient, chronic and structural poverty 57
3.3.1 Transient poverty 57
3.3.2 Chronic poverty 58
3.3.3 Structural poverty 59
3.3.4 Types of poverty: A summation and evaluation 60
3.4 Dynamic poverty and policy 61
3.5 Dynamic poverty in South Africa .....**Error! Bookmark not defined.**
3.6 Analysing the dimensions of dynamic poverty 64
3.6.1 Gendered dimensions 65
3.6.2 Geographical dimensions 66
3.6.3 Racial dimensions 68
3.7 Poverty reduction policy on gendered, spatial and racial dimensions of poverty 69
3.7.1 Policy responses to gendered dimensions of poverty 69
3.7.2 Spatial poverty and policy targeting 70
3.7.3 Race and poverty policy 73
3.8 Poverty Lines 73
3.8.1 Absolute poverty line 74
3.8.2 Relative poverty lines 75
3.8.3 Subjective poverty lines 76
3.8.4 Poverty lines in South Africa 77
3.9 Poverty trends and the relevance of a cohort-based pseudo panel in South Africa 80
3.10 Conclusion 84

Chapter 4 86

4 Making sense of poverty statistics in South Africa 86
4.1 Introduction 86
4.2 Methodology 86
4.3 Selection of research approach 87
4.4 Research design rationale 88
4.4.1 Key informants 89
4.4.2 Case Studies 90
4.4.3 Sampling method 91
4.4.4 Data collection summary 91
4.4.5 Data coding and analysis 92
4.5 A brief history of statistical institutions in South Africa 94

4.6 The evolution of poverty measurement in South Africa 99
4.6.1 Increase in the number of household surveys.....100
4.6.2 Evolving data collection methods 103
4.6.3 Embracing multidimensional poverty 105
4.6.4 Evolution of poverty lines in South Africa.....107
4.6.5 Analysis of poverty dynamics 108
4.7 Challenges 110
4.7.1 Funding and infrequency of survey data in South Africa.....111
4.7.2 Issues with the Population Census data.....112
4.7.3 Data disaggregation 113
4.7.4 Quality and comparability of survey data 114
4.8 Future survey conduct 116
4.9 Conclusion.....117

Chapter 5 119

5 Constructing A Cohort Based Pseudo Panel For Poverty Analysis: Data Issues And Methodology 119
5.1 Introduction 119
5.2 Constructing a cohort-based pseudo panel 120
5.2.1 Description of pseudo panel approach 121
5.2.2 Assumptions and matching techniques 121
5.3 Weighting the pseudo panel 122
5.4 Application of poverty lines 123
5.5 Age restriction of birth cohorts.....123
5.6 Selecting a household survey for the pseudo panel.....124
5.7 Selection of the LCS surveys and IES of 2010 127
5.8 Comparability of the IES 2010 and the LCS of 2008 and 2014.....130
5.9 Real and nominal expenditure values.....136
5.10 Descriptive analysis of pseudo panel 136
5.11 Conclusion.....142

Chapter 6 144

6 Estimating Expenditure Patterns and Poverty Dynamics Using a Cohort Based Pseudo Panel in South Africa 144
6.1 Introduction 144
6.2 Estimating the changes in mean expenditure per capita for birth cohorts in the pseudo panel 144
6.2.1 A gendered analysis of per capita expenditure.....147
6.2.2 A geographical analysis of mean per capita expenditure 150
6.2.3 An analysis of mean per capita expenditure by race 153
6.2.4 Discussion: Distribution of mean per capita expenditure for birth cohorts.....158
6.3 Cohort analysis: Age, cohort and time effects.....162
6.3.1 Regression Analysis: Exploring the Changes in Expenditure for Different Birth Cohorts (2008- 2014) 165

6.4 Changes in poverty status of birth cohorts in pseudo panel 168
6.4.1 A gendered analysis of poverty estimates 169
6.4.2 A geographic analysis of poverty incidence.....170
6.4.3 A racial analysis of poverty incidence.....171
6.4.4 Discussion: Poverty analysis for the birth cohorts 174
6.5 Conclusion 175

Chapter 7 176

7 Poverty measurement using alternative data: Lessons and future analysis176
7.1 Introduction 176
7.2 Summary of findings 179
7.2.1 Expansion of survey data for poverty measurement 180
7.2.2 Paucity of data for poverty measurement.....182
7.2.3 Role of NSO's in data availability for poverty measurement 183
7.2.4 Alternative data for poverty measurement 183
7.2.5 Birth cohort pseudo panel methodology.....184
7.2.6 Pseudo panel findings.....185
7.3 Significance and contribution of research 186
7.4 Relevance of study to policy 187
7.4.1 Poverty and social protection in South Africa.....188
7.4.2 Poverty and structural forms of employment 188
7.5 Final remarks and recommendations.....189
7.6 Conclusion 189

References 191

Appendix 220

List of Tables

TABLE 2.1: SURVEY AVAILABILITY FOR POVERTY ANALYSIS BETWEEN 2002-2011 27
TABLE 2.2: PANEL SURVEY CONDUCT IN DEVELOPING COUNTRIES 28
TABLE 3.1: SUMMARY OF EARLY RESEARCH WORK ON DYNAMIC POVERTY IN DEVELOPING COUNTRIES (1980'S-1999) 55
TABLE 4.1: SUMMARY OF STATS SA PUBLICATIONS ON POVERTY IN SOUTH AFRICA 101
TABLE 5.1: SURVEY AVAILABILITY FOR POVERTY ANALYSIS IN SOUTH AFRICA 125
TABLE 5.2: DATA COLLECTION FOR THE LCSS AND IES OF 2010 131
TABLE 5.3: AVERAGE EXPENDITURE IN THE LCS AND IES SURVEYS (NOMINAL VALUES) 133
TABLE 5.4: RESPONSE RATES FOR THE IES AND LCSS 135
TABLE 5.5: URBAN AND RURAL MAKE-UP OF SOUTH AFRICAN PROVINCES 136
TABLE 5.6: NUMBER OF BIRTH COHORTS IN PSEUDO PANEL 137
TABLE 6.1: MEAN EXPENDITURE PER CAPITA (OVERAL) 147
TABLE 6.2: TOTAL MONTHLY HOUSEHOLD EXPENDITURE PER CAPITA FOR COHORTS (MHH) 149
TABLE 6.3: TOTAL MONTHLY HOUSEHOLD EXPENDITURE PER CAPITA FOR COHORTS (FHH) 150
TABLE 6.4: TOTAL MONTHLY HOUSEHOLD EXPENDITURE PER CAPITA FOR COHORTS (URBAN) 152
TABLE 6.5: TOTAL MONTHLY HOUSEHOLD EXPENDITURE PER CAPITA FOR COHORTS (RURAL) 153
TABLE 6.6: MEAN PER CAPITA EXPENDITURE BY RACE WITH CI (2008) 155
TABLE 6.7: MEAN PER CAPITA EXPENDITURE BY RACE OF HEAD CI (2010) 156
TABLE 6.8: MEAN PER CAPITA EXPENDITURE BY RACE OF HEAD CI (2014) 157
TABLE 6.9: PERCENTAGE CONTRIBUTION OF SALARY/WAGES AND SOCIAL GRANTS TO TOTAL INCOME 159

List of Figures

FIGURE 3.1: POVERTY HEADCOUNT FOR AGE COHORTS (2006-2015). 82
FIGURE 3.2: POVERTY HEADCOUNT BY RACIAL GROUPING (2006-2015). 83
FIGURE 3.3: POVERTY HEADCOUNT FOR URBAN AND RURAL HOUSEHOLDS (2006-2015) 84
FIGURE 3.4: POVERTY HEADCOUNT FOR MHH AND FHH (2006-2015) 84
FIGURE 4.1: SUMMARY OF QUALITATIVE METHODOLOGY APPLIED IN THIS CHAPTER. .. 94
FIGURE 5.1: DISTRIBUTION OF BIRTH COHORTS (2008-2014). 138
FIGURE 5.2: DISTRIBUTION OF BIRTH COHORTS FROM THE 2011 POPULATION CENSUS 10% SAMPLE. 139
FIGURE 5.3: DISTRIBUTION OF BIRTH COHORTS BY GENDER. 140
FIGURE 5.4: DISTRIBUTION OF BIRTH COHORTS BY GEOGRAPHY. 141
FIGURE 5.5: MEAN HOUSEHOLD SIZE OF BIRTH COHORTS. 142
FIGURE 6.1: MEAN EXPENDITURE PER CAPITA FOR ALL BIRTH COHORTS. 145
FIGURE 6.2: MEAN EXPENDITURE PER CAPITA (IN RANDS) BY GENDER OF HOUSEHOLD HEAD. 148
FIGURE 6.3: MEAN EXPENDITURE PER CAPITA BY GEOGRAPHY. 151
FIGURE 6.4: MEAN PER CAPITA EXPENDITURE BY RACE OF HEAD (2008). 154
FIGURE 6.5: MEAN PER CAPITA EXPENDITURE BY RACE OF HEAD 2010. 155
FIGURE 6.6: MEAN PER CAPITA EXPENDITURE BY RACE OF HEAD (2014). 157
FIGURE 6.7: PERCENTAGE CONTRIBUTION OF SALARIES AND SOCIAL GRANTS TO INCOME FOR ALL BIRTH COHORTS. 161
FIGURE 6.8: CONDITIONAL MEAN OF LOG EXPENDITURE BY AGE. 166
FIGURE 6.9: CONDITIONAL VARIANCE OF THE LOG OF EXPENDITURE FOR YEARLY BIRTH COHORTS OVER TIME. 167
FIGURE 6.10: POVERTY HEADCOUNT FOR ALL BIRTH COHORTS (POVERTY HEADCOUNT). 169
FIGURE 6.11: POVERTY INCIDENCES BY HEAD OF HOUSEHOLD (POVERTY HEADCOUNT). 170
FIGURE 6.12: A COMPARISON OF POVERTY INCIDENCES FOR BIRTH COHORTS IN URBAN AND RURAL HOUSEHOLDS (POVERTY HEADCOUNT). 171
FIGURE 6.13: POVERTY HEADCOUNT FOR ALL BIRTH COHORTS BY RACE (2008). 172
FIGURE 6.14: POVERTY HEADCOUNT FOR ALL BIRTH COHORTS BY RACE (2010). 173

FIGURE 6.15: POVERTY HEADCOUNT FOR ALL BIRTH COHORTS BY RACE (2014).....173

Declaration

I hereby declare that this thesis titled, '*An application of synthetic panel data to poverty analysis in South Africa*' is a product of my own research and that not any part of it has been part of any degree elsewhere. All references have been duly acknowledged.

Rejoice Mabhena

……..

Student

Professor Julian D. May

…….

Supervisor

Dedication

I dedicate this thesis to my daughter, Roxanne Tarisai for her undying and pure love despite my absence. May this thesis inspire you to be a woman of influence and power; you are truly God's gift to me. I love you.

Acknowledgements

I am eternally grateful to God for giving me life, good health and for taking me this far in my life. I would like to thank my Supervisor and role model, Professor Julian May for his brilliance and support throughout the writing of this thesis. Your extensive work of poverty dynamics in South Africa inspired my thesis, thank you. I value greatly your academic, financial and moral support that made it possible for this thesis to be completed. I also appreciate immensely the mentorship I received from Professor Murray Leibbrandt and Dr. Corretta Jonah. Your guidance has paid off.

To DAAD, I am grateful for the financial support that enabled me to research extensively. I also appreciate the work from staff members from the South African German Centre for Development Research (SA-GER-CDR) Dr. Stefan Buchholz, Belinda du Plessis and former DAAD Coordinator Dr. Ina Conradie, thank you for all your efforts in ensuring the success of the DAAD Programme at UWC. I also extend my appreciation to the rest of the DAAD PhD students who I have worked with, as well as ISD PhD students, Fru Wanka, Tsega Tefera, Alex Boakye, Evans Boadu Kwasi, Kwesi Aloysius Agwani, Keisha, Adeyemi Badewa, Grace Nkomo, Shamiso Mandioma, Michael Nguatem, Jonas Nzabamwita and Post-Doctoral Fellow Dr. Daniel Adeniyi. I further extend my thanks to the DST-NRF Centre of Excellence in Food Security for the extended financial support. Special thanks to Dr. Camilla Adelle, Elaine Petersen and Noluntu Didiza and the rest of the CoE team. To Priscilla Kippie, thank you for your assistance over the years. I appreciate your input.

Finally, a special thanks goes to my family and friends for the much-needed support they gave me in the writing of this thesis. Thank you to my mother Mrs C. Mabhena, and younger siblings Audrey Mabhena, Faith Mabhena and Kimberley Mhunduru for believing in my vision and never leaving my side. Special mention to my closest friends Moline Rusere, Francis Chiparawasha, Takunda Satumba and Irene Nyamu for their insights and moral support. My thesis would not have been complete without the guidance of my spiritual parents Pastor Jolly Ighodharo and Justine Ighodaro and my spiritual leaders Professor Apostle Johnson Suleman, Dr. Lizzy Suleman, Dr. Sule Emmanuel, Dr Tomi Sule and Pastor Christian Vainqueur Kasweka. Thank you for your prayers and spiritual guidance during the writing of this thesis. To everyone else who took time to listen to my PhD stories, I appreciate you.

Acronyms and abbreviations

AMPS	All Media Products Survey
ASGISA	Accelerated and Shared Growth Initiative for South Africa
BT10	Birth to Ten birth cohort
BT20	Birth to Twenty birth cohort
CAPI	Computer Assisted Personal Interview
CAPS	Cape Area Panel Study
CPRC	Chronic Poverty Research Centre
CS	Community Survey
EPWP	Expanded Public Works Programme
FAO	Food and Agriculture Organization
FHH	Female Headed Household
FPL	Food Poverty Line
GEAR	Growth, Employment and Redistribution
GDP	Gross Domestic Product
GHS	General Household Survey
IES	Income and Expenditure Survey
IFAD	International Fund for Agricultural Development
KIDS	KwaZulu Natal Dynamic Study
LBPL	Lower Bound Poverty Line
LCS	Living Conditions Survey

LFS	Labour Force Survey
MDGs	Millennium Development Goals
MHH	Male Headed Household
MIF	Mo Ibrahim Foundation
NSO	National Statistical Office
NSS	National Statistical System
NDP	National Development Plan
NIDS	National Income Dynamic Study
NSSD	National Statistical System Development
OHS	October Household Survey
OECD	Organization for Economic Cooperation and Development
PRSP	Poverty Reduction Strategy Paper
PSLSD	Project for Statistics on Living Standards and Development
QLFS	Quarterly Labour Force Survey
SALDRU	Southern African Labour Development Research Unit
SAMPI	South African Multidimensional Poverty Index
SASQAF	South African Statistical Quality Assessment Framework
SSI	Survey to Survey Imputation
Stats SA	Statistics South Africa

SSA	Sub Saharan Africa
TBVC	Transkei Bophuthatswana, Venda and Ciskei
UBPL	Upper Bound Poverty Line
UN	United Nations
USAID	United States Agency for International Aid
WFP	World Food Programme

Abstract

There is a wide-reaching consensus that data required for poverty analysis in developing countries are inadequate. Concerns have been raised on the accuracy and adequacy of household surveys, especially those emanating from Sub-Saharan Africa. Part of the debate has hinted on the existence of a statistical tragedy, but caution has also been voiced that African statistical offices are not similar and some statistical offices having stronger statistical capacities than others. The use of generalizations therefore fails to capture these variations. This thesis argues that African statistical offices are facing data challenges but not necessarily to the extent insinuated.

In the post-1995 period, there has been an increase in the availability of household surveys from developing countries. This has also been accompanied by an expansion of poverty analyses efforts. Despite this surge in data availability, available household survey data remain inadequate in meeting the demand to answer poverty related enquiry. What is also evident is that cross sectional household surveys were conducted more extensively than panel data. Resultantly the paucity of panel data in developing counties is more pronounced. In South Africa, a country classified as 'data rich' in this thesis, there exists inadequate panel surveys that are nationally representative and covers a comprehensive period in the post-1995 period. Existing knowledge on poverty dynamics in the country has relied mostly on the use of the National Income Dynamic Study, KwaZulu Natal Dynamic Study and smaller cohort-based panels such as the Birth to Twenty and Birth to Ten cohort studies that have rarely been used in the analysis of poverty dynamics.

Using mixed methods, this thesis engages these data issues. The qualitative component of this thesis uses key informants from Statistics South Africa and explores how the organization has measured poverty over the years. A historical background on the context of statistical conduct in the period before 1995 shows the shaky foundation that characterised statistical conduct in the country at the inception of Statistics South Africa in 1995. The organization since then has expanded its efforts in poverty measurement; partly a result of the availability of more household survey data. Improvements within the organization also are evidenced by the emergence of a fully-fledged Poverty and Inequality division within the organization. The agency has managed to embrace the measurement of multidimensional poverty. Nevertheless, there are issues surrounding

available poverty related data. Issues of comparability affect poverty analysis, and these are discussed in this thesis. The informants agreed that there is need for more analysis of poverty using available surveys in South Africa.

Against this backdrop, the use of pseudo panels to analyse poverty dynamics becomes an attractive option. Given the high costs associated with the conduct of panel surveys, pseudo panels are not only cost effective, but they enable the analysis of new research questions that would not be possible using existing data in its traditional forms. Elsewhere, pseudo panels have been used in the analysis of poverty dynamics in the absence of genuine panel data and the results have proved their importance.

The methodology used to generate the pseudo panel in this thesis borrows from previous works including the work of Deaton and generates 13 birth cohorts using the Living Conditions Surveys of 2008/9 and 2014/15 as well as the IES of 2010. The birth cohorts under a set of given assumptions are 'tracked' in these three time periods.

The thesis then analysed the expenditure patterns and poverty rates of birth cohorts. The findings suggested that in South Africa, expenditures are driven mostly with incomes from the labour market and social grants. The data however did not have adequate and comparative variables on the types of employment to further explore this debate. It also emerged that birth cohorts with male headship as well as birth cohorts in urban settlements and in White and Indian households have a higher percentage share of their income coming from labour market sources. On the other hand, birth cohorts with female headship and residing in rural, African and in Coloured households are more reliant on social grants. The majority of recipients of social grants receive the Child Social Grant and its minimalist value partly explains why birth cohorts reporting social grants as their main source of income are more likely to be poor when compared to birth cohorts who mostly earn their income from the labour market. Residing in a female-headed household, or in a rural area as well as in Black African and Coloured increases the chances of experiencing poverty. This supports existing knowledge on poverty in South Africa and confirms that these groups are deprived. The results of the pseudo panel analysis also show that poverty reduced between 2006 and 2011 for most birth cohorts but increased in 2015. Policy recommendations to reduce poverty therefore lie in the labour market. However, given the high levels of unemployment in the country today, more rigorous labour incentives are required.

Chapter One

An application of synthetic panel data to poverty analysis in South Africa

1.1 Introduction and Rationale

Existing knowledge on global poverty is reliant upon the data sources that are available and whether these are adequate and able to estimate poverty within reasonable margins of error. Household surveys undertaken by National Statistical Offices (NSOs) are a widely accepted source of data and are used to make national and global comparisons on who is lagging behind and who is advancing in poverty reduction. Nonetheless, concerns have been raised on the accuracy and adequacy of household surveys, especially those emanating from Sub-Saharan Africa (Jerven, 2009; 2013; 2014; 2016).

The contention that data required for poverty analysis in developing countries are inadequate is not new and is well documented (Beegle et al. 2016; Dabalen et al. 2014; Douidich et al. 2013; Grosse et al. 2009; Hentshel et al. 2000; Jolliffe et al. 2014; Serajuddin et al. 2015). Equally, the centrality of data in poverty measurement in Sub Saharan Africa was placed in the limelight over the period of the Millennium Development Goals (MDGs). Efforts to assess progress made towards achieving the MDGs brought scrutiny on the statistical capabilities of different African countries and increased the demand for quality and reliable statistics from developing countries. This call for improvements in existing data collection methods placed pressure on the National Statistical Systems (NSSs) of developing countries that were already frail (Sanga, 2011).

Apart from work that focused on the inadequacy of poverty related data, arose another group of critics who questioned the reliability and accuracy of existing household survey data in developing countries. Such critical work includes the works of Devarajan (2013); Jerven (2009; 2013a; 2013b; 2014a, 2014b); Jerven and Johnston (2015); Sandefur and Glassman (2015); Young (2013). In particular, Jerven (2009; 2014a) questioned the reliability of African GDP estimates, and alerted that such figures should be used with caution, as they can be misleading Interestingly, Jerven (2016) also notes the existence of a paradox as part of the data debate associated with African statistics - developing countries have inadequate data, yet they need it the most.

Conceivably, one of the most contested arguments in the validity of African statistics is pursued at length in research conducted in Jerven (2012; 2013a; 2013b; 2014a; 2014b). Jerven (2013a:84) scrutinized African statistics and highlighted the existence of immense discrepancies in the methodologies used for data collection and ultimately in the sources of data in African countries. African statistical capacities have been criticized at length with part of the literature arguing the existence of a 'statistical tragedy' on the continent (Devarajan 2013; Jerven 2016; Jerven and Johnston 2015).

This critique on African statistics did not emerge in the post-2000 period. Tracing back on earlier critical work on African household survey data, Frankel (1953) likened African statistics that were produced in the immediate period following independence as 'fictional'; marking one of the earliest criticisms levelled at statistics from developing countries. More recently, Young (2013:696) also contended that statistics used to support the economic growth and development in poor countries is 'minimal or non-existent'.

What is apparent in the views expressed in the literature above is the low regard given to statistics coming from developing countries. However, it should be noted that African countries have dissimilar statistical capacities and as such should be viewed differently. Jerven and Johnston (2015) in line with this caution acknowledge that African statistical offices are not homogenous, noting that some statistical offices in African countries have stronger statistical capacities than others.

The contention of this thesis is that existing criticisms on African statistics may have exaggerated the poor performance of official statistics from African NSOs and are based on a few case studies that are not necessarily representative of the entire continent. A data challenge exists but not necessarily to the extent insinuated in the previous arguments. I argue that effective poverty monitoring and its reduction commences with accurate, frequent and reliable statistics. Yet, given the stringent budgets faced by most African central statistical offices, it will be a long time before the demand for such statistics are fully met. Further, Jerven (2017) warns that accurate statistics do not necessarily translate into improved policy choices. This lack of a linear relationship between accurate data and right policy choices is partly explained by the fact that the main constraints in wrong policy choices are not necessarily centred on a general lack of accurate statistics (Jerven 2013a). Notwithstanding this, this study

argues that capturing poverty trends and dynamics more accurately remains a starting point in fighting this challenge.

This thesis critically engages with the 'data issues' associated with poverty measurement in Sub Saharan Africa using South Africa as a case study. The study then advances the analysis of poverty by utilizing recent data management techniques that expand the potential usage of existing official household survey data. The term 'data issues' in this thesis is applied to refer to the unavailability of frequent, comparable, or adequate household survey data for poverty analysis. The term is used in reference to both cross sectional and panel data in this thesis.

The objective of this study is to attempt to move beyond these critiques to explore solutions. As one option to improving the use of existing data, a pseudo panel is generated using 3 cross sectional household surveys from South Africa in this thesis. Apart from the use of pseudo panel methods to enable the analysis of poverty dynamics, the use of Survey-to-Survey Imputation (SSI) methods will also be explored as one of the ways of leveraging existing household survey data to create new 'alternative' data sources for poverty analysis.

Prior to the application of such data manipulation techniques, this study also explores how existing data types used in poverty analysis contribute to poverty analysis using qualitative methodologies. Based on in-depth interviews with informed respondents in Statistics South Africa (Stats SA), the objective of these comparisons is to highlight the data issues and gaps associated with poverty related household surveys and how they can be dealt with.

In short, the arguments made in Gibson (2016) suggest that there is inadequate data on poverty measurement in development countries and we know less than we think. More surveys have been conducted in the post-1990 period, but such efforts remain inadequate.

1.2 The context: declining global poverty?

The conceptualization of poverty has changed significantly over the years and the consensus reached concurs that it is multidimensional, incorporating both monetary and non-monetary dimensions and is non-static. Despite this realization, poverty remains a problem that most developing countries are still battling with. Globally, existing

evidence from the Millennium Development Goals (MDGs) reports suggest that poverty has significantly declined in most regions. United Nations (UN) (2015) in its final MDG report stated that globally, the population living on less than $1.25 a day has declined from 47% in 1990 to 14% by 2015. This decline translates to a reduction of people living in extreme poverty from 1.9 billion to 836 million during this reference period. Using the most recent extreme poverty line of $1.90 a day, The World Bank (2016:3) estimated that 10.7% of the global population or 767 million people were poor in 2013. Atamanov et al. (2018) also support a similar trajectory of global poverty using a similar measure and estimated a total of 783 million poor people globally in 2013. This is a significant reduction from a globally poor population estimate of more than 1.8 billion in 1990 (World Bank 2016). This concludes that poverty has been reducing at an average rate of 1.1 percentage points per year since 1990 (World Bank 2016).

Notwithstanding this progress evident in global poverty statistics, the number of people living in extreme poverty has increased in some regions of the world. For instance, the headcount of people living in extreme poverty in Africa increased by more than 100 million between 1990 and 2012, despite a reduction in the total share of the African population living in extreme poverty during the same time period (World Bank 2016).

Part of this increase in extreme poverty incidences is partially explained by the increase in the African population. United Nations Economic Commission for Africa (UNECA) (2016) notes that there has been an increase in the global population by 2.9 billion between 1980 and 2015. UNECA (2016) found that Africa's population during the last 35 years increased from 478 million to close to 1.2 billion during the same reference period. Future projections by UNECA (2016) also suggests that despite the fact that population growth rates have slowed down in other parts of the world, Africa's population will continue to grow to over 2.4 billion by 2050 with a growth rate of at least 1.5% in the next ten years.

Despite the reductions recorded in global poverty between 1990 and 2015, these advances in poverty reduction have been unequally distributed as most of the progress was reported in Asia[1] (Food and Agriculture Organization (FAO), International Fund

[1]India and China's progress in reducing extreme poverty from 61% to 4% in 2015 resulted in a significant decline poverty of global poverty (UN 2015c; WFO, IFAD and FAO (2002) www.mdgmonitor/mdg-1-eradicate-poverty-hunger)

for Agricultural Development (IFAD), World Food Program (WFP) 2002.). In Sub-Saharan Africa (SSA), evidence shows that there were 58 million more poor people in 1999 as compared to 1990 (UN 2015).With this evidence relative to other regions, it is not surprising therefore that in 2015, SSA continued to lag behind the rest of the world in fighting poverty.

The World Bank (2016) states that SSA has the highest poverty headcount when compared to other regions. To be more specific, 41.0% of the poor or 389 million people from a total of 767 million globally poor are located in this region (World Bank 2016). UN (2015) also supports similar estimates and states that more than 40% of the population in SSA still lives in extreme poverty and statistics showed a poverty rate of 28% in the same year (UN 2015). Despite the 15-year effort that has been exerted in achieving poverty reduction across the world, more than 800 million people today are still experiencing hunger and poverty with the most affected people in SSA (UN 2015). SSA's continued stagnation in poverty reduction can be partly explained by the fact that 28 out of the 31 low human development countries in the world are located in this region (Handley et al. 2009).

Statistics such as these have increasingly become central in poverty diagnostics and are key to the design of programs and projects. Yet, the inadequacy of poverty related survey data in developing countries is evident. Alkire (2014:1) concurs to this effect and adds that 'data on poverty are severely limited in terms of both coverage and frequency' in most African countries.

The recent Sustainable Development Goals' (SDG's) framework continues to highlight the importance of statistics in the monitoring of development related goals. Beegle et al. (2016) found that African countries only managed to conduct half as many surveys as did the developed world. They state that African countries conducted an average of 3.8 household consumption surveys since the inception of the MDGs and only 1.6 of these are comparable.

1.3 Background and Contextualization of the Study

The roots of South Africa's poverty have been well documented over the years. The accounts show that since the inception of democracy in 1994, the country prioritized poverty reduction as a key component in its policies (Triegaardt 2009). This was not

unanticipated given the fact that prior to 1994, the Apartheid government provided services based on racial lines and resultantly created higher poverty rates for the African population more than the other racial groups in South Africa, a phenomenon that was unacceptable at the time of transition and that is still evident today.

Johannes et al. (2005) state that South Africa inherited immense inequalities in the access to all basic services as a result of Apartheid. During this era existing legislation crafted policies that promoted racial segregation with homelands as the epitome of such policies. The result was that at the onset of democracy, these homelands where a majority of the African population were concentrated had limited economic activity, translating to high levels of unemployment in the country. Thus, the government in its bid to reverse these inequalities introduced a myriad of policies to attempt to address these imbalances. [2]

Of significance as well in the analysis of poverty in South Africa is that the country in the post-1994 period has managed to capture its poverty trends more laudably than other countries in Sub Saharan Africa due to the availability of a significant number of poverty related surveys in the country. Yu (2013) states that since 1994 the number of surveys that have been used in poverty analysis in South Africa has been growing. However, in as much as the availability of these surveys makes South Africa a 'data rich' country, issues relating to the comparability of these surveys discredit their meaningful contribution to poverty measurement.

1.4 Data for poverty measurement

The post-1990 period has witnessed a lot of data being produced to monitor national and international development related quantifiable goals in developing countries. Despite this increasing need for quality data, most developing countries cannot meet the current demand for poverty relevant statistics (Kamanaou et al. 2005). More recently, Jolliffe et al. (2014) and Serajuddin at al. (2015) acknowledge that data availability in these countries improved, but data gaps still exist. Evidence from existing poverty literature concurs that in most developing countries the available data sources on poverty estimation are not frequent and accurate (Grosse et al. 2009; Douidich et al. 2013). As stated by Dabalen et al. (2014), the need for such high-quality

[2] Such policies include the Rural Development Program of 1994; the Growth Employment and Redistribution program of 1996.

data in the developing world continues to outperform its production. The high costs associated with the conduct of household surveys used in poverty analysis remains at the centre of these data inadequacies.

Two broad categories of poverty analyses can be distinguished: cross sectional poverty measurement and analyses that use panel data. The distinction between these 2 broad categories in this thesis enables for gaps to be identified for each type of data. Evidence shows that most developing countries have inadequate poverty relevant data but there is a general paucity of panel data.

1.4.1 Poverty Analysis using Cross Sectional Data in South Africa

Cross sectional data can be defined as data that comprises of a sample of individuals or households taken from a given point of time (Wooldridge 2009). Cross sectional data allows for the analysis of poverty on households or individuals at given points in time. These data are the most frequently used in poverty analysis.

The assertion made earlier that South Africa today can be viewed as 'data rich' blurs out the issues associated with the available household surveys, but rather enumerates the number of household surveys that enable poverty enquiry. This is confirmed by that fact that there exists a significant number of household surveys that have been conducted in the post-1994 period that have contributed to what we know about poverty today in South Africa. This increase is also documented in Yu (2013) in the post-1994 period in South Africa.

This was not always the case, prior to 1994, South Africa had a paucity of nationally representative poverty related household survey data. In line with this assertion, Finn et al. (2014b:1) note that the early 1990s were characterized by 'little data on household income and expenditures. May (2010) in Smokes and mirrors gave a brief historical context of poverty in the country on the poverty trends in the pre-1994 period. His findings suggest that South Africa had high poverty rates in the 1960s and 1970s and these high levels were mostly concentrated in disadvantaged areas.

In a turn-around, South Africa has managed to conduct a significant number of cross-sectional surveys that are nationally representative and relevant to poverty analysis in the period after 1994. A significant body of literature that analyses how poverty has changed over the years in South Africa has also grown in the post-1994 period. The

majority of such studies used cross sectional data (Bhorat and Kanbur 2005; Bhorat and Westhuizen 2012; Finn et al. 2014a; Leibbrandt et al. 2010; May and Hunter 2004; van der Berg and Louw 2004; van der Berg et al. 2006; van der Berg 2011; The World Bank 2018).

The expansion of cross-sectional household surveys in the post-1994 period in South Africa has allowed for different modules on household expenditures, income and welfare to be captured differently. What also stands out in the analysis of poverty in South Africa is that the poverty trajectory of the country in the post-1994 period is contested.

There is a broad consensus in the post-1994 period that South Africa experienced positive economic growth (Bhorat and Westhuizen 2012; du Plessis and Smit 2006; Finn et al. 2014a; Hanival and Maia 2008). The expectation therefore is that such positive figures are to be associated with a significant reduction in the country's poverty rates, but this has not been the case. Finn et al. (2014a), for instance, argue that South Africa had 'sluggish' poverty reduction achievements with poverty rate reducing from 34% to 30% between 1993 and 2008[3]. In a separate study May (2010:25) found that the late 1990s and early 2000s saw a decline in poverty rates in South Africa.

These disagreements in the country's poverty trajectory are well captured in Leibbrandt et al. (2010) and in Bhorat and Westhuizen (2012). In a report made for the Organization for Economic Cooperation and Development (OECD), Leibbrandt et al. (2010) divided existing poverty literature into 2 broad categories; studies conducted in the 1995-2000 period and the second category with studies conducted in the post-2000 timeline.

Between 1995 and 2000, poverty estimates in Hoogeven and Ozler (2006) and Simkins (2005) suggest that poverty increased in South Africa. Hoogeven and Ozler (2006) found, for instance, that the number of South Africans living below the poverty line increased from 12.6 million in 1995 to 14.4 million in 2000 using the PPP$1 per day poverty measure. The number of poor South Africans increases when the poverty measure changes to the PPP $2 a day, from 22.9 million people to 25.2 million between 1995 and 2000. Simkins (2005) also found results that support this trajectory. Using the

[3] This poverty rate uses the US$2 a day poverty line

IES of 1995 and 2000 and the Population Census of 1996 and 2001, Simkins (2005) also concluded that poverty increased from 29% to 34% using a poverty line of R800 household level poverty line.

Despite these findings that show that poverty between 1995 and 2000 increased in South Africa, another group of studies found results on the contrary. United Nations Development Programme (UNDP) (2004); van der Berg and Louw (2004) and van der Berg (2006) found that poverty either reduced or stabilized in a similar period. The findings from a report by the UNDP (2004) concluded that poverty between 1995 and 2000 declined slightly, but this result is not sustained if the poverty gap is used. Instead the poverty gap estimates show that poverty actually increased between 1995 and 2000 (UNDP 2004). Van der Berg and Louw (2004) and van der Berg et al. (2006) also concluded that poverty either stabilized or declined between 1995 and 2000. It is worth noting these two studies adjusted the IES of 1995 and 2000 to match estimates from other data sources including data from National Accounts in South Africa.

Additional disagreements on poverty trends in the post 1994 period in South Africa are highlighted by the debate between van der Berg et al. (2005) and Meth (2006). Van der Berg et al. (2005) explored poverty and income distribution in the post-transition period in South Africa using the All Media Products Survey (AMPS) and a combined dataset that utilized data from a number of surveys and data from national accounts, employment related surveys including the Labour Force Surveys and data on social grants. They found a downward trend in poverty incidence starting in 1996 and becoming steeper in 2000. Their findings also suggest that the poverty headcount in South Africa reduced from 18.5 million to 15.4 million between 2000 and 2004.

On the contrary, Meth (2006) refuted the findings by van der Berg et al. (2005), arguing they were too low. Meth (2006) argues that the data used by van der Berg et al. (2005) underreports income and expenditure and as result criticizes the ultimate findings as flawed. Instead, Meth (2006) found that the poverty headcount fell from a higher poverty headcount of 19.5 million to 18 million during a similar time period.

The debate continues, as the application of different poverty lines and surveys do not always show converging results on poverty. For example, Finn et al. (2014a) examined poverty in South Africa for a period covering more than 20 years. Their conclusions suggest the existence of a 'sluggish' poverty reduction trajectory in South Africa. They

found that in 2012, the share of people living below a US$2 a day only decreased from 34% in 1993 to 30% in 2008.

Common problems with all of these studies include the use of different poverty lines, the use of inappropriate methods such as household income levels that are not adjusted for household size, and the use of incomparable surveys. The most immediate challenge therefore is that there is no broad consensus on the poverty trends in South Africa in the post-1994 period. Instead there exists varying statistical analogies on the country's performance.

Addressing these problems through the establishment of a suite of national poverty lines, Stats SA (2014a) has emerged as an increasingly authoritative voice on South African poverty trends. Using the IES survey of 2005 and 2010/11, they argue that the country between 2006 and 2011 successfully reduced poverty in South Africa from 57.2% to 45.5% using the Upper Bound Poverty Line (R322). They also found that there was a significant reduction in the poverty gap ratio in South Africa. Using the Upper Bound Poverty Line, they argue that the gap reduced from 26.7% in 2006 to 19.6% in 2011. The analysis by Stats SA suggests that poverty has significantly reduced in South Africa and they attribute this reduction to various government policies relating to poverty alleviation including the country's social nets that have been expanded in the post 1994 period (Stats SA 2014a).

Based on these findings, one may argue that the availability of cross-sectional surveys in South Africa relevant in poverty analysis has paved way for extensive poverty analysis in the country. But this approach presents only a static snapshot of poverty trends and does not address the underlying dynamics.

1.4.2 Poverty dynamics in South Africa

Poverty analysis using panel data allows for a deeper analysis of poverty since it follows similar individuals or households over time. In South Africa, Carter and May (2000); Finn and Leibbrandt (2013); Finn et al. (2012); Gradin (2012); Leibbrandt et al. (2010); May and Woolard (2007); May et al. (2000); May et al. (2011); Woolard and Klasen (2005); May, Woolard and Baulch (2011); von Fintel and Zoch (2015); Yu (2013) are part of the growing body of studies that used panel data in poverty analysis.

One of the first forms of panel data used in poverty analysis in South Africa followed 70 households in KwaZulu-Natal between 1982, 1985 and 1993 and was created by Elisabeth Ardington. The panel followed households that were selected in rural KwaZulu-Natal over a period of 11 years on the social and economic circumstances of households (Ardington 1995). Analysis emanating from this panel analysed the factors that explain why households experience poverty and not necessarily on the extent of chronic poverty (Carter and May 2001).

A majority of studies that analyses poverty dynamics in South Africa have used the KIDS or NIDS. The KIDS was a panel survey that was conducted in KwaZulu Natal province and to date has three waves; 1993, 1998 and 2004. The sample for KIDS was drawn from the Project for Statistics on Living Standards and Development (PSLSD) original sample of 1993 (May et al. 1999). The Project for Statistics on Living Standards and Development (PSLSD) survey in 1993 was the first household survey in South Africa that was comprehensive of all population groups (May et al. 1999). (PSLSD 1994: i) states that the survey sought to collect data on the living conditions of South Africa. However, the PSLSD of 1993 could not allow the analysis of poverty dynamics since it was a cross sectional survey resulting in the introduction of the KIDS (May et al. 1999).

The KIDS used data from the PSLD as the baseline and then collected information from 1200 households in 1993, 1998 and 2004. This study provided insights on poverty dynamics in the first decade of democracy in South Africa (Aguero et al. 2007). It became the first large-scale panel survey that enabled the analysis of inter-temporal comparisons of similar households in the country (Aliber 2001; May 2001).

Adato et al. (2006); Aguero et al. (2007); Baulch (2011); Carter and May (2001); May and Woolard (2007); May et al. (1999); May et al. (2001); May et al. (2007); Roberts (2000); Woolard and Klasen (2005), are some of the studies that have used KIDS data to analyse poverty dynamics in South Africa. Analysis from the first two waves of KIDS shows the existence of chronically poor households in 1998. May et al. (1999) found that two thirds of households that were below the poverty line in 1993 remained in poverty in 1998 suggesting this proportion of poor households were either persistently or chronically poor in 1998. Carter and May (2001) found that of the households in the sample, 22% were poor in 1993 and 1998 and that households that

were found just below the poverty line in 1993 had a higher chance of moving above the poverty line in 1998 than those households that were further from the poverty line (Carter and May 2001).

Disaggregating the analysis by geographic area, Roberts (2000) using the KIDS data also found that a majority of chronically poor households were in rural areas when compared to metropolitan or urban households and were more likely to be in female-headed households. Roberts (2000) attributed this high prevalence of chronically poor households in rural areas to be partly a result of the low-income generating potential of agriculture that is the dominant economic activity in rural areas.

Adato et al. (2006) using the first two waves of KIDS analysed dynamic asset poverty in South Africa between 1993 and 1998 and found that a significant proportion of the South African population was trapped in asset poverty. Adato et al. (2006) further pursued this study and confirmed the existence of poverty traps as well as limited upward mobility of poor households using qualitative data.

A gendered analysis of the first two waves of KIDS found female headship to have increased in the 1998 KIDS since 1993 (Roberts 2000). Chronically poor households were also found to have higher illiteracy rates, and this is despite the reduction in illiteracy rates between 1993 and 1998 (Roberts 2000). As expected, chronically poor households had a higher food share and had lower income and expenditure figures (Roberts 2000).

The NIDS is the first national household panel survey in South Africa with its first wave completed in 2008 (Bhorat et al. 2009). As the KIDS remains restricted to KwaZulu-Natal province, NIDS is that only source of panel survey in South Africa that has a sample that is nationally representative and is relevant to poverty analysis.

The NIDS is conducted by the Southern African Labour and Development Research Unit (SALDRU) at the University of Cape Town (NIDS 2019). In 2008, 7305 households were identified with about 28 255 individuals and this became the first wave of NIDS (Bhorat et al. 2009). The first wave of NIDS commenced in February 2008 and the subsequent waves are undertaken every 2 years collecting information on income/expenditure (NIDS 2019). To date, NIDS has been conducted in 2008, 2010/11, 2012, 2014/15 and 2016. The data is collected from households in all the nine

provinces of South Africa. The availability of the NIDS data since 2008 has enabled the analysis of poverty dynamics in the post-2000 period.

Argent et al. (2009); Finn and Leibbrandt (2013, 2017); von Fintel and Zoch (2015) and Gradin (2012); Schotte et al. (2017a, 2017b, 2018); and Zizammia et al. (2018) have used NIDS data in the post 2008 period to analyse the movements of households in and out of poverty as well as the characteristics associated with such households. Existing literature from NIDS have utilized different waves of this survey. Finn et al. (2013) used the first two waves of NIDS and analysed absolute and relative transitions into and out of poverty in South Africa between 2008 and 2010/11. They concluded that households that were poor in 2008 remained poor in 2010/11. An extension of this analysis using three waves of NIDS by Finn and Leibbrandt (2013) found that there were higher incidences of poverty exits between waves two and wave three than there were between waves one and two. Since NIDS is the only panel that has a nationally representative sample in South Africa, the survey limits the analysis that can be done on poverty using panel data. Finn and Leibbrandt (2013) in their analysis acknowledged the 'paucity' of research in this field.

This paucity of panel data which is characteristic of most developing countries is partly explained by and Tangamodzi (2016) who argues that panel data require more human and financial resources than cross sectional data. In understanding these high costs associated with panel data, the argument is that the first wave may have the same costs as cross sectional surveys of similar size and data collection methods, but the real costs of a panel survey are experienced in the subsequent waves where the individuals or households identified in the first wave are tracked down (Phiri and Tangamodzi 2016).

One of the most recent studies which used NIDS data analysed the determinants of South African households moving in and out of poverty between 2008 and 2014/15 (Finn and Leibbrandt 2017). They analysed absolute poverty transitions using a balanced panel for the first four waves of NIDS. They found that wave two and three households had higher chances of exiting poverty in waves three and four than was the case for households in wave one and two. They also found that close to 50% of households were below the poverty line in all the four waves and 54% of the households that were part of the balanced panel were poor in both wave one and four, with more than 50% of these households experiencing severe poverty.

Von Fintel and Zoch (2015) used the first three waves of NIDS to analyse the child poverty dynamics in South Africa. Whereas, Finn and Leibbrandt (2017) used the absolute poverty lines to analyse poverty dynamics, von Fintel and Zoch (2015) used an asset-based poverty line to identify structural poverty. They found that between 2008 and 2012, child poverty had decreased by more than half in their sample. The availability of assets in a household also showed significance in the ability of children to exit the poverty trap, they found that 40% of the children who were in a poverty trap were from households with poor asset endowment (von Fintel and Zoch 2015). Finn et al. (2013) used the 2008 and 2010 NIDS to analyse the transitions of households in and out of poverty during this period. Their findings show that an estimated 75% of households who were below the poverty line in 2008 remained poor in 2010, but they also found that fewer households entered into poverty when compared to those who managed to exit. Existing evidence using both the KIDS and NIDS shows that South Africa experiences chronic and transitory poverty.

Another less mentioned panel survey conducted in the post-1994 period is the Birth to Twenty-cohort study (BT20). Restricted to the Soweto township in the Gauteng province of South Africa, the BT20 sample was limited to children who made the sample in Soweto with a total sample of 3 273 children in 1989. The study was initially named a Birth to Ten (BT10) cohort study and it followed birth cohorts for ten years, before the study was extended to 20 years (Richter et al. 2007). The BT10-BT20 followed subpopulations of children born in a specific time period and residing within specified confines (Richter et al. 2004). Therefore, the sample of a cohort study is restricted to only incorporate specified parameters; in this case birth related specifications whilst a panel such as NIDS selects a sample that is representative of a broader population. The BT20 therefore can only be used to make inferences to individuals born within the time frame for the BT20 and residing within the residential confines specified. On the other hand, NIDS allows for poverty analysis for the whole country since its sample is nationally representative.

The BT20 collected data of a broad spectrum, but of importance in this study is the study's contribution to poverty since it collected information on the well-being of the respondents. Despite the fact that it was a relatively small cohort-based panel, the BT20 contributed to knowledge of poverty dynamics in the country in the post-1994 period. BT20 also demonstrated the potential contribution that can be made by cohort data.

However, the BT20 panel was not specifically designed to analyse poverty dynamics and as such does not have a comprehensive poverty module. Its major contribution perhaps marks an important step on the building of statistical capacity of panel data construction in the country.

1.5 Comparison of cross-sectional and panel data in poverty analysis

Poverty analyses have been conducted using cross sectional and panel data types. Davies (1994), Hsiao (2014) and Ruspini (2000) argue that panel data are capable of demonstrating the direction of causality. This is in contrast to the use of cross-sectional data that only allows for 'snap shots' of poverty at specific points to be analysed (Haughton and Khandker 2009:204). Panel data in this regard are therefore preferred and further allow for poverty to be categorized as transitory or chronic, an important classification that allows for policies to be designed more effectively for poverty reduction.

Recently, the case for using panel data in poverty analysis instead of cross-sectional types of data has been well documented (Baigrie and Eyal 2013). Adequate literature exists that shows that panel data are advantageous in poverty analysis. For instance, Hsiao (2005) states that panel data helps to deal with omitted variables and enables the analysis on the transition of households in and out of poverty over a given period. To be more specific, Baigrie and Eyal (2013) argue that using panel data allows for the observation of variables over time and this eliminates measurement error and omitted variable bias associated with cross sectional data.

The use of panel data in poverty measurement is also sometimes preferred because it reduces measurement error. Even though both cross sectional and panel data are affected by measurement error especially when using income or expenditure data, panel data allows for measurement error to be isolated and the actual changes in income or expenditure over time to be analysed (Kamanou et al. 2005). Moreover, May et al. (2000) also argue that panel data is sometimes preferred in poverty analysis since it allows for unobserved time invariant characteristics of households to be controlled, thus minimizing bias.

However, the availability of such data in the developing world is scarce, which is also true of South Africa. Martinez et al. (2013) argues that one of the reasons that explains

the scarcity of panel data is that it is more expensive to collect when compared to cross sectional survey. Also, Scott and Marrioti (2014) argue that in the post 2015 era, the need for panel data in poverty analysis has significantly increased. Yet this data gap is still to be addressed. Instead, the analysis of poverty trends emanating from cross sectional surveys remain dominant and may show an increase or reduction of poverty, but it largely remains unknown which households transitioned in or out of poverty for the period under analysis.

Despite the fact that panel data are sometimes preferred in the analysis of poverty, panel data have their own challenges encountered in poverty analysis. The most cited problems of panel data include problems arising from sample attrition and panel conditioning. In South Africa, on top of such widely observed limitations are other issues including limited coverage in the case of KIDS and small sample problems associated with the NIDS.

One of the effects of attrition is that ultimately the validity of the estimates made from the panel is affected as attriting households may introduce bias into the analysis (Baigrie and Eyal 2013; Baulch and Quisimbing 2009). Attrition, for instance, attrition may introduce bias as the missing households may be representative of the poorest households in the sample (Roberts 2000). The panel in that case is likely to underreport poverty and if policy is informed by such data, it becomes less effective as the targeting is misinformed.

In another study, Roberts (2000) argues that this attrition may introduce bias as the missing households may be representative of the poorest households in the sample. To try and understand the causes of attrition, Lee (2003) argues that there are three main causes; migration, mortality, and non-response. Migration is specifically a problem in developing countries as the processes of industrialization often result in high mobility of households (Lee 2003).

Narrowing down the problems associated with attrition to South Africa, attrition in both KIDS and NIDS further limit the analysis of poverty over time. For example, attrition in KIDS is well documented. As presented earlier, there are three waves of KIDS that have been conducted over the years; in 1993, 1998 and 2004. In 1998, KIDS re-interviewed 1178 households, which constituted 85% of the PSLSD sample, 41 additional split off households were also followed up in 1998 to bring the sample size

to 1219 (Roberts 2000). May et al. (1999) found that in explaining attrition, the mobility of households was part of the reason why some households could not be reinterviewed in 1998. On a positive note, the 1998 wave of KIDS experienced fewer refusals with only an 11% refusal rate, which May et al. (1999) argues is in line with the general trend in developing countries. The 2004 KIDS managed to interview 62% of the original households interviewed in the PSLSD in 1993, an attrition rate that hints on the possibility of bias (Aguero et al. 2007). What is clear is that the 2004 wave of KIDS was less representative of the original sample used in the first wave. And this is expected in panel data surveys, because households break away and new ones are created. It then becomes increasingly difficult to retain the same households identified in the first wave, thereby introducing bias into data generated from later waves.

In attempting to understand the attrition rates of the KIDS 1998, Roberts (2000) argues that households that were 'lost' was partly a result of the household roster design that only captured first names making future identification a challenge. But there is no evidence that supports this design fault as explaining the 15% attrition rate. Instead, Roberts (2000) found that households that moved did not leave forwarding addresses. Lastly, Roberts (2000) found cheating to have occurred during the original 1993 PSLSD survey as some households were non-existent as part of the factors that explain the attrition experienced in the 1998 KIDS. Finn and Ranchold (2013) also found cheating amongst fieldworkers working with NIDS. Close to 7% of the sample was affected and fieldworkers were fabricating interviews, and this compounds bias within surveys

Von Fintel and Zoch (2015) using the first three waves of NIDS found higher attrition rates to exist among the white population with an attrition rate of 53% compared to the black population group with 13.39%. Their findings show that close to 40% of the children in their sample remained in structural poverty between 2008 and 2012. They also identified low education and low asset holding as some of the characteristics associated with households experiencing poverty traps.

Parker et al. (2014); Hsiao (2007); Houghton and Khandker (2009) state that panel surveys are more expensive to administer when compared to cross sectional surveys. Due to the sensitivity of information relating to the actual costs associated with the running of official surveys, this research fails to provide the different estimates

associated with the conduct of panel surveys when compared to cross sectional data in South Africa.

Attrition is perhaps one of the most cited weaknesses of panel data in poverty literature. Haughton and Khandker (2009) argues that panel data are affected by attrition bias as households pull out of the panel resulting in those households becoming less representative over time of the population they are supposed to represent. This is actually one of the weaknesses associated with panel data; it becomes less representative of the population as the design does not account for new households being formed or reformed (Haughton and Khandker 2009). Parker et al. (2014) argue that one of the limitations of using panel data is that it poses the risk of increasing survey fatigue which may result in some of the households dropping out. Long and complex questionnaires on cross sectional surveys as well may introduce fatigue to respondents.

Cross sectional data on the other hand are not altogether always a 'second best' option. For instance, the availability of cross-sectional surveys widely in developing countries brings convenience in poverty analyses, as they are more readily available than panel data in developing countries. Part of this is a result of the relatively lower costs associated with the conduct of cross-sectional data when compared to panel data. Moreover, compared to panel surveys, cross sectional surveys allow for large samples to be used. This use of large samples enables generalizability to small spatial units, allowing for the analysis of characteristics and distribution of poor households.

Cross sectional survey data in poverty analysis also allows for the measurement of poverty trends, an attribute that is evident in existing national and global poverty statistics distributed by national statistical offices and shared broadly. Visser et al. (2000) also attributes this type of data as enabling the analysis of relationships between variables and can also be used to analyse the differences between sub-groups.

However, the major drawback when using cross sectional data in poverty analysis is that such data does not control for the distinction between transient and chronic poverty (Kamanou et al. 2005). Wijesekere (2009) further discredits cross sectional data arguing that it does not give meaningful information on the dynamic aspects of the population under scrutiny; and this type of data also fails to precisely show causality (Davies 1994). Therefore, when analysing poverty using this type of data poor

households cannot be analysed beyond the fact that they fall below a given poverty line at the time when the survey was conducted. Despite this inherent weakness in cross sectional surveys, they are the most commonly administered type of survey in most developing countries.

May et al. (2000) state that the use of cross sectional data fails to address crucial questions on poverty dynamics; cross sectional data manages to capture whether poverty increased, decreased or remained the same but it does not follow the same households over time and this missing dynamics may not be effective in informing poverty reduction policy.

The use of different types of data in poverty analysis complements our understanding of poverty and resultantly each type is important. Visser et al. (2000) also suggest that to maximize on the strengths of cross sectional and panel data, both forms can be combined by administering the same questionnaire as a panel and an independent cross section.

The methodology proposed in this study will not necessarily be affected by attrition, as it does not follow the same households over the years. Instead it uses birth years of household heads to identify similar household and these are refreshed with each additional year added to the panel.

1.6 Statement of the Problem

Orthodox survey analysis methods are inadequate in keeping up with the increasing data demands in poverty measurement. A majority of poverty assessments depend heavily on data that has already been collected. The limitation of this is that such data are not necessarily collected to suit the different research questions specific to poverty analysis. As such the data are not necessarily customized to answer specific poverty related query and at the same time the data still needs to remain comparable over time. Therefore, there is need to go beyond just analysing available household survey data, to creating other additional data sources.

As much as South Africa can be categorized as a 'data rich' country, it has a data gap as far as nationally representative panel data are concerned. In South Africa, NIDS remains the only panel with a sample that is nationally representative that is related to poverty measurement and as such leaves more room for research.

Extensive research has been conducted using the KIDS and NIDS in the analysis of poverty dynamics in South Africa. Despite the existence of such studies, this thesis argues that there is a paucity of panel data in South Africa and leaves room for alternative data methods. Synthetic panel data, also termed pseudo panels, are one such alternative that can be used to pursue poverty related research questions that would otherwise be impossible with existing survey data.

Pseudo panels provide longitudinal analysis of poverty without incurring costs associated with genuine panel data. The use of pseudo panels in econometric analysis is rapidly expanding since its inception in the early 1980s with the work of Deaton (1985). Pseudo panels are sometimes preferred because they are less affected with non-attrition (Antman and McKenzie 2007; and Moreno 2015; Cruces et al. 2013; Dang and Lanjouw 2013; Deaton 1985). Panel surveys apart from being costly are not available for longer time periods, a limitation that is easily overcome by pseudo panels (Antman and McKenzie 2007; Bourguignon and Moreno 2015; Cruces et al. 2013; Dang and Lanjouw 2013). In line with the arguments above, this study uses the Living Conditions Surveys of 2008 and 2014 as well as the Income and Expenditure Survey (IES) of 2010, to examine the feasibility of pseudo-panels in the South African context and to demonstrate their potential contribution to poverty analysis.

1.7 Objectives and Research Questions

The research interest of this thesis is twofold. The first seeks to investigate actions and intentions that have been taken by Stats SA to improve poverty measurement in South Africa. To this end a qualitative analysis was conducted using data collected using semi-structured interviews. To extend this analysis, existing data related challenges in poverty analysis in South Africa are also reviewed. The analysis discusses how Stats SA has captured poverty estimates in the country as an attempt to understand better the current gaps in poverty measurement. Under this broad objective the analysis seeks to also illuminate how official statistics relating to poverty analysis in South Africa has been collected and then analysed and illuminate the gaps that have been identified by Stats SA that still require attention.

Based on the constraints that limit Stats SA's options, the second research interest seeks to promote the use of alternative data sources in poverty analysis to complement existing poverty studies. In this regard, a pseudo panel data is created using data from

cross-sectional household surveys in South Africa. This will enable the investigation of poverty and expenditure patterns of birth cohorts between 2008 and 2014. This study is a contribution to the growing field of studies that advocates for the use of simulated data in poverty analysis. The paucity of panel data in poverty analysis is acknowledged to exist in most countries in Sub Saharan Africa. In a bid to overcome such inadequacies, there has been a growing use of pseudo panel data in poverty analysis. But this has not extended fully to African countries and in South Africa more specifically. The creation of pseudo panels in poverty analysis is cost effective and is a viable alternative data source in the absence of actual panel data. The generation of a pseudo panel in this study seeks to complement existing poverty analysis that uses panel data in South Africa.

The broad objectives in this study are then:

i. To review how poverty trends can and have been measured in South Africa?
ii. To examine factors and issues surrounding the frequency, volatility, comparability and reliability of data sources in South African surveys related to this study.

iii. To investigate the robustness of poverty trends and characteristics in South Africa using pseudo panel data

iv. To examine the linkages that exists between poverty and gender in South Africa using pseudo panel data.

Specifically, the study aims to address the following research questions:

1. What are the poverty trends in South Africa and how have these been measured over the years?

2. How can the frequency, comparability, consistency, reliability and accuracy of data available in poverty analysis in South Africa be addressed and improved?

3. Can the use of pseudo panels in poverty analysis be an effective alternative poverty measurement tool in South Africa?

4. Can pseudo panel data be utilized in exploring the poverty and gender nexus in South Africa?

5. What are the gender differences in the risk of poverty between Female Headed Households (FHHs) and Male Headed Household (MHHs) in South Africa that can be identified using pseudo panel data?

1.8 Research Design

The choice of a research design is influenced by the research objectives and questions that are key in the study. This factor was crucial in the choice of the appropriate research design in this study. Creswell (2014:21) defines a research design as 'the procedures of inquiry'. In another study, de Vaus (2001:9) provides a rather more elaborate definition and defines research design as the structure of research before data collection and analysis. Its function is essentially '*to ensure that the evidence obtained enables us to answer the initial question as unambiguously as possible'* (De Vaus 2001:9). The following subheadings will explore the different concepts involved in the research design that were used in this study.

1.8.1 Research Methodology

This study will use mixed methods, that is both qualitative and quantitative analysis. The motivation in using this methodology in this study emanates from the research questions in this study, which are both quantitative and qualitative in nature. Part of this complexity is explained by Morse and Niehaus (2009:15) who argue that,

'*…the complexities of the phenomena studies in social science, and the limitations within methods means that there are occasions when a phenomenon cannot be described in its entirety using a single method'*.

The thesis investigates the perceptions on the statistical preparedness of South Africa in providing frequent, accurate and comparable household surveys in poverty measurement. In this section, the discussion explores at length Stats SA's efforts over the years to measure poverty. The semi-structured interviews explore further the gaps that exist in poverty measurement and justify the need for more analysis of poverty using existing survey data in the country. Purposive sampling is used in the selection of respondents that were interviewed. Patton (2002) defines purposive sampling as a sampling method that identifies and selects 'information-rich' cases for the most effective use of available resources. Creswell and Plano Clark (2011) further elaborate that purposive sampling identifies and selected individuals who have specific

knowledge and experience in a subject under study. The interviews that will be conducted in this study identify individuals within Stats SA who have immense knowledge and experience in dealing with surveys related to this study. Ten in-depth interviews were conducted with data producers in Stats SA to provide meaningful insights on the different issues related to poverty related data. The main instrument for data collection for the qualitative analysis was an interview guide

In the quantitative component, the study seeks to create pseudo panel data using South African household data to investigate the linkages that exist between poverty and gender using quantitative methods. The quantitative part of this analysis makes use of a pseudo panel created using data from the LCS and IES surveys. This will allow for the analysis in the movements in and out of poverty of birth cohorts in South Africa to be investigated from an alternative data source. The quantitative part of this study was done in two stages. The first level uses pseudo panel data to analyse expenditure patterns of households in the pseudo panel before descriptive statistics are applied to analyse poverty dynamics. STATA 14 will be used to conduct a quantitative analysis using multivariate analyses and descriptive statistics to explore the relevant research questions.

1.9 Structure of the study

Chapter One is an introduction to the study and discusses the importance and significance of this research as well as providing the background to the study. This chapter also gives the methodology and research objectives and questions that are central in this study. The chapter ends by providing an outline of the research.

Chapter Two is a literature review on the use of data imputation methods applied to cross sectional survey data and panel data alike in the absence of adequate household survey data required for poverty analysis. This chapter explores the data inadequacies and gaps that are encountered in poverty analysis and conducts an in-depth discussion on the relevant survey-to-survey imputation method used to date. Pseudo panel methodologies are also discussed as part of the methods used to provide alternative panel data sources where they are inadequate or unavailable. One of the broad research interests in this chapter is to acknowledge that poverty data are inadequate, and this is partly a result of the fact that poverty is multidimensional and as such it is difficult to have adequate household survey data that captures its diverse forms comprehensively.

This justifies the need for imputations to maximize existing survey data without incurring further costs on already strained budgets.

Chapter Three discusses the theoretical frameworks in poverty measurement over time. The chapter pursues at length the use of poverty lines in poverty measurement as well as the importance of chronic and structural poverty. Of significance in this are the policy implications in understanding the distinctions between transient, chronic and structural poverty.

Chapter Four is a qualitative analysis chapter that discusses the perceptions of data producers on issues related to current poverty related surveys from South Africa. The chapter explores official poverty statistics as well as the challenges faced in poverty measurement by Stats SA.

Chapter five introduces the quantitative research component of this thesis. The chapter discusses the decisions made in the generation of the pseudo panel. The methodology and assumptions used in the generation of the pseudo panel are also presented. The chapter ends by a descriptive analysis of the birth cohorts used in the pseudo panel.

Chapter six then uses the pseudo panel to analyse the expenditure patterns and poverty incidences of the birth cohorts in the pseudo panel. The purpose is to test the usefulness of the pseudo-panel for the analysis of poverty dynamics.

Chapter seven is a conclusion of the thesis. It sums up the investigation of the numbers used in poverty analysis in South Africa. The chapter also provides relevant conclusions in the use of pseudo panels in poverty analysis in South Africa. The last section of the concluding chapter presents the recommendations emanating from this study.

Chapter Two:

Analyzing poverty in the absence of appropriate survey data: A literature review of alternative data approaches

2.1 Introduction

This chapter explores data gaps that exist in poverty analysis that result from a general unavailability of adequate survey data in poverty analysis whether in cross sectional or in longitudinal form. The chapter explores pseudo panel approaches and Survey-to-Survey Imputation methods (SSI) as possible alternatives in leveraging cross sectional data for the analysis of poverty dynamics and trends. The objective of the chapter is to promote the use of alternative data sources in the analysis of poverty. The term alternative data source is used to refer to data that are generated by manipulating existing household surveys to meet specified research enquiry. Such data manipulation approaches have been applied in the generation of poverty maps and extend to the construction of pseudo panels. An alternative data source in this study can be differentiated from orthodox household surveys (cross sectional or in panel data form), which in this study are considered as the 'traditional forms' of data used in poverty analysis.

There is a growing body of work that focuses on the development of alternative data sources as an attempt to deal with existing data gaps in poverty analysis. This literature review chapter discusses data gaps that exist in the analysis of poverty trends and dynamics. After the identification of these gaps, a discussion on pseudo panel methods for the analysis of poverty dynamics and SSI approaches used in creating simulated cross-sectional data for the analysis of poverty trends is pursued at length.

The assertion that Millennium Development Goals (MDGs) era had more 'data gaps than observations' (Jerven (2017:46), illuminated the mismatch between data available and data required to monitor development progress in developing countries. This leaves a lot of doubt on the ability of developing countries to monitor current poverty reduction development targets. Therefore, in light of an increasing demand for quality and comparable data, this thesis argues for the creation of alternative sources of data to accommodate the analysis of new research questions on poverty using existing household surveys.

2.2 Household survey availability for poverty measurement

There has been a growing interest on the type and quality of data that is used in poverty analysis in the post-1990 period. Part of this has been a direct result of the pursuit of global targets such as the MDGs (1990-2015) and the recent Sustainable Development Goals (SDGs) (2016-2030) that have prioritized poverty reduction efforts. This has resulted in a growing demand for additional data sources to effectively monitor existing targets and a growing scrutiny on the quality of data being produced.

In response to the rising demand of development statistics indicated above, evidence shows an expansion of survey data sources in the post-1990 period in developing countries (Alkire 2014; Chen and Ravallion 2010; Demombynes and Sandefur 2014; Gibson 2016). To explore this expansion in perspective, poverty measurement efforts significantly grew in the last 20 years (Gibson 2016). Chen and Ravallion (2010) argue that the improvement in the technologies used universally in the construction of national poverty lines, as well as in representativeness of household surveys made this process easier. To support this increase in poverty measurement efforts globally, Beegle et al. (2016) and Serajuddin et al. (2015) found an increase in the production of comparable surveys in the same period.

Using poverty survey data from 155 countries and territories monitored with the World Bank, an increase was observed in the number of countries with consumption data from 13 countries in the 1990s to 40 in 2001 and 62 in 2011 (Serajuddin et al. 2015:2-3). Moreover, between 2000 and 2005 the number of countries with survey data available for poverty analysis increased from 37 to 60 countries (Serajuddin et al. 2015:5), with data points increasing in Sub Saharan Africa from 5.1 data points to 7.5 between the early 1990's and 2000s. Table 2.1 illustrates this increase in data points of poverty related data in Sub Saharan Africa in relation to the rest of the globe. The table highlights two factors. First, it shows the total number of surveys available per region, labelled as the total number of surveys in the table below. Secondly the table also shows the proportion of poverty specific data in relation to the total number of surveys conducted in every region represented below.

Table 2.1: Survey availability for poverty analysis between 2002-2011

Region	Poverty data points	Total number of surveys
Sub Saharan Africa	1.7	16.8
South Asia	2.4	28.8
Middle East and North Africa	1.9	11.5
Latin America and Caribbean	5.2	19.5
Europe and Central Asia	7.7	21.7
East Asia and Pacific	2.1	16.0

Source: Serajuddin et al. 2015:19

Sub Saharan Africa has the least poverty data points when compared to other regions. Despite this, East Asia and the Pacific as well as the Middle East and North Africa have fewer total number of surveys per year when compared to Sub Saharan Africa. The existence of the least number of poverty specific surveys in Sub Saharan Africa is rather a contradictory outcome - as the region in 2015 was home to more than 50% of the global poor (World Bank 2018b).

Evidence from recent studies also shows that the collection and use of panel data have expanded recently in developing countries (Hsiao, 2007). In a more recent study, Dang and Dabalen (2017) also confirm this expansion of panel data, stating that there has been an improvement in the availability of panel data in developing countries in the last 10 years. However, an analysis of existing panels within this region shows that a majority are sub-national panels and offers limited analyses on poverty. To illustrate the existence of such data, Baulch (2011) found that in Africa sub-national panels have been successfully conducted in Egypt, Ethiopia, Ghana, Kenya, Lesotho, Madagascar, Malawi, South Africa, Tanzania, Uganda, Zambia and Zimbabwe. Sub-national panels are an indication of the inability to conduct bigger panel surveys that would capture bigger samples reflective of the entire country's population. The foregoing is partly a result of stringent budget allocated to statistics in developing countries. Baulch (n.d.) in World Bank Workshop also found that panels in developing countries have shorter time periods with a majority having only two time periods. On the other hand, developed countries have panels that stretch for longer periods.

Table 2.2 below summarizes some of the panel surveys conducted in developing countries. The table attempts to provide an updated list of countries originally presented

in Baulch (2011). Despite the fact that it was difficult to get a comprehensive list of all panel surveys conducted in these countries, the available data from public domains indicates an expansion of panel data between 2012 and 2018.

Table 2.2: Panel survey conduct in developing countries

Country	Survey name	Type	Years of availability
Argentina	Argentina Enterprise Survey	Panel (National coverage)	2006, 2010, 2017
	Permanent Household Survey	Rotating panel	1990-2002, 2002-2018
Bangladesh	Bangladesh Enterprise Survey	Panel	2006, 2010, 2013
	Livelihoods Systems in Bangladesh	Panel	1987, 2000, 2004,2008
	Greeley Survey	Panel	1994-2001
	Livelihoods Systems in Bangladesh	Panel	1987, 2000, 2004, 2008
	BIDS 64 Village panel	Panel	1987/88,2000
Bolivia	Bolivia Enterprise Survey	Panel	2006, 2010, 2017
	Integrated Child Development Project Evaluation Data (Extension of the Comprehensive Development Project (PIDI)	Panel	1995/96,1998, 2003, 2004
Brazil	Enterprise Survey	Panel	2009
	Longitudinal Survey data of households	Panel	1996-2009, (Annual)
	Sustainable Agriculture in the Amazon Study	Panel	1996, 2000, 2005
	Bolsa Family evaluation	Panel	2005, 2009
Chile	Enterprise Survey	Panel	2006, 2010
	National Socioeconomic Characterization	Panel	1996, 2001,2006,2013, 2015/16
China	China Family Panel Studies	Panel	2010, 2011, 2012, 2014
Egypt	Enterprise Survey	Panel	2010
	Panel of Young People of Egypt	Panel	2009, 2014
	Egypt Integrated Household Survey	Panel	1997, 1999
Ethiopia	Ethiopia Socioeconomic Survey	Panel	2011/12, 2013/14, 2015/16
	Rural Household Survey	Panel	1989-2004
Hungary	Enterprise Survey	Panel	2002, 2005, 2009, 2013
	Hungarian Household Panel	Panel	1992-1997
Ghana	Socioeconomic Panel Survey	Panel	2009/10, 2014

	Urban Household Panel Survey	Panel	2004, 2005, 2006
	Rural Household Survey	Panel	1994-2004
	Agricultural Innovation and Resource	Panel	6 weekly intervals
Guatemala	Enterprise Survey	Panel	2017
	INCAP	Panel	
India	India Human Development Survey	Panel	2004/5, 2011/12
	National Council of Applied Economic Research panel	Panel	1970/71, 1981/82, 1998/99
	ICRISAT	Panel	1968/69, 1970
	ICRISAT New	Panel	2001-2006
	Rural Economic Development Survey (REDS)	Panel	1968/9,1981/82, 1998/9
	Enterprise Survey	Panel	2002,2005, 2014
Indonesia	Indonesia Family Life Surveys	Panel	1993/4, 1997/8, 2000, 2007, 2014/15
	100 Village Survey	Panel	1998, 1999
Iran	Data from Statistical Centre of Iran	Panel	1992, 1993, 1994, 1995
Kenya	Enterprise Survey	Panel	2007, 2013/14, 2018
	Mpesa Panel Survey	Panel	2008, 2008/9, 2010, 2011, 2014
			2007, 2009, 2011
	The Transfer Project	Panel	
	Tegemeo Rural Household Panel	Panel	
	Quarterly Survey of 6 cities		
	USAID Panel		
	KICS		
	STEP Skills Measurement Household Survey		
Lao	Lao Expenditure and Consumption Survey	Panel	
	STEP Skills Measurement Household Survey	Panel	
Lesotho	Enterprise Survey	Panel	2009-16
Madagascar	1-2-3 survey		1997/98, 1998/99
Malawi	Integrated Household Panel Survey	Panel	2010, 2013, 2016
	Enterprise Survey	Panel	2009-2014

	Complementary Panel Survey	Panel	2000, 2001
Mali	Enterprise Survey	Panel	2007, 2010, 2016
Mexico	Mexico Family Life Survey	Panel	2002, 2005
	Mexico Panel Survey	Panel	1997, 2000, 2006, 2012
Nepal	LSMS	Panel	1995/96,2003/04, 2010.11
Nicaragua	LSMS	Panel	1998,2001,2005
Pakistan	IFPRI Household Food Security Panel	Panel	1986-1990/91(Annual), 2002/3, 2005/6
	North West Frontier Province	Panel	1995/6,1998/9
	Pakistan Rural Panel Household Survey	Panel	1990/91, 2004, 2010
Peru	National Household Survey	Panel	1997-1999
	Enterprise Survey	Panel	2006,2010,2017
Phillipines	IFPRI and RIMCU Bukidnon panel	Panel	1984/5 (4 rounds), 1992, 2004
	FIES/APIS	Panel	1997,1998,1999
Poland	Household Budget Survey	Panel	1993-1996
	Polish Panel Survey	Panel	1988-2013(7 waves)
Rwanda	National Budget Consumption Survey	Panel	1982-1983 (Quarterly)
Russia	Russian Longitudinal; Monitoring Survey	Panel	1992-2019 (Annual) 27 waves
South Africa	KwaZulu Natal Dynamic Study	Panel	1993, 1998, 2003
	National Income Dynamic Study	Panel	2008, 2010, 2012, 2014, 2016, 2018
	Community Agency for Social Enquiry study of urban informal settlements	Panel	1997, 2001
	USAID survey	Panel	May/June 2001, Nov/Dec2001, July/Aug 2002
Sri Lanka	Uda Walawe Irrigation Area Impact Assessment	Panel	2001-2002
Tajikistan	Tajikistan Living Standards Measurement Survey	Panel	2007, 2009
Tanzania	Tanzania National Panel Survey	Panel	2008/9, 2010/11, 2012/13, 2014/15
	Enterprise Survey	Panel	
	Kagera Health and Development Survey	Panel	2006-2013
	Tanzania Urban Household Survey 2005	Panel	1991-2004

Uganda	Uganda National Panel Survey	Panel	2005, 2010, 2013, 2014
Venezuela	Household Sample Survey	Rotating Panel	1997,1998
Vietnam	Vietnam Living Standards Survey	Panel	1992/93-1997/98
	Vietnam Household Living Stands Survey	Rotating panel	2002,2004,2006,2008
Zambia	Rural Agricultural Livelihoods Survey	Panel	2012, 2015
	Enterprise Survey	Panel	2007, 2013
Zimbabwe	Manufacturing Firm Survey	Panel	1993, 2015/16
	Rural Household Dynamics Study	Panel	1986/87, 2015/16

(Sources: Baulch 2011, the Bureau for Research and Economic Analysis of Development (BREAD); International Household Survey Network (IHSN), World Bank Central Data Catalog).

The update of the panel surveys in the table above was dependent on the availability of data in the public domain and therefore remains incomprehensive. The table however shows a continuation of some panel surveys after 2011 as well as the introduction of new panel surveys in other countries.

2.3 Data inadequacy in poverty measurement

Notwithstanding the progress recorded in the availability of cross sectional and panel surveys noted above, data inadequacies continue to exist for both cross sectional and panel data. At cross sectional level, the paucity of comparable and adequate household survey data for the analysis of poverty trends is widely acknowledged in poverty literature. On the other hand, panel data are scarcer when compared to cross sectional surveys and where they exist, they often have small samples. This section discusses inadequate poverty related survey data in developing countries.

Data inadequacies in developing countries are well documented (Beegle et al. 2016; Dabalen et al. 2014; Douidich et al. 2013; Grosse et al. 2009; Hentshel et al. 2000; Jevern 2017). Jerven (2017) conducted a comparative analysis on survey availability for all African countries and found that surveys used in the collection of poverty statistics might be fewer than anticipated. Jerven (2017:34) states that from 48 African countries, 27 had at least 2 data points to analyse poverty, but 21 countries were found to have only one household surveys and or no household survey data at all. This means there is no way of knowing how the 21 countries actually fared in their poverty reduction efforts during the MDG era.

In another study, Beegle et al. (2016) compared surveys availability between developed and developing countries and found that the latter group of countries only managed to conduct half the number of consumption surveys conducted in developed countries. In fact, the data inadequacy is worse in some scenarios. For instance, Jerven (2014c:5) found that 6 from a total of 49 countries in Sub-Saharan Africa have never conducted a household survey between 1990 and 2015. There is therefore no way of knowing whether poverty increased or decreased in these countries.

To explore part of this debate further, Grosse et al. (2009) argue that many developing countries do not have adequate time series data that is nationally representative, and this makes the study of poverty trends an impossible undertaking. Further arguments in this body of literature are postulated by Dabalen et al. (2014), Douidich et al. (2013) and Grosse et al. (2009) who argue that the available data sources in most developing countries on poverty estimation are infrequent, inaccurate and inadequate for poverty monitoring.

To further illustrate the presence of inadequate data in developing countries, Gibson (2016) argues that despite the increasing availability of household survey data, we know less about poverty today than we used to previously. This is partly due to the use of poor survey design which makes the analysis of food consumption and costs of living difficult to gauge accurately (Gibson 2016). In a similar argument, Ravallion and Chen (1997) argue that despite the recorded improvements in the availability of household surveys in developing countries, these are limited by the availability of conceptual and practical challenges. Therefore, the increase in household surveys that collect information on consumption and expenditure has been useful in our understanding of poverty trends in the post-1990 period but has also brought into the spotlight issues of data quality.

Panel data is less available than cross sectional data in poverty analysis in developing countries (Dang et al. 2018). The observed improvement in survey availability also seems to be skewed towards middle to high-income developing countries resulting in a limitation on the availability of survey data in lower income developing countries (Harttgen 2017).

Apart from a general lack of appropriate cross sectional data, nationally representative panel data are scarce in the developing world and this makes the analysis of poverty

dynamics challenging (Bourguignon et al. 2004:2; Bourguignon and Moreno 2015; Cruces et al.2013; Dang et al. 2014; Douidich et al. 2013; Urzainqui 2017; Verbeek 1993, 2008). This is contrary to the fact that developing countries experience higher levels of poverty when compared to developed countries and would benefit from nationally representative panel data. But perhaps the fact that they are still battling with poverty eradication agendas partly explains their failure to conduct adequate surveys for effective poverty monitoring. Due to stringent budgets, their allocation of funds for statistical purposes remains limited.

Dang and Dabalen (2017) and the World Bank (2017) show that the coverage of panel data in African countries remains limited. Dang and Dabalen (2017) found that panel data with a national coverage existed only in 7 countries in Sub Saharan Africa and where it was available, it was collected for relatively short periods. In a more detailed analysis on panel data availability, Yaqub (2000) found that from a total of 44 developing countries only five had panel data, with a majority of data available allowing cross sectional poverty analysis. The conduct of panel data has improved since the work of Yaqub (2000), but the absence of panel data in earlier periods also leaves gaps in our understanding of changes in poverty dynamics.

Of importance as well is the fact that the expansion of panel data discussed earlier is mostly in sub-national panels. For instance, the recorded increase in the number of panel surveys conducted in Egypt, Ethiopia, Ghana, Kenya, Lesotho, Madagascar, Malawi, South Africa, Tanzania, Uganda, Zambia and Zimbabwe discussed earlier in Baulch (2011) was in sub-national panels. Due to the small sample sizes of sub national panels and their availability in limited time periods, there is more panel data forms that are required to fully understand poverty dynamics at country level in developing countries. Therefore, a gap remains in panel data relevant for the analysis of poverty dynamics in developing countries.

In South Africa, data inadequacy is less evident, given the availability of a commendable number of cross-sectional household surveys that have been used in the analysis of poverty trends. However, panel data remain less adequate, a data gap that will be partly addressed by a pseudo panel in this thesis. The analysis of poverty dynamics in South Africa is limited to the availability of few available panel surveys such as the KIDS data, which only allows for the analysis of poverty dynamics in 1993,

1998, and 2004 in KwaZulu-Natal province, and to the five waves of NIDS (2008, 2010, 2012, 2014, 2016). The two panel surveys have been criticized on a number of levels. The KIDS for instance can be criticized on its coverage of only one province in South Africa, limiting the analysis of poverty dynamics only to KwaZulu-Natal. Little is therefore known on the poverty dynamics experienced in the post-independence period in South Africa from 1994 to 2007 prior to the introduction of NIDS. This translates to missing information on poverty dynamics for more than a decade. Therefore, despite the pro-poor government policies that have been passed in South Africa during this period, there is uncertainty on the actual movements of households in and out of poverty. However, cross sectional surveys in South Africa during 1994 and 2007 has shown that the country was battling with high levels of poverty and poverty reduction where it was achieved was marginal during this period.

2.4 Analysis of poverty trends with incomplete cross-sectional data: A discussion of Survey-to-Survey Imputation approaches

Data inadequacies in poverty analysis discussed earlier in this chapter are an indication of the increased importance of using alternative data sources in poverty measurement. What is clear is that traditional forms of household survey data are sometimes either inadequate or imperfect to adequately address poverty related enquiry. SSI approaches have been defined by Cuesta (2017) as the use of consumption/income household surveys to make imputations of consumption/income in another survey such as a labour force survey. But it should be noted that the creation of alternative data sources using Survey-to-Survey Imputation (SSI) methods, will not necessarily provide a solution to all these problems but rather provides alternative data sources and ways for analysing new research questions in poverty analysis.

This use of alternative data sources in poverty analysis allows for the formulation of new research questions that are relevant to today's poverty reduction agenda. An analysis of poverty using existing traditional data sources to date in developing countries has been useful in the understanding of poverty trends and dynamics but has left gaps especially where such surveys are absent, less frequent, incomparable or too narrow. Therefore, alternative data sources provide the possibility of dealing with some of these data challenges faced in poverty analysis in developing countries.

The use of SSI methods has become widespread and has been used many times in poverty analysis (Douidich et al. 2013; Elbers et al. 2002, 2003, 2005; Stifel and Christiansen 2007). Originally, SSI methods were designed to work with missing data but have been adopted to suit different types of missing data problems over the years (Douidich et al. 2013). The SSI approaches used in creating alternative data sources have also evolved over the years.

Moreover. SSI approaches have been applauded for their cost effectiveness in poverty analysis and their ability to create frequent and timely poverty estimates from surveys (Newhouse et al. 2014). Their ability to create comparable time series data, and comparable expenditure/consumption data for poverty analysis have enabled new dimensions of poverty to be analysed, which would not have been possible using the original data forms. But the use of SSI approaches to leverage existing survey data in poverty analysis is not without criticisms and weaknesses. The weaknesses surrounding the use of SSI approaches are mainly rooted in the assumptions that have to be fulfilled to enable the imputations.

2.4.1 Asset Indices

Asset indices recently have been used as an SSI method. Conventionally, they are used as a proxy to welfare or wealth and are most frequently used in Demographic Health Surveys (DHS) since these surveys do not collect expenditure and income data but rather collect extensive information on assets. Johnston and Abreu (2014:2) define an asset index as 'any composite indicator such that the underlying indicators on which it is based reflect an individual's, more often a household's ownership (or lack thereof) of a range of assets'. Corsi et al. (2012) estimates that the DHSs have been conducted in more than 85 countries since 1984 and were designed to explore demographic, fertility and family planning in low and middle-income countries. Therefore, it is not surprising that traditionally, asset indices have been associated with DHS data since they do not collect income/consumption data. However, in the SSI sense, asset indices are generated using DHS data with other household survey data that have income/ consumption data to fill in data gaps where such income/consumption data are less comprehensive.

Filmer and Prichett (2001) estimated the relationship between household wealth and school enrolment of children in India using asset indices. In their method they

constructed a linear index using indicators of asset ownership. They were able to analyse the relationship between wealth and school enrolment without expenditure data and instead maximised on asset ownership. The results also showed that using assets instead of consumption expenditures as proxies for long-run wealth, the results have less measurement error. Sahn and Stifel (2000) in their application of asset indices analysed poverty in different time periods using DHS data. The welfare measures that they used were generated using a factor analysis comprising of household characteristics, durable goods, and education level of the household head. The study was successful in analysing intertemporal and interregional welfare using DHS data.

For example, in Liberia there is a general lack of income and consumption survey data for monetary poverty measurement. Dabalen et al. (2014) found that household surveys in the country have income/consumption modules that are inadequate to measure poverty aggregates as a result of the country's limited ability to conduct adequate surveys. In 2014, only two official poverty surveys existed in the country; the Core Welfare Indicator Survey (CWIS) of 2007 and the 2010 Demographic Health Survey (DHS). The CWIS of 2007 did not collect detailed consumption data required for the creation of aggregate poverty lines. Since there were no other poverty related surveys conducted during this period, Dabalen et al. (2014) created an asset index to fill in consumption data gaps to assess poverty trends using data from the DHS and the CWIS.

The SSI methodology used by Dabalen et al. (2014) used proxy poverty quintiles using factor analysis since the DHS does not have consumption information. A set of variables was chosen using factor analysis and a similar aggregate to the DHS was constructed using data from the CWIS survey. This methodology allowed the comparison of rankings of households using asset indices and consumption data. This SSI approach used asset indices with a consumption/expenditure household survey as a way of filling up gaps where such income/expenditure data were inadequate for poverty analysis. This approach requires strict data requirements that have to be fulfilled. For instance, the CWIS and the DHS had similar questionnaires (despite the consumption section) and their samples were drawn from the same population of households living in the same cluster to allow the assumptions used in the SSI to create robust results (Dabalen et al. 2014).

Despite the advances made in the analysis of welfare dynamics, there has been concerns on such application of asset indices approaches especially on their statistical validity (Merola and Baulch 2019). Merola and Baulch (2019) instead applies a more parsimonious and easier to interpret method using the categorical sparse principal components method.

2.4.2 Small area estimation methods

Small Area Estimation (SAE) has an early history, Brackstone (1987) traces its origins to as early as the 11th century in England. In its early application it was referred to as small area statistics and applied data derived from the population census or administrative data (Ghosh and Rao 1987:73). Ghosh and Rao (2003:55) apply the term 'small area' to refer to a 'local geographical area that is small or to describe any small subgroup of a population'. The problem that SAE seeks to answer is how to produce reliable estimates where only small samples are available. SAE methods have been applied in the analysis of poverty and recently existing approaches have used poverty maps and regressions and these applications are briefly discussed below.

2.4.2.1 Poverty Mapping

Poverty mapping is arguably one of the earliest forms of SAE approaches used in poverty measurement (Cuesta 2017). Poverty mapping produces non-conventional data for poverty analysis. Alderman et al. (2001:1) define a poverty map as 'spatial description of the distribution of poverty'. In another study, Hentshel et al. (2000:161) define a poverty map as 'a geographic profile of poverty'. Poverty mapping has become a vital tool for policy makers in allocating transfers and they provide crucial information on spatial distribution of living standards (Hentschel et al. 2000).

To understand their uses in poverty enquiry, Davis (2003) highlights the purposes of poverty mapping to include the spatial identification of the poor as well as the targeting of poverty reduction programs. Small area estimation, multivariate weighted basic-needs index and participatory approaches are some of the methods used in the creation of poverty maps (Davies 2003).

In their earliest forms, poverty maps were constructed from household survey data and this only managed to generate limited information due to the small sample sizes of the survey data used to construct them (Benson et al. 2007; Alderman et al. 2001; Hentschel et al. 2000). Thus, the use of household surveys alone in the construction of poverty is

bound to face poor disaggregation. Hentschel et al. (2000) argue that it is difficult to find a household survey that has a large enough sample and at the same time have detailed income/expenditure data to allow high-resolution poverty maps.

Recent advances in poverty mapping allow the combination of census data with household survey data and resultantly permit for the disaggregation of analysis to lower levels (Benson et al. 2007). The combination of census data and household data allows for the strengths of each data type to be combined. Whereas household survey data have relatively small sample sizes, census data are not affected by this limitation. However, census data have limited information on income/expenditure for meaningful poverty analysis that is present in household surveys (Alderman et al. 2001) or if available are measured poorly (Elbers et al. 2003). In South Africa for instance, the Income and Expenditure Surveys as well as the Living Condition Surveys collect detailed income/expenditure surveys but creating poverty maps based on such data alone will provide poverty maps that are disaggregated to provincial level. This does not provide detailed information on where the poor are located within such provinces. Alderman et al. (2001) argues that the combination of such household level surveys with population census data allows the 'maximization' of the two data types. The use of poverty mapping to leverage existing poverty data offer an alternative data source in the analysis of poverty in developing countries. One of the gains of poverty maps to date is their ability to quickly give a detailed spatial distribution of poverty (Henninger 2008), which has been useful in policy design and implementation.

One of the most classic applications of poverty mapping that maximises on census data is applied in Elbers at al. (2003). In their method they, took combined the fact that household surveys have detailed income and consumption data as well as the comprehensive coverage of Census data in Ecuador. Using a per capita expenditure variable as the proxy measure of wellbeing, they estimated models that enabled them to analyse welfare by subregion.

Poverty mapping has been used in Bolivia to leverage existing poverty related data in the country. Bolivia also suffers from a similar case of limited data sources in poverty and inequality analysis. In Bolivia, nationally representative data only started to exist in the country from 1997 and not before this period and as a result there is incomplete income and consumption data for poverty analysis (Grosse et al. 2009). Instead the

DHS in Bolivia had been conducted since 1989. To compensate for this incomplete income and expenditure data, Grosse et al. (2009) developed an alternative method of measuring poverty derived from an exercise used in poverty mapping. Their approach linked urban household income survey data with DHS data to create time series household income data from 1989-2002. The approach also performed well on validation tests and it managed to impute incomes from assets in the DHS thereby generating new information on poverty and inequality analysis in Bolivia. The results suggest that the poverty mapping derived methodology is able to reproduce trends differential in poverty where there is comparable data (Grosse et al. 2009). Their findings suggest that it is possible to use DHS surveys in the absence of consistent income and expenditure nationally representative surveys.

Poverty mapping exercises have also been used in South Africa. Alderman et al. (2001) used the Income Expenditure Survey, the October Household Survey and the Population Census to create poverty maps. Their methodology provided alternative imputed income estimates for all households enumerated in the census. They derived prediction models consumption/income data as their endogenous variable, and their exogenous variables were restricted to variables found in the surveys and the census. They managed to set parameter estimates to the census data allowing them to derive poverty and inequality statistics from the two data sets.

This approach is not unique to South Africa alone as it was applied in Ecuador in Elbers et al. (2001). Alderman et al. (2001) argue that this derivation of alternative data sources by combining two surveys has been conducted with instances of literature from the 90's. Arellano and Meghir (1992); Angrist and Krueger (1992) and Lusardi (1996) are cited as part of the body of research that has combined two survey data sets but not with census data.

2.4.2.2 Regression based SSI approaches

Dabalen et al. (2014) used a regression-based SAE approach that uses the relationship between the covariates of poverty in the survey that has consumption data (CWIS) to predict the level of poverty levels in the survey that does not have consumption data (DHS). In their survey-to-survey imputation method, they constructed a comprehensive list of variables from the DHS and the CWIS surveys and the variables chosen were correlated to poverty. This was followed by a stepwise regression run on the CWIS data

that had consumption variables with capital household expenditure as the dependent variable. Then variables were regressed against the log of household consumption. The beta coefficients from this model were then applied to the same variables in 2010 CWIQ to predict poverty levels in that survey (Dabalen et al. 2014). This methodology has its merit since it is based on empirical relationships between consumption and poverty correlates and as a result their correlation was more significant. However, its major drawback is that it requires at least one comparable survey that has consumption variables (Dabalen et al. 2014).

Douidich et al. (2013) used SSI approaches with the Moroccan Labour Force Surveys in estimating poverty from household expenditure between 2001 and 2010. Their argument is that there is data gap experienced during the usual four-five year period time when a survey is released. Their study seeks to address the information data gaps left for the period between the releases of two surveys. To cater for this shortcoming, they use imputation methods to fill in these data gaps. Imputation methods have been used in statistics in dealing with missing data problems, but they are also being used in cross-survey imputations whereby one survey is used to fill in the gaps of a different survey with the same population.

The SSI approach they used accurately reproduces official market conditions and poverty statistics in Morocco. Their methodology effectively makes use of LFS data to estimate poverty in the absence of household Surveys. They used cross-survey imputation methods to estimate poverty rates in Morocco. They created their first consumption model using the HCS data and a second model using the 2007 HCS data. In their analysis, they then used the 2001 model to estimate poverty with the 2001 and 2007 data from the HCS and did the same with the 2007 model. Douidich et al. (2013) state that the imputed poverty estimates found with this method were not very far from the official poverty estimates in Morocco during this period for both models. They further extended the analysis using the 2001 and 2007 consumption models to estimate poverty from the LFSs for the period 2001-2010 and their results were in line with the quarterly poverty trends in Morocco. The results derived from the imputation method used by Douidrich et al. (2013) shows how imputation methods can be used to enhance comparability between datasets. Apart from solving comparability problems in survey data, imputation also allows for new analyses to be observed that would not have been possible with the available data. For instance, Douidrich et al. (2013) argue that the use

of imputation methods provides findings in Morocco's poverty research that would not have been visible when only two points in time were considered.

The use of the alternative methodologies discussed in this review apart from allowing for statistical inference, are useful in re-estimating poverty rates as was done by Tarozzi (2007) in India in a bid to validate official figures. There are other countries that had the same data problems in Latin America, Africa and Asia and Hentschel et al. (2000) argues that they have managed to overcome this problem by combining basic census information and have used this information to construct poverty maps.

The literature in this section has demonstrated the importance of frequent and high-quality data in poverty monitoring and analyses. However, as stated by Dabalen et al. (2014) the need for such data in most developing countries will continue to outpace its production. As a result, alternative methods will continue to grow in importance in poverty and inequality literature. The methodologies reviewed in this section show that asset indices, cross-survey imputation and poverty mapping derived methods can actually be effective in improving poverty and inequality related statistics for policy making and effective poverty reduction strategies. Such methodologies make it possible to study poverty trends using available data sources that would not have been possible. Of importance as well in the methods discussed in this section is the fact that these alternative methods of measuring poverty and inequality also help improve the comparability of household survey data. For instance Douidrich et al (2013:2) found that cross survey imputation methods 'resolve problems of comparability between surveys of the same type over time, for example due to changes in the questionnaires ' This gives a new possibility in poverty research as it provides a way of comparing poverty statistics in time frames that otherwise would not have been possible.

Comparability of data in survey imputations is jeopardised when the definitions and coding of variables is not consistent. Bourguignon and Moreno (2015:18) stressed the need to 'strictly adhere' to similar variable definitions in all the years used in the synthetic panel, an important prerequisite that ought not to be compromised in pseudo panel construction. This requirement perhaps remains the most important in SSI approaches. Alderman et al. (2001) and Tarozzi and Deaton (2008) also highlight the importance of having similar definitions of variables in line with other poverty mapping prerequisites.

2.5 A brief review of synthetic/pseudo panel data approaches applied in the analysis of poverty dynamics

The availability of nationally representative panel data relevant for poverty analysis remains limited despite the fact that the demand for such data in the post 2015 era has significantly increased (Scott and Mariotti 2014). Baulch (2011) discussed at length the different types of panel data available in the developing world, but what emerged from the analysis is that existing panels are not always nationally representative, sometimes not well spaced and, in some instances, have been discontinued.

Cross sectional data is dominantly used in poverty analysis in most developing countries. But Bourguignon and Moreno (2015) state that such data only captures snapshots of poverty at different points in time and making conclusions based on such information more likely to be biased. This is because it remains unclear how over time households have moved in and out of poverty when using cross sectional data in poverty analysis. Static poverty measurements therefore fail to capture the characteristics associated with households trapped in poverty over long periods of time, and thus cannot effectively inform poverty reduction policy.

It is against this context that this section reviews the methods used to create synthetic panel data over the years. The main objective of this section is to provide an overview of the different issues that are considered when creating synthetic panel data as well as highlight the limitations of the approaches used to date. Cruces et al. (2013) argue that the different approaches developed to date are highly dependent on varying data demands, assumptions on structural parameters as well as the research questions that they seek to answer. Thus, all these have to be considered when choosing a synthetic panel data methodology.

This study uses two broad categories to summarize existing approaches in the creation of synthetic panel data. The first type creates data that enables analysis to be conducted at cohort level (Deaton 1985; Deaton and Paxson 1994; Banks et al. 2001; Pancavel 2007), whilst the second category enables analysis to be conducted at household level and individual level, such as the works of Bourguignon et al. (2004); Guell and Hu (2006); and Dang et al. (2014). The use of synthetic panels in poverty analysis has moved from country level to regional analysis. Dang et al. (2017) created a synthetic

panel for 21 Sub-Saharan African countries, and similar panels have been created in the Middle East, Latin America and North Africa.

2.5.1 Cohort based synthetic panel data approaches

One of the earliest forms of synthetic panel data was created by Deaton (1985). His study used household survey data in the United Kingdom (UK) to create cohort based synthetic panels. The unavailability of panel data in the UK on consumer expenditure and household level labour in the 1980s made it impossible to analyse household labour related activity over time. Instead there were a number of household surveys in the UK that had been conducted annually. Deaton (1985) created cohorts that he 'tracked' using household surveys. He was able to analyse the changes over time of income and consumption of different age cohorts in the UK. The household surveys that he used were large enough to generate random samples of individuals from each of the cohorts. A summary of the statistics from these random samples was then used to generate a time series that would work like a panel to infer behavioural relationships for the cohorts.

One of the key assumptions made by Deaton was the use of time invariant characteristics to define the subgroups (Deaton 1985). Time invariant variables are variables that do not change over time (Williams 2018). For instance, the race, or birth year of an individual is assumed to remain the same at any period. This is in contrast to time varying variables that change over time such as an individuals' income or settlement type at two different periods that are subject to change and not fixed. The second major assumption of his approach was the treatment of cohort means as the observations of the synthetic panel. The pseudo panel generated in Chapter Five in this thesis uses similar birth years to generate cohorts and these are 'followed' in a 3-time period (2008, 2010 and 2014).

The use of cohort based synthetic panel data approaches was later further pursued by Deaton and Paxson (1994); Banks et al. (2001); Penchavel (2007); Moffit (1993); Collado (1997); McKenzie 2004. The underlying assumptions of cohort-based approaches to synthetic panel data divides the population using common characteristics into groups where each cohort mean is assumed to be the error-ridden measurement of the population cohort mean (Deaton 1985).

But the use of cohort means is highly dependent on the number of cohorts generated, the time periods as well as the asymptomatic theory applied (Deaton 1985). Cohort based synthetic panels are also affected by sample size, whereby smaller sample sizes are more likely to increase bias. Morever, Meng et al. (2014) also argues that one has to compromise between increasing the number of cohorts which increases the possibility of increased heterogeneity of the individuals in the synthetic panel and increasing the number of individuals in a cohort which increases the robustness of the estimates in the cohort. But one cannot experience robustness and heterogeneity at the same time, hence the dilemma and whether to create more cohorts or fewer cohorts with more individuals.

Verbeek (2008) argues that the use of cohort means only allows for analysis to be conducted at cohort level and therefore fails to provide useful information on the mobility of certain groups. The use of cohort based synthetic panels as applied by Deaton (1985) and subsequent research also had the major limitation of requiring many rounds of repeated cross sections (Bourguignon et al. 2004) and as such, it means it cannot be used in cases where a few rounds exist.

One other limitation of this group of studies as stated by Dang et al. (2011) is that the analysis emanating from such techniques cannot be disaggregated below the level of the cohort meaning little is known on lower level of analysis. For instance, Dang et al. (2011) states that using cohorts cannot give insights on issues relating to racial groups that may be useful in understanding poverty especially in the South African context.

2.5.2 Household/individual level approaches to synthetic panel data

The second type of synthetic panel data approaches allow analysis to be made at household and individual level (see for instance, the works of Bourguignon and Moreno 2015; Bourguignon et al. 2004; Dang et al. 2014; Guell and Hu 2006). Bourguignon et al. (2004) used a unique approach in creating synthetic data to estimate the vulnerability of households to poverty. Their methodology applied synthetic panel data techniques to recover some parameters of individual earning dynamics from cross sectional data to study individual earning dynamics from cross sectional data. These parameters were then used to calculate estimates on the vulnerability to poverty of households. Tracking cohorts of randomly selected individuals born in five-year intervals in successive cross-sectional surveys created their synthetic panel. Their findings suggest that the

parameters of the models from the synthetic panel data were close to the estimations made from true panels.

Dang et al. (2011) also used cross sectional data to make inferences on movements of households in and out of poverty. Their study sought to overcome the limitations experienced when tracking cohorts whereby analysis could not be disaggregated to lower levels. Their methodology enabled them to analyse the movements of households in and out of poverty instead of focusing only on cohort means. They also sought to use an approach that was less data demanding when compared to other early pseudo-panel methodologies that required more than two cross sections. Their approach successfully enabled the analysis of income mobility within cohorts as well as across cohorts unlike previous traditional studies that only showed cohort means. Dang et al (2011) in their approach adopted small area estimation methods used in Elbers et al. (2003) in the creation of poverty maps. They estimated a model of consumption using the first round of cross-sectional data using time invariant covariates. The parameter estimations used in the first model were then applied to the same time invariant covariates regressors in the second round to allow for the estimation of the unobserved consumption of income in the first period for individuals surveyed in the second round. The analysis of mobility was then based on the actual consumption observed in the second round with the estimate of the first round.

Dang et al (2011) managed to create upper and lower bounds to entry and exit from poverty of households. They used two approaches in the derivation of these estimates. The first approach was a non-parametric approach that used no structure on error distribution. The width of the bounds that were created by this approach were dependent on the extent to which time invariant and deterministic characteristics explained cross sectional income or consumption. The major limitation in using a non-parametric approach is that the exact autocorrelation remains unknown and it may be within a narrower range. This limitation is addressed by the application of parametric bounds. The second approach by Dang et al. (2011) is a parametric bound approach and it has more assumptions which enables the narrowing of the width bounds which are broader using the non-parametric approach.

Dang et al. (2011) applied both approaches on data from Vietnam and Indonesia where actual panels are available to allow for the validation of their methods. The conclusion

by Dang et al. (2011) was that in Vietnam and Indonesia where they applied analysis, the 'true' estimate on the extent of the mobility as shown in the actual panels is generally between the upper bound and lower bound assessments of mobility. They also concluded that the use of the parametric approach allows for the mobility estimates to be narrowed down as the models became more specified when compared to the non-parametric approach.

Cruces et al. (2013) created synthetic data using a slightly different approach to the ones discussed above. In their approach, they used panel data from Chile and treated each round of the panel as an independent cross section survey. They used three different approaches in their creation of synthetic panels. The first approach was a mean based approach that identified cohorts and followed cohort means over time. Their second approach estimated vulnerability to poverty measures without panel data and was pioneered by Bourguignon et al. (2004). In this methodology, Bourguignon et al. (2004) simulated individual earning dynamics to estimate individuals' vulnerability to poverty using cross sectional data. The third approach by Cruces et al. (2013) follows a methodology by Dang et al. (2011) in predicting lower and upper bounds for the probability of poverty status. These three methods had different data requirements and assumptions. But the results derived from their estimations concluded that the synthetic panels they created performed poorly when compared to 'true' estimates. For instance, the synthetic panels failed in predicting the mobility pattern observed in Chile and they also performed poorly in predicting a broader set of income mobility measures. Using Dang et al.'s approach, they calculated the joint probabilities of poverty status for 2 time periods. However, this isolated study cannot render all synthetic panels useless, the use of different data and assumptions also translates variability in the proximity of the results to actual panels.

Bourguignon and Moreno (2015) extended the methodology applied by Dang et al. (2014) and Dang and Lanjouw (2013). They eliminated the bi-nominal assumption used by Dang et al. (2011) and Dang and Lanjouw (2013). They introduced a more robust estimation of the correlation coefficients across two periods by adopting methods pioneered by Deaton (1985) and Browning et al. (1985). The third adjustment they made was the use of calibration techniques to allow for the reproduction of the targeted distribution by the synthetic distribution residuals. They also probed a household level estimation of synthetic that enables the analysis of the entire income distribution and

not just the share of the population transitioning in and out of a given threshold. The last adjustment makes the method by Bourguignon and Moreno (2015) unique because it allows for a more disaggregated level of longitudinal analysis of income inequality and income mobility more than previous methods.

The methodology used by Bourguignon and Moreno (2015) split the sample survey into two equal sized parts with one considered as two successive cross sections and the other a genuine panel. A comparison of the two types of panels shows that the synthetic panel resembles closely the actual panel estimates with insignificant statistical estimates.

Household level synthetic panel data approaches have been criticized based on the fact that they are rooted in strong assumptions that are not easily fulfilled on the cross section being used (Dang and Lanjouw 2013). For instance, the reliability of Bourguignon et al.'s (2004) approach is highly dependent on the quality of the matching methodologies used (Bourguignon and Moreno 2015). The use of synthetic panel data has allowed developing countries to analyse poverty dynamics in the absence of genuine panel data.

2.6 Motivation for using pseudo panel data in developing countries

In response to the data gap left by the limited availability of panel data, there has been a growing body of work that creates synthetic panel data for poverty analysis. The availability of a myriad of approaches in the creation of approaches is encouraging since it allows for different ways in creating synthetic panel data. As expected, each approach has its own limitations and strengths. Perhaps the most important factor in the creation of synthetic panel data is the fact that each synthetic panel is created using different assumptions since it seeks to address specific research questions.

Synthetic panels are sometimes advantageous over actual panel data as far as issues related to attrition and non-response are concerned (Cruces et al. 2013; Dang and Dabalen 2017; Deaton 1985; Verbeek 2007; Urzainqui 2017). In support of this argument, Deaton (1985) in his study stated that attrition was not a problem when using synthetic panels as all samples are drawn every year therefore making sure that representativeness is constantly maintained. Deaton's methodology also had the merit that it could recognize measurement error from the start, thus controlling for it from the beginning. To expand on this, attrition reduces the value of a panel and affects the

validity and generalizability of results (Baulch and Quisimbimg 2009; Baigrie and Eyal 2013.

Dang and Dabalen (2017) state that the use of synthetic panel data apart from being exempt from the limitations faced by genuine panel data such as attrition, provide newer and more data for the analysis of poverty dynamics. Urzainqui (2017) also argue that the use of synthetic data types in the analysis of poverty dynamics creates relatively larger datasets that represents the population better than the restricted sample sizes offered by genuine data due to the creation of balanced panels.

Moving on to the relevance of alternative data sources in poverty analysis in South Africa, nationally representative surveys in South Africa only started emerging after 1995. Woolard and Leibbrandt (2006) argue that nationally representative surveys in South Africa are fairly recent, having only emerged in the mid-90s. Despite the existent of a significant number of cross-sectional household surveys to date in South Africa, only one nationally representative panel survey exists. The use of synthetic panel data forms has been pursued in South Africa but not at length especially in poverty research.

Scott and Mariotti (2014) argue that issues on the data type and quality have gained a lot of interest as a result of the Sustainable Development Goals (SDGs). They argue the need for widespread use of panel data in poverty analysis. One of the major motivations in the need for panel data is based on the achievements made in poverty reduction during the Millennium Development Goals (MDGs) era. Existing evidence shows that between 1990 and 2015, there was a reduction of people living in extreme poverty by 50%, but little is known on the individuals who moved in and out of poverty during this period (Scott and Mariotti 2014). The unavailability of panel data in a majority of developing countries makes it a difficult task to understand these poverty dynamics during this period. It is also worth mentioning that the MDGs have been critiqued as statistically flawed (Kenny and Sumner 2011), as they are argued to have provided a 'faulty yardstick' due to their use of absolute measures and disregard of absolute gains in the countries under analysis (Vandermoortele 2009:356).

As much as synthetic panel data may be able to overcome some of the problems associated with genuine panel data, they also have their own limitations. These weaknesses are a result of the assumptions that have to be made, the type of cross-sectional data to be used, as well as the research questions that the synthetic panel seeks

to answer. They are mainly rooted in the assumptions that have to be met which in some cases have proven to be impossible. In fact, Urzanqui (2017) argues that some of these stringent assumptions in the absence of genuine panel data cannot be tested. But the fact they allow for new research questions to be formulated justifies their position in poverty literature. There are cases where synthetic panels have performed poorly when compared to a genuine panel. For instance, Cruces et al. (2013) found their synthetic panel performed poorly in estimating income mobility in Chile.

But there is a reasonable size of literature where synthetic panel data has produced results that are close to those of a genuine panel. Such studies include, Bourguignon and Moreno (2015); Bourguignon et al. (2014a); Antman and McKenzie (2007); Betti and Chelli (1998); Dang and Lanjouw (2013); Dang et al. (2011) are part of the literature that has successfully created synthetic panel data that closely resembles results from a genuine panel. Bourguignon and Moreno (2015) created a synthetic panel to analyse poverty transitions of households and their approach effectively captured this and turned out to be ideal when analysing well-being using cross sectional data.

2.6.1 A place for pseudo panel data in South Africa

The NIDS and the KIDS remain the only key panel data sources for the analysis of poverty dynamics in South Africa to date. This existence of only two panel surveys in South Africa with detailed income/expenditure modules for poverty analysis pre-empties on the existence of a data gap in panel survey availability in the country. This gap is heightened by the fact that only one of these panel surveys, the NIDS is nationally representative. This is in contrast to the relatively abundant availability of cross-sectional surveys that have been used in poverty analysis in the country. The use of synthetic panel data as a substitute for genuine panel data emerged in the 1970s with Deaton and has been used extensively over the years.

In a review written by Bamberger (2009) on a chapter by Kamanaou et al. (2005), he argues that majority of poverty assessments depend heavily on data that has already been collected but the main limitation in this is that such data are not necessarily collected to suit the different research questions specific to poverty analysis. As such the data is not necessarily customized to answer specific poverty related query and at the same time the data still needs to remain comparable over time. Therefore, there is

need to go beyond just analysing available panel survey data, to creating synthetic data forms to meet specific research questions.

Finn and Leibbrandt (2013; 2017) noted the paucity of studies that analyse poverty transitions in South Africa. In fact, they confirm that more is known on poverty trends in South Africa, but little is known on the existing poverty dynamics in the country (Finn and Leibbrandt 2017). The NIDS apart from having a small sample is further limited by attrition, like any other panel survey.

In terms of the KIDS, apart from the fact that it collected data on one province only in South Africa, it tends to have an unrepresentative sample of existing racial groups within Kwa-Zulu Natal. The 1998 wave of KIDS excluded white and coloured households due to their small sample sizes and clustering (Aguero et al. 2007; May 2008; May and Woolard 2007; May et al. 2011). May (2008) argues that the samples of these two racial groups were too small to enable any racial comparative analysis using KIDS and this was compounded by their location in small clusters. The KIDS sample was also not representative of the African and Indian population in Kwa-Zulu Natal province (Aguero et al. 2007).

Inherent weaknesses within existing panel data used in poverty analysis in South Africa motivate for other alternative panels to be explored. A pseudo panel does not follow similar households or individuals, but it follows individuals or households with similar characteristics. Therefore, the use of synthetic panel data types provides new possibilities on poverty research in South Africa. This study seeks to complement existing poverty literature in South Africa that has utilized panel data in exploring the movements of households in and out of poverty and at the same time explores the possible contributions. Antman and McKenzie (2007) states that the use of synthetic panel data reduces potential bias that can be introduced when using a genuine panel due to attrition.

2.6.2 Reservations in the application of pseudo panels

Few studies to date have reviewed the application of synthetic panel methods in poverty analysis, with such appraisals limited to the works of Urzanqui (2017) and most recently Herault and Jenkins (2019). The findings from the review of synthetic panels from the afore-mentioned work suggests that the use of synthetic panel data needs to be done with caution and has its own drawbacks. Urzanqui (2017) for instance

evaluated the Dang and Lanjouw (2013) approach and the Bourguignon and Moreno (2015) method by making comparisons with data from genuine panel data and concluded that approaches that utilize cohort means and not their variance can estimate more correctly the autocorrelation of income residuals. This is besides that fact that there is need for strong assumptions at the cohort and individual level on the comparability of income residual persistence. Urzanqui (2017) also found that synthetic panels tend to be reliable when the estimate of individual income residual autocorrelation is accurate in a short-term analysis of poverty transitions.

Herault and Jenkins (2019) in a recent study found that the Dang and Lanjouw (2013) method when applied to high quality data from Australia and Britain provide estimates that are less accurate when compared to real panel data. Lastly, Herault and Jenkins (2019 concluded that the validity of estimates id dependent on the choices made in the construction of the pseudo panel such as the age of the household head, the choice of poverty line and the years analyze among other choices.

Despite the availability of research that shows the possible drawbacks of using synthetic panels, such data methods remain key in developing countries where genuine panel data are absent and have the possibility of making know the poverty dynamics where only cross-sectional data are available.

2.7 Conclusion

This chapter acknowledges that there exists inadequate cross sectional and panel data types required for poverty analysis in developing countries and in South Africa more specifically. The existence of data gaps in most developing countries is well documented. Therefore, the analysis of both poverty trends and dynamics is hampered by this lack of appropriate survey data in most developing countries leaving the poverty picture incomplete. Cross-sectional data where it is available has been found to be less frequent and/or incomparable. On the other hand, panel data with nationally representative samples remain scarce or non-existence for some countries.

The chapter has demonstrated the importance of frequent and high-quality data in poverty monitoring and analyses and justified the use of alternative data sources. As a result, the application of SSI approaches will continue to grow in importance in poverty literature. The methodologies reviewed in this section show that asset indices, cross-

survey imputation and poverty mapping derived methods can actually be effective in improving the comparability and frequency of cross-sectional survey data. SSI approaches have also been used to create time series data allowing poverty analysis to be conducted for longer periods. Of importance as well in the methods discussed in this section is the fact that these alternative methods of measuring poverty also cater for problems of non-comparability between datasets. This gives a new possibility in poverty research as it provides a way of comparing poverty statistics in time frames that otherwise would not have been possible.

Moving on to the application of SSI approaches in the analysis of poverty dynamics, the creation of synthetic panels has allowed for the analysis of poverty dynamics in settings where the absence of genuine panel data had hampered such analysis. The use of synthetic panel data in poverty analysis has been growing over the years and has been applied in developing countries where there is a dearth of panel data. But, as is expected in quantitative analysis, synthetic panel data approaches have their own limitations and challenges. What is evident in synthetic panel data approaches is the methods used have evolved but they remain problematic. This does not dismiss the fact that they allow for relatively new sources of data for the analysis of poverty dynamics.

Chapter Three

Conceptualizing dynamic poverty

3.1 Introduction

This chapter conceptualizes the measurement of dynamic poverty as a precursor to the task of this thesis: to propose options for the analysis of dynamic poverty using existing data in South Africa. Apart from reviewing the analysis of dynamic poverty, of major importance in this chapter are the distinctions that exist between transient, chronic and structural forms of poverty as well as an understanding of the application of poverty lines in South Africa.

This chapter is made up of three sections; the first section discusses the different types of poverty exploring the differentiations in the use and conceptualization of transient, chronic, and structural poverty. This discussion apart from illustrating the dissimilarities that exist between these types of poverty justifies how such an understanding is key for poverty reduction policy. Secondly, the chapter engages the spatial, gender and racial dimensions of poverty. This part of the analysis extends the policy relevance debate within these dimensions. The analysis of these dimensions is a precursor to the analysis done later in this thesis on birth cohorts in the pseudo panel, where the mean expenditure per capita and poverty experiences of the birth cohorts are analysed within similar dimensions. The section briefly summarises some of the key findings from official poverty reports from Stats SA. The last section with a specific focus on South Africa introduces the use of poverty lines, tracing the history of their usage to build a discussion on the existence of three poverty lines in the country.

In this chapter, the term dynamic poverty is used to refer to poverty that occurs over time. To quote a few definitions, the Chronic Poverty Research Centre (CPRC) (2005:5) defines dynamic poverty as the analysis of changes of people's poverty status over time, differentiating dynamic poverty from poverty trends, with the latter only giving snapshots of poverty at different points in time. Specifically, CPRC (2005:5) defines dynamic poverty as an analysis of how households' poverty status changes or remains the same over a given period of time. Interestingly, Baulch (2013:1) defines the same concept as the 'intra and inter-annual changes in welfare' that causes households to move across a poverty line in different time periods. In short, the

assertion that 'poor people are not poor all the time' in (Yaqub 2000:1) summarises the concept of dynamic poverty more concisely as applied in this chapter.

It should be noted at this point that there is a clear distinction between studies that analyse poverty for a given period of time using static data and those that use longitudinal data to analyse dynamic poverty of similar households over time. This study is concerned with the latter and it is within this understanding that dynamic poverty arguments in this chapter will be referring to.

3.2 Dynamic poverty

To begin the conceptualization of dynamic poverty, the term has not always been used in poverty research. Whereas the analysis of static poverty analysis can be traced as far back as the 1900s with the work of Rowntree (1901), the analysis of dynamic poverty emerged in the early 1960s. The assertion that 'poverty is dynamic' by Townsend (1962:219), is one of the earliest studies to introduce the time element to the analysis of poverty. Research in the 1970s that analysed poverty dynamics remained limited and included research by Coe (1978); Levy (1977) and Morgan (1974). In the 1980s Bane and Ellwood (1985); Coe et al. (1982); Gottschalk (1980); Hill (1981); Rainwater (1982) are part of the early authorship on dynamic poverty.

What is evident from an analysis of this early research in dynamic poverty is an absence of literature from developing countries in the 1960s and 1970s. Elsewhere, the analysis of dynamic poverty exists but is limited and only started gaining momentum in later years. Clark and Hulme (2005:5) state that 'serious work' on dynamic poverty started emerging in the 1980s. However, during this period there was inadequate conceptual progress on the analysis of dynamic poverty. This dearth of literature on dynamic poverty during this period was partly a result of an absence of panel data (Deaton 1997). The consensus that poverty is non-static also started growing after this period and research on dynamic poverty extended to include developing countries.

Table 3.1 below summarizes some of the earliest research on dynamic poverty that occurred between the 1980s and late 1990s in developing countries. The table illustrates an absence of literature on dynamic poverty in developing countries between the 1960s until the 1980s with a majority of studies in this region emerging in the 1990s.

Table 3.1: Summary of early research work on dynamic poverty in developing countries (1980's-1999)

Author(s)	Title of publication	Year	Waves	Country
Gaiha, R.	Income mobility in rural India	1988	3	India
Noponen H.	The dynamics of work and survival for the urban poor: A gender analysis of panel data from Madras	1991	5	India
Dreze, J. Lanjouw, P Sern, N.	Economic mobility and agricultural labour in rural India: A case study	1992	4	India
Gaiha, R. and Deolaiker, A.B.	Persistent expected and innate poverty: Estimates for semi-arid rural South India	1993	9	India
Chaudhuri, S Ravallion M	How well do static indicators identify the chronically poor?	1994	-	India
Scott, C.S Litchfield J.A	Inequality, mobility and the determinants of income among the rural poor in Chile 1968-1986	1994	-	Chile
Grootaert, C. Kanbur, R	The lucky few amidst economic decline: Distributional change in Cote dIvoire	1995	3	Cote dIvoire
Reardon, T. Taylor, J.E	Agro-climatic shock, income and poverty: Evidence from Burkina Faso	1996	2	Burkina Faso
Glewwe, P Hall, G	Are some crops more vulnerable to macroeconomic shocks? Hypothesis tests based on panel data from Peru	1996	3	Peru
Muller, C.	Transient, seasonal, and chronic poverty of peasants: Evidence from Rwanda	1997	-	Rwanda
Baulch, B McCulloch N	poverty transitions in rural Pakistan Being poor and becoming poor: Poverty status and	1998	2	Pakistan
Jalan, J Ravallion, M	Transient poverty in Post-reform China	1998	-	China
Baulch, B McCulloch, N	Distinguishing the chronically poor from the transitorily poor: Evidence from rural Pakistan	1999	5	Rural Pakistan
Dercon S Krishnan P	Poverty, seasonality and intra-household allocation	1999	2	Ethiopia
Gunning, J. Hoddinottt J Kinsey, B Owens T	Revisiting forever gained: Income dynamics in the Resettlement areas of Zimbabwe 1983-1997	1999	4	Zimbabwe
Jalan J	Do transient and chronic poverty in rural China share common causes	1999	6	China

Author(s)	Title of publication	Year	Waves	Country
Ravallion M				
Rajshiri, J Lanjouw, P	The evolution of poverty and inequality in Indian villages	1999	-	India
Cumpa, M. Webb, R	Mobility and dynamic poverty in the 1990's	1999	-	Peru
Mallucio, J Haddad L May, J.D.	Social capital and income generation in South Africa, 1993-1998	1999	-	South Africa

Source: Various sources

The list presented in the table above is by no means exhaustive. The publications included depended on availability in public search domains. Despite this, the summary of research work presented in the table above shows that in developing countries, panel data available during this early period covered short periods and was not nationally representative.

Apart from a lack of panel data in developing countries, it also seems that the concern in the pre-2000 period in poverty research in developing countries was more focused on debates on the use of monetary and non-monetary approaches to static poverty analysis. For instance, Hulme and McKay (2005:2) argue that the time element of poverty was 'neglected' during this period whilst poverty debate was predominantly on poverty depth and multidimensional aspects.

In the period starting 2000, more studies that analysed dynamic poverty emerged in developing countries. Such extensive work on dynamic poverty includes Addison et al. (2009); Barrett et al. (2006); Bigsten and Shimles (2004); Baulch and Hoddinott (2000); Baulch and Massett (2003); Baulch and Vu (2011); Baulch and Davis (2008); Behrman (2006); Carter and Barrett (2006); Calvo and Dercon (2009); Davis and Baulch (2011); Dercon and Porter (2011); Duclos et al. (2010); Ellis and Freeman (2004); Foster (2009); Gunther and Klasen (2007); Krishna (2004); Krishna et al. (2004); Krishna et al. (2006a); Krishna et al. (2006b); Kristjanson et al. (2010); Green and Hulme (2005); Lohano (2011); Machado and Ribas (2010; Lokshin and Ravallion (2004); May (20030; May et al. (2011); McKay and Lawson (2003); Nargis and Hossain (2006); Neilson et al. (2008); Parker and Kozel (2005); Porter and Quinn (20080; Quisimbing (2011); Ribas and Machado (2007); Yaqub (2000). Extensive research on chronic

poverty has also been conducted by the Chronic Poverty Research Centre (CPRC). In South Africa in the post-2000 period, Adato et al. (2006, 2007); Aliber (2003); Carter and May (2001); Finn and Leibbrandt (2013, 2017); May (2008); Schotte et al. (2017, 2018); Woolard and Klasen (2005) are part of the research that has analysed dynamic poverty in the country.

3.3 Understanding the distinctions between transient, chronic and structural poverty

Authorship on dynamic poverty lacks a universal definition in the conceptualization of transient, chronic or structural poverty. However, despite these variations what is clear is that in understanding dynamic poverty, the time duration and severity is the least for transient poverty, and it increases as one moves to chronic and structural forms of poverty. The argument by the CPRC (2011c:9) that 'not all poverty is the same' illustrates the complexity of poverty. It varies in terms of severity, duration, dimensions and social, economic categories.

3.3.1 Transient poverty

Transient poverty has been described as poverty that is 'occasional' (Punton and Shepherd 2015), or as 'fluctuating' poverty (CPRC 2005:5) and similarly as 'short stay' poverty in Bane and Ellwood (1985:21). In Baulch (1996:40), transient poverty is referred to simply as 'temporary poverty'. In a later study, Baulch and Hoddinnott (2000:6) refers to households experiencing this type of poverty as the 'sometimes poor', a term also applied in Carter and May (2001); Hulme et al. (2001); May et al. (1999).

To cite a few examples on the definitions used of transient poverty, Yaqub (2000) does not give a specific time-bound definition of transient poverty, but instead argues that transient poverty is a result of 'short-run variations' in welfare. Elsewhere, transient poverty has also been defined as a temporary shock in consumption (Gunther and Klasen 2007; Jalan and Ravallion 1998) or as a concept that is a result of short-term risks and shocks (Gunther and Klasen 2007). Similarly, Jalan and Ravallion (1998) associate transient poverty with shocks that affect households' consumption abilities in the short term. Lastly, Malaba (2006:6) defines transient poverty as a 'type of poverty that emanates from cyclical or temporary factors and is experienced in shorter periods of time'. In its variations, transient poverty is a short-term phenomenon and affects consumption in short periods of time.

The causes of transient poverty are well documented; Baulch and McCullogh (2003) as already stated earlier, found this type of poverty to be caused by temporary shocks, driving households to fall into poverty but escape in a year or two. Transient poverty has also been attributed to be a result of 'hard years' (Baulch 1996:40). Other causes of transitory poverty include family size, migration and seasonal employment (Ribas and Machado 2007) as well as ill-health, and weather (Yaqub 2000); factors that affect a household's consumption pattern in a short-term period. Jalan and Ravallion (1998) and Greene and Hulme (2005) interpret transient poverty as emanating from short-term variations in consumption or a result of 'seasonality', temporary shocks or downturns on business cycles.

3.3.2 Chronic poverty

Chronic poverty on the other hand has also been referred to as 'long term' poverty (Jalan and Ravallion 1998), 'persistent poverty' (Barrett et al. 2006), intergenerational poverty (Aliber 2003) and sometimes used synonymously with terms such as the 'always poor' (Baulch and Hoddinott 2000:6; CPRC 2005; Hulme et al. 2001:2). In defining chronic poverty, it is evident that the time dimension at the core of this concept is longer than the reference period applied in transient poverty with time frames ranging from five-year periods (Hulme et al. 2001; Hulme and Shepherd 2003) to intergenerational time frames (CPRC 2005; Hulme and Shepherd 2003; Malaba 2006). Whilst transient poverty is associated with a household's inability to deal with shocks in consumption, Baulch and Hoddinnott (2000:9) define chronic poverty as occurring when a 'household's inter-temporal average welfare' is below a given poverty line.

The CPRC (2005:6) in what they term 'poverty chronicity', argue that this concept is a longitudinal aspect that shows the 'persistence of poverty'. In a different study, Greene and Hulme (2005) define chronic poverty as a form of poverty that is associated with limited social and economic mobility of households. Even though there is no direct link between severe poverty and chronic poverty, CPRC (2011a), states that severe poverty is a good indicator of chronic poverty. However, not all severely poor households experience chronic poverty, and the distinction lies in the fact that severe poverty is a static concept that captures the depth of poverty and not the duration (CPRC 2009). Severe poverty experiences are therefore not enough proxies of chronic poverty (CPRC 2011c).

The duration of chronic poverty is often tied to a five-year reference period. This choice has been justified extensively and, in this study, coincides with the use of birth cohorts separated by five years later in this thesis. Corcoran (1995) and Yaqub (2000) argue that individuals who experience poverty for a period of five years or more are more likely to remain poor for the rest of their lives. Hulme and Shepherd (2003) distinguish between chronic and transient poverty and justify that chronic poverty requires interventions that go beyond temporary employment policies, they also applied a five-year period to define the presence of chronic poverty in a household. The premise behind using a five-year period in Hulme and Shepherd (2003) is based on the assumption that this period is adequate to capture changes in a person's lifetime. Baulch and Massett (2003); Carter and May (1999); Yaqub (2000) are some of the studies that have also applied this definition of chronic poverty.

The causes of chronic poverty, unlike transient poverty are complex. The CPRC (2005) found that the chronically poor are not a homogenous group and are found all over the world and this is partly due to the fact that chronic poverty is multilayered and complex. The chronically poor tend to be left out of ongoing development processes and are often found in 'low potential areas' or 'marginalized' areas and resultantly do not benefit from development processes, leaving them behind (CPRC 2005:28).

3.3.3 Structural poverty

Apart from the conceptualization of poverty into transient and chronic forms, structural types have also emerged. Structural poverty in its application suggests a more severe form of poverty that moves beyond what May and Woolard (2007:17) term 'short term windfalls' associated with transient poverty and also goes beyond the conceptualization of some definitions of chronic poverty. Part of the scholarship that have written on this form of poverty include Adato et al. (2006); Baulch and McCulloch (2000); du Toit (2005a); Green and Hulme (2005); May (2007,2010,2014); May and Woolard (2007).

Structural poverty takes into account the processes that lead poor households/individuals to remain poor for long periods (Bevan 2003; Carter and May 2001; Du Toit 2005a.) The level of deprivation associated with structurally poor households tends to go beyond income/consumption deprivation. Adato et al. (2006); Carter and May (2001) found that structurally poor households need to accumulate

assets over time to escape poverty. This suggests that structurally poor households are asset deprived.

What this translates into is that structurally poor households need more than income to escape poverty. In Du Toit (2005a), the argument on structural poverty shows that an increase in income may slightly improve welfare standards, but it may fail to make these households escape structural poverty if their access to resources remains restricted. In a nutshell, the description of structural poverty goes beyond a lack of income but considers the factors that help understand how households remain poor for long such as asset ownership. Structural poverty also is distinct when compared to chronic poverty; it is more embedded in the ability of households to accrue assets adequate to escape poverty.

Malaba (2006) attributed the existence of structural poverty to be tied to the political economy of the country - the dual economy with a modern urban sector and a 'neglected' rural sector a remnant of colonialism. As is the case in South Africa, structural poverty therefore is a result of decades of artificially generated deprivation for selected groups, mostly Africans. Therefore, such groups with a poor asset base find themselves trapped in poverty for long periods of time. In his assertion regarding assets and dynamic poverty, McKay (2009) notes that asset accumulation is key in exiting poverty and a failure to accumulate assets may result in chronic poverty; poor people have a poor asset base and often fail to participate in asset accumulation activities.

3.3.4 Types of poverty: A summation and evaluation

The previous section has demonstrated the variability that exists in the conceptualization of dynamic poverty. Du Toit (2005) states that the distinctions that exist between transient and chronic poverty are not universally agreed on, and various contentions exist. Despite the noted variations, there is an agreement in the analysis of dynamic poverty, that chronic poverty is more severe than transient poverty. There are a number of characteristics that are associated with chronic poverty. McKay and Lawson (2003) give these characteristics to include a lack of human capital, lack of ownership of physical assets, low earnings from paid work, as well as geographic disadvantage. On the other hand, transitory poverty is more associated with short-term events that upset a household's ability to make a living such as ill-health (May 2008; Yaqub 2000).

In developed countries, poverty is mostly a temporal position (Baulch and McCulloch (2003:3). On the other hand, poverty is more permanent in developing countries (Baulch and McCulloch 2003). Baulch and McCulloch (2003:3) also mentioned the existence of structural poverty in developing countries. This suggests that poverty in developing countries is tied with long-term social and economic disadvantage and reflects a general lack of assets that traps households in longer periods of poverty.

On the use of the five-year reference period, criticisms have emerged especially on the application of specific time periods to distinguish between transient and chronic poverty. For instance, the use of the five-year time period used to distinguish between the transiently and chronically poor households has been criticized on the basis that it ignores the time period it might have taken poor households to exit poverty, a factor that is key in addressing poverty reduction policy on households that remain trapped in poverty over time (du Toit 2005a). In short, a household that has been in poverty for a period exceeding five years and one that has been poor for five years cannot be differentiated using this definition.

Instead of focusing on time periods, du Toit (2005a; 2005b) argues that focus should rather be on the processes that cause the mobility of households in and out of poverty. This means there is need to explore the structural factors that keep households poor for long periods of time as well. Conceptualizing chronic and transient poverty in line with du Toit's arguments is likely to be more useful to policy as it gives a more detailed description of poor households over time. At the same the use of the five-year period to define chronic poverty is sometimes preferred due to the ease of application and ease of comparability in different places.

3.4 Dynamic poverty and policy

This chapter argues that understanding the different types of poverty is key in designing poverty reduction policy. There remains no gold standard that is widely agreed on pertaining to the exact time duration associated with transient, chronic and structural poverty, and the applications of existing approaches are dependent on the nature of research enquiry, the data available and other country specific settings. Despite these contentions, there seems to be an agreement within the conceptualization of dynamic poverty among scholars; that distinctions in the analysis of dynamic poverty are key for policy (Clark and Hulme 2005; Foster and Santos 2012; CPRC; 2011a; 2011b;

Thornbecke 2008; Baulch 1996; Baulch and Hoddinnott 2000; McKay and Lawson 2003; United Nations (UN) 2010; Yaqub 2000). For instance, Baulch (1996) argues that the distinction between transient and chronic poverty allows for relevant policies in poverty reduction. Whereas chronic poverty requires more costly interventions that aim to increase the opportunities of the poor, transient poverty may require less costly interventions that may only seek to smooth the income of the poor (UN 2010). Baulch (1996) hints that transient poverty may require temporary solutions such as allowing poor households to borrow or give them insurance on their crops, stable prices or secure forms of employment. On the other hand, chronic poverty requires interventions that are expensive including long-term investments in human capita and land reform and resettlement (Baulch 1996).

Differentiating poverty into its types (i.e. into transient, chronic and structural therefore aids policy), CPRC (2011a) shows that the identification of chronic poverty shows the existence of poor growth in the economy. If we follow the 'economic growth is good for poverty reduction' posed by Son and Kakwani (2004), then part of the solution to reducing chronic poverty would not only involve growth, but, pro-poor growth as suggested by Klasen (2005). However, Dollar and Kray (2002) also warn that the growth of income of the poor is not necessarily a result of 'pro-poor' policies, instead findings show that growth does have a positive effect on poor people's income as well as everyone else and such benefits are dependent on other variables including fiscal discipline and openness to international trade.

Dealing with severe forms of poverty that have longer duration also suggests the existence of overlapping causes and in some cases multiple interacting factors (CPRC 2005). The CPRC (2005) also highlight what they term 'maintainers of chronic poverty' which trap poor people for a long time. At the same time, the CPRC reports suggest that policy also needs to address the 'drivers' of poverty that drive the vulnerable into poverty. Whether they become transitorily or chronically poor depends on their ability to escape poverty, a move that is difficult.

An application of the transient and chronic components of poverty over time in Jalan and Ravallion (2000) found that both forms of poverty significantly reduced if households accumulated physical capital. But whereas fluctuations in ownership of physical wealth resulted in increased transient poverty, chronic poverty remained

relatively lower. The overall picture from Jalan and Ravallion (2000) confirms findings in similar studies that analyse transient and chronic poverty; the two require different policy designs.

Reports from the World Bank and United Nations have shown a reduction in absolute poverty in the past 30 years. Despite these gains, CPRC (2011c) indicates that what remains is chronic poverty. Understanding dynamic poverty is key to poverty reduction policy. For instance, the CPRC (2005) argues that the chronically poor are not a homogenous group. Instead the chronically poor, is composed of groups of people and in some countries, there are certain attributes that are associated with this group even though they are not universal.

Policies that improve economic growth clearly are good for poverty reduction, but whereas their overall impact can reduce the absolute numbers of poor people, chronic poverty is less affected (CPRC 2011a). Therefore, the recent global estimates from the World Bank showing that poverty between the 1990s and 2015 has reduced was a reduction in the absolute number of poor people according to the international poverty line but misses the performance of chronically poor over the same period.

Currently, Sub-Sahara Africa lags behind in terms of poverty reduction and has the highest number of countries still living in poverty (The World Bank 2018b). CPRC (2005) found that Sub-Saharan Africa was the worst affected by chronic poverty globally and there is no recent update on whether these numbers in a more recent period have improved or worsened. However, given the fact that the overall number of people living in extreme poverty between the 1990s and the end of 2015 from recent MDG reports reduced we also expect chronic poverty to have reduced but not significantly.

It also seems that presence of largely stagnant economies, conflicts, exploitative gender relations and unequal class structures found in Sub Saharan Africa partly explains why chronic poverty may remain a challenge in this region (CPRC 2005).

Recent studies in South Africa also show that the country experiences high levels of chronic poverty. In a recent World Bank report, it was reported that close to 50% of the country's population are experiencing chronic poverty[4] and a higher percentage are prone to experiencing transient poverty (World Bank and Stats SA 2018). Using data

[4] The Report used an Upper Bound Poverty Line of R992

from NIDS, the World Bank and Stats SA (2018) found that between 80% and 90% of the poor are in chronic poverty and this group tends to be more dependants on social grants than labour market income. Such households also tend to be female headed, are more likely to reside in rural areas and are dominantly African or Coloured and have the least education (World Bank and Stats SA 2018). Some of the drivers of chronic poverty in South Africa include structural forms of unemployment and geography (Woolard and Klasen 2005; Schotte et al. 2018). In Schotte et al. (2017), transient poverty was found to be more associated with households residing in urban areas and partly a result of movements in and out of the labour income. This could indicate the presence of insecure forms of employment resulting in periods of cyclical unemployment which affects consumption in the short-term period. These findings therefore suggest a multi-pronged approach in the fight of chronic poverty. Clearly different policies are required to deal with structural unemployment as well as dealing with job security to alleviate both chronic and transient poverty.

Existing knowledge on dynamic poverty is tied to the existence of longitudinal data that enable the analysis of poverty into transient, chronic and structural forms. Such data remains inadequate in some developing countries. This study through the use of a pseudo panel contributes to the knowledge of dynamic poverty in South Africa.

3.5 Analysing the dimensions of dynamic poverty

Apart from an understanding of the types of poverty, this chapter also explores three dimensions of poverty; gender, race and geography. The argument made in this chapter is that in South Africa, an analysis of dynamic poverty would be incomplete without incorporating these dimensions as they are well pronounced in poverty debate in the country. However, the selected dimensions discussed in this section are by no means universal; they are specific to the poverty narrative of South Africa, but they have also been relevant elsewhere. The gender and spatial dimensions of poverty share common attributes in some developing countries, with poverty often embedded within these two dimensions. The racial dimension of poverty is less evident in other developing countries but is more pronounced in South Africa as a result of the country's colonial and apartheid history. But related to race are ethnic and social classification cultural systems that often trap entire tribes or societies in poverty as a result of marginalization and discrimination.

3.5.1 Gendered dimensions

The 'feminization of poverty' has become somewhat of a buzzword that has triggered debates in poverty research. Early studies that have pursued the gendered dimension of poverty attempted to analyse the incidence of poverty in female headed households and these include Allen (1992); Appleton (1996), Buvinic (1990); Buvinic et al. (1978), Buvinic and Gupta (1993,1997); Chant (1997a,1997b;1999); Gimenez (1987); Graham (1987); Kumari (1989); Lewis (1993); Millar (1992), Moghadam (1996,2005); Pearce (1978); Youssef (1978). Recently, Horrell and Krishnan (2007); t et al. (2015); Milazzo and van de Walle (2015); Minot et al. (2006); Posel and Rogan (2012); Quisimbing et al. (2001). World Bank (2011) have further pursued this feminisation of poverty debate.

The 'feminization of poverty' debates are argued to have emerged in the 1990s (Chant 1997a; 1997b). There are two broad groups of studies on the 'feminization of poverty'; the first is concerned with the differences between men and women's incidences of poverty whilst the second group of studies analyses poverty within FHH and MHH (Klasen et al. 2015). This study is concerned with the latter. To explore part of this scholarship, Bolivinic and Gupta (1997), Chant (1997a; 1997b), and The World Bank (2001) found that FHH are more susceptible to poverty. More recently, World Bank (2011) also found the FHH when compared to MHH are disadvantaged in terms of access to credit and land. In a similar argument, Klasen et al. (2015) also found FHH more susceptible to cultural and social discrimination. In most developing countries women have no access to land (The World Bank 2007; 2011).

The susceptibility of FHH to poverty when compared to MHH is not always straightforward. Klasen et al. (2015) argue that as much as there is an association between FHH with poverty, FHH are not homogenous and not all FHH are poor, an important distinction that is central in understanding the 'feminization of poverty' debates. Milazzo and van de Walle (2015) concur with the caution initially voiced in Chant (1997a; 1997b), that the 'feminisation of poverty' cannot be generalized. Instead, they are of the view that FHH are transitory, and this is further compounded by the variability in household size and composition. Whether FHH experience more poverty than MHH is not clear and often tied to the variations that exist within FHH (Klasen et al. 2015). Bolivinic and Gupta (1993) argue that understanding the feminization of poverty debate is inadequate in the absence of panel data because the static data does not provide an adequate understanding of the differences within FHH.

Despite these cautionary notes, more recently, Milazzo and van de Walle (2015) revisited the 'feminisation of poverty' and found that the share of FHH in Africa has been rising. In South Africa, poverty has a gender dimension (Aliber 2001:10). Aliber (2003) found that the majority of the poor in South Africa were residing in FHH. Posel and Rogan (2009) found that FHH are more vulnerable to poverty than MHH. In their research, they also found that in South Africa, poverty has been reducing in South Africa, but the decline has been less pronounced for FHH. Despite the fact that the findings have mixed results, the association of poverty with female headship is key when designing poverty reduction policy in South Africa.

3.5.2 Geographical dimensions

The identification of poverty within certain geographies suggests that poverty is often a case of geographic disadvantage. Research has also explored this spatial dimension of poverty (Aliber 2001; Bird et al. 2010a, 2010b; CPRC 2005; Daimon 2001; Grant 2010; Higgins et al. 2010; Hill 2002; Minot and Baulch 2005; Minot et al. 2003; Oelbaum 2010; Shah 2010).

Higgins et al. (2010) using a spatial dimension to poverty analysis, argue that location and geography are key in understanding poverty. In fact, in their findings poverty was associated with 'disadvantaged places', cities and slums as well as place with 'low potential', 'weakly integrated' or marginal (Higgins et al. 2010:2). They justify spatial poverty to be partly a result of 'geographic disadvantage' (Higgins et al. 2010.2).

Bird and Shepherd (2003:4) also found that poverty was mostly found in 'semi-arid regions', or places that are 'neglected', 'unmonetized' and where markets are not well monetized. Minot et al. (2006) in a dynamic poverty study in Vietnam found poverty to be concentrated in areas with poor infrastructure, less urbanised and highly dependent on agricultural activities. CPRC (2005:28) also found that chronic forms of poverty are more associated with people residing in 'low potential rural areas', 'slums', and 'marginalized places'. UNHABITAT (2016) found that one in every eight people globally live in slums and this translates to a significant population. In what Garland et al. (n.d.) calls the urban poor, increasing urbanization in developing countries has made poverty an urban problem. Marx et al. (2013) defines slums as poverty traps and a long-term problem. Extensive research shows the existence of slums in Kibera, Kenya (Desgroppes and Taupin) and in parts of India (Cronin and Gunthrie 2011).

Spatial evidence of poverty has also been observed in China where a higher percentage of the poor are in Western China, a region regarded as lagging behind in the WDR of 2009. The Southern region of China on the other hand has the highest number of poor people in absolute numbers (World Bank 2009).

The World Bank and Stats SA (2018:xxiiii) recently concluded that poverty in South Africa has a 'strong spatial dimension. Earlier studies such as Aliber (2001) found poverty to be more prevalent in rural areas. Earlier studies in South Africa in the immediate period after apartheid, found that rural households located in the former homelands had no land ownership, and livestock increasing their risk to poverty (Aliber 2001). The 1970 Bantu Homelands Citizenship Act arguably was the onset of extreme spatial inequalities of opportunities resulting in the African race being the most deprived. The Act placed Africans in 10 homelands and Zuberi et al. (2005) states that the Transkei Bophuthatswana, Venda and Ciskei (TBVC) states were set up as 'racial and ethnic specific geopolitical units' that were dominantly African.

In South Africa, the spatial dimension of poverty was a result of 'separate development', a direct result of Apartheid (Aliber 2003:476). The result was that due to geographic location homelands were denied proper infrastructure and services and when the country gained democracy in 1994, poverty was 'inherited' (Aliber 2003:476).

Aliber (2003) also found that poverty was more dominant in rural areas in South Africa partly a result of limited productive resources and a general lack of employment opportunities. The CPRC (2005) found that in some countries there are some geographies that are associated with chronic poverty-slums, rural areas and even in some cities. To illustrate the magnitude of poverty at the end of apartheid, Stats SA (1999) found that close to 1 million Africans residing in the former homelands had no arable whilst 1.4 million had no livestock and more than half a million owned neither.

A recent study revealed that Eastern Cape, KwaZulu Natal and Limpopo have the highest poverty rates in the country (The World Bank and Stats SA 2018). This is no coincidence as these provinces also had the highest share of homelands during the Apartheid era (Higgins et al. 2010). Stats SA (2018a) also concur that the spatial dimension of poverty in the country reflect Apartheid spatial policy designs.

3.5.3 Racial dimensions

Whereas spatial and gender dimensions of poverty seem to be well captured, the racial dimension is less documented in developing countries. Instead, poverty seems to dominantly reflect geographic and gender specific disadvantage. In other parts of the world instead of race, poverty tends to affect some ethnicities more than others and in the case of India there exists a caste system with lower castes the most affected.

In Bolivia, Brazil, Ecuador, Guatemala, Mexico and Peru, indigenous populations tend to be the most affected by poverty (The World Bank 2015). Since such poverty incidences has been extended over generations, it also tends to be chronic and intergenerational (The World Bank 2015). Often some indigenous tribes as is the case in Guatemala face social exclusion. As a result, they become marginalized and this increases their chances of remaining poor for long periods.

Poverty in South Africa has a distinct racial dimension particularly affecting the African and Coloured racial groups (Aliber 2001:10). A majority of poverty research in the post-1994 period in South Africa almost always has a racial component in analysing poverty. During the Apartheid era as a result of Apartheid legislation, the African and Coloured races failed to accumulate wealth. For instance, the Native Lands Act of 1923 put in place restrictions on the buying and renting of land for Africans in certain areas. Further legislation such as the Mines and Workers Act of 1911, Natives Labour Regulation Act 1911, further restricted earnings potential of black Africans. In the immediate period following the end of apartheid, Stats SA (1999) found that close to one million Africans that were located in the former homelands had no arable land and 1.4 million owned no livestock, reflecting a poor asset base.

As a result of years of deprivation and disadvantage, poverty for the black African race has become structural, intergenerational and chronic. The stark differences that exist between the high poverty rates experienced by the black African race compared to the almost-non-existent poverty rates for the White racial category illustrates the country's history. This is also manifested in the country's high rates of inequality.

3.6 Poverty reduction policy on gendered, spatial and racial dimensions of poverty

Poverty reduction policy in developing nations has witnessed a mix of policy designs in different countries and the results have been varied. There is a clear distinction between policies that are universalistic in nature from those with specific targeting. Often there is a mix of targeting and universalism to meet the levels of deprivation for specific geographies, races and genders. This section extensively uses evidence from the World Development Report (WDR) of 2009 as well as analyses of various Poverty Reduction Strategy Papers from various countries.

3.6.1 Policy responses to gendered dimensions of poverty

Understanding poverty from a gender dimension is argued to go beyond the provision of descriptive statistics and uncovers the causes (Human and Sciences Research Council (HSRC) (2010). Whilst geography is mostly a natural phenomenon, gender is a social construct influenced by power relations and is often multi-layered in various aspects of culture (HSRC 2010), making it an evasive concept. Understanding how to design poverty reduction policy to deal with the gendered notions of poverty is equally complex as a result. What is clear is that the position of women globally has improved but the gains have been uneven and gender inequalities still exist (HSRC 2010). Women have limited access to assets and other key resources related to well-being which makes them more vulnerable to poverty (Chant (2003a;2003b;2006; Kabeer 2003). A recent conclusion by UN Women (2017) that there are more women than men experiencing poverty is therefore a confirmation that the problem has lasted for so long and still exists.

Policies in developing countries have incorporated the 'gender equality' buzzword in their PRSPs, however the results have been mixed and their failure often pinned to poor implementation. For instance, in India, to attempt to promote women's participation in government and reduce their vulnerability, the government passed a law that reserves a certain number of seats for women in local government (The World Bank 2011). Gender targeted poverty reduction policy has also supported women's access to finance (The World Bank 2011). Despite the mixed results that came out of these polices the efforts are clearly worth applauding. In another case study, Bangladesh's policy mix

has tended to favour employment and education biased policies to promote women (The World Bank (2011).

In South Africa there has been laudable progress on the representation of women in politics, yet past policies have not fully embraced the gendered dimensions of poverty. The Reconstruction and Development Programme (RDP) for instance is argued to have stated on paper a gender-based development approach but this was never properly implemented (HSRC 2010). Polices that followed such as the Growth, Employment and Redistribution (GEAR) and Accelerated and Shared Growth Initiative for South Africa (ASGISA) have also been argued to be gender insensitive (HSRC 2010). Despite this, the current spread of social grants in the country with a special focus on caregivers has had a positive impact on women's poverty status. For example, Patel et al. (2012) found that even though the Child Social Grant was not specifically designed to promote gender equality; it has resulted in a positive outcome on women's vulnerability to poverty. Moreover, women's representation in parliament also shows significant improvements. In 2004, 33% of Member of Parliaments were women and this increased to 43% in 2009, 40% in 2014 and 46% in 2019 (Gender Links 2019).

3.6.2 Spatial poverty and policy targeting

The World Bank (2009) in the World Development Report (WDR) of 2009 showcases one of the most comprehensive analyses of geography and poverty. The conclusions from this report suggest that geography has a strong association with the welfare of an individual (World Bank 2009:1). In a clear example, a child born in New York is more likely to have a better life when compared to a child born in a similar time in a third world country -and strongly argues that a person's future income is predicted more with 'where' they are and less by 'whom' or 'what' they know (World Bank 2009:1). Geography therefore often limits the opportunities that people have in securing jobs and ultimately a good standard of living.

Essentially, the World bank (2009) report shows that spatial inequalities exist at local, national and at a global level, however, the manifestations of poverty over geography has different characteristics and causes. This heterogeneity translates to the application of a mix of dissimilar poverty reduction policy and as will be shown later in this section, implementation tends to be problematic. One of the policy recommendations in the

WDR report (2009) for areas that are described as 'lagging behind'[5] is the implementation of polices that promote the migration of people from 'lagging areas' to more 'affluent' places (World Bank 2009). However, the application of such policy is effective if the population of people living in the lagging behind area has a lower percentage share of poor people in that specific country (World Bank 2009). Moreover, the assumption made for this policy directive to work is that labour and capital are mobile (World Bank 2009). Such a policy would work for instance in Laos where there are sparse populations of people in the areas that are lagging behind; however, the presence of these sparsely populated poor pockets further complicates implementation of such policy because it makes labour mobility difficult (World Bank 2009). Poverty reduction policy therefore in this case should rather be universal since targeting is unclear (World Bank 2009).

In another example the World Bank (2009) shows that poverty reduction policy that promotes the mobility of people from lagging areas also gets ineffective if such areas have a significant proportion of the country's population (World Bank2009). Essentially movements to 'leading' regions would generate new pockets of poverty as more people compete for jobs and resources. The World Bank (2009:230) instead argues that what is required are 'spatially constructive infrastructure' which ultimately are assumed to lead to 'spatial convergence' between the 'affluent' and 'lagging regions'.

India has pronounced spatial poverty; however, the circumstances are not the same as observed elsewhere. The World Bank (2009) for instance found that poor people living in the lagging areas of India self-select to remain there because in those areas they have access to good land for agriculture and have stronger social capital than elsewhere (World Bank 2009). The policy response given by the Indian government attempted to create 'balanced growth' in the entire country- this policy mix largely failed to reduce poverty and India later adopted policies that sought to promote access to basic and essential service in the lagging behind places as well as infrastructural support in such regions (World Bank 2009).

[5] 'Lagging behind' in the World Development Report (2009) is used to refer to areas that are economically distant from more affluent regions and often have higher poverty rates.

Higgins et al. (2010) using PRSP from 14 developing countries found that China, Ghana, Vietnam, Mexico, Laos and India have applied policies that attempted to deal with the spatial dimension of poverty. In Ghana for example, there is a geographical divide between the Northern states and the Southern states with the latter more developed as a result of a concentration of agricultural activities and minerals (Higgins et al. 2010). One of the policy responses that sought to address educational inequalities between the two regions targeted schools in the Northern region viewed as 'lagging behind' and attempted to upgrade all schools to the standards observed in the Southern regions.

What is clear from the application of 'spatially blind policies'[6] in poverty reduction policy is that they might miss certain geography pockets of poverty and become ineffective. However, targeting does not always produce positive results. Geographic targeting in some cases is ineffective especially in a country where such places are densely populated or are located in various places and sparsely populated as indicated earlier. Spatial poverty targeting however enables the isolation of the poorest and when implemented effectively yields good results in some countries. The World Bank (2009) found that spatial targeting in poverty reduction policy should address land and market distortions and often focus on the removal of barriers that restrict the movement of people as well as the provision of basic services as some of the policies that have been applied with success.

What makes spatial poverty targeting policy difficult is that areas regarded as lagging behind are not the same; some are sparsely populated (Laos) whilst some have dense populations (Brazil). At the same time applying spatially blind poverty reduction policy may also be ineffective. In Nigeria for example, the World Bank (2009) found that the country does not have a well-defined targeting policy and a majority of the funds available are mostly equally shared. Nigeria has significantly high poverty rates and geographically the divide between the Southern and Northern Nigeria is hard to miss. Poverty reduction policy therefore is not one-dimensional and often includes the fusion of targeting and universalism and often-applies sectorial interventions (World Bank 2009). Higgins et al. (2010) also made similar conclusions and argued for the need of

[6] The World Bank (2009:240) in the WDR defines spatially blind policies as those that are not specifically designed to consider geographic factors.

a balance between universalism and targeting as well as the need for short term and long-term policy choices.

3.6.3 Race and poverty policy

Poverty reduction policy where some population groups are more likely than others to experience poverty is very sensitive and often policy is not bold enough to target specific racial groups. In South Africa for instance where this is the case it would be impossible for a specific poverty reduction policy directly addressing only the African black population. Policy on poverty reduction has therefore been less obvious. However, the government through the various social grants provided in the country partly addresses the problem. The means tests associated with social grants in the country reaches the poorest households. Since the Black African racial category is the most affected by poverty in South Africa, the conclusion reached by Patel et al. (2012) that the CSG for instance is well targeted seems to partly address the racial component of poverty.

However, there are other countries that have designed specific policies to address certain population groups. For instance, in India, the Caste system results in lower mobility of other Castes and people within such castes remain in poverty traps for long (The World Bank 2011). To address this, the government introduced laws that prohibit discrimination on the basis of tribe or caste and also generated public employment and educational policies to try and create employment opportunities and educational opportunities for everyone (The World Bank 2011). However, policy on paper is inadequate, the implementation of these policies has proven to be problematic in India and poverty remains associated with some castes more than others (World Bank 2011).

3.7 Poverty Lines

Poverty lines are central in the analysis of monetary poverty (Statistics South Africa 2008). A basic definition of a poverty line is given by Kamanou et al. (2005), who define it as a specification of income or a given level of spending required to purchase a bundle of essential goods. Coudouel et al. (2002) also define a poverty line as a cut-off point that separates the poor and non-poor households. One of the key arguments of using a poverty line is that it allows for specific targeting of resources to those who are located below the line. The use of the poverty line also allows for the analysis on the extent of poverty (Kamanou et al. (2001). In understanding the concepts

surrounding poverty lines, there are 3 approaches to poverty line conceptualization that have been applied over the years; absolute poverty, relative poverty and subjective poverty.

The use of poverty lines remains relevant in poverty measurement today. In justifying their position in the poverty debate, Woolard and Leibbrandt (2006:18) argue that poverty lines are not perfect constructions, but rather are useful in poverty measurement since they allow for a line to be drawn between the poor and non-poor. In this line of argument, Greely (1994:11) also states that 'an absolute and objective poverty line is a form of information that empowers the poverty reduction agenda and encourages appropriate resource allocation'.

Baulch (1996) argues that the use of poverty lines is surrounded by uncertainty. For instead, the use of the minimum energy requirements may not be the same for different people due to differences in activity levels, conversion efficiencies and size (Baulch 1996). There also exist variations in the allowances for non-food consumption between different individuals in different places with non-food expenditure patterns tending to be high in urban areas than rural areas (Baulch 1996).

3.7.1 Absolute poverty line

Absolute poverty is a category of money metric approaches to poverty measurement that uses income and expenditures as a measure of well-being in absolute terms. There are a number of explanations that have been given by different scholars for this concept but all of them suggest a failure to meet a minimum standard of living as an indicator for the existence of poverty. For instance, May (2008:39) defines absolute poverty as 'the inability to attain an absolute minimum standard of living' according to a given minimal standard which is quantifiable.

The absolute poverty line is linked to a basket of goods and services that are essential for a minimal standard of living (Lanjouw 2001) and Foster (1998:336) defines an absolute poverty line as a 'fixed cut-off level applied across all potential resource distributions' which remains unchanged over time despite the changes that might happen in the economy. In Kamanaou et al. (2005:32), an absolute poverty line is defined as one that indicates deprivation in the absolute sense; as the value of a set level of resources necessary to maintain a minimal standard of well-being. Using this definition, poverty is eliminated when all households are equal or are above the poverty

line (Kamanaou et al. 2005). In understanding the usage of absolute poverty lines, Stats SA (2008a) links them to a specific welfare level and interpreted as fixed standards of living.

3.7.2 Relative poverty lines

Money metric poverty measurement can also be understood in relative terms. May (2008) argues that conceptualizing poverty using the relative approach explains poverty of an individual or household relative to the well-being of other people in that society. More specifically, Stats SA (2008a) define a relative poverty line as a cut-off point in the welfare distribution. But the application of this poverty line is not useful when making comparisons over time because there will always be a certain population that falls below the demarcated threshold even after standards of living improve over time Stats SA (2008a). Houghton and Khandker (2009) associate relative poverty to the 'the poor are always with us' quotation, because there is always a population defined as poor using this approach.

What is apparent in the conceptualization and measurement of relative poverty is the comparison of society into two groups demarcated by a given cut off point. Lanjouw (2001) argues that the construction of a relative poverty line has the merit that it is simple and transparent since its construction is determined by a percentage cut off point in a welfare distribution. But, using the relative poverty line is problematic when monitoring poverty over time and space since there is always a certain percentage of people who are at the bottom of a distribution even after living standards have risen (Lanjouw 2001). As a result, the challenge posed by using the relative poverty line in poverty measurement is that it cannot trace poverty over time and space (Lanjouw 2001; May 2008). As a result of the difficulties in comparing relative poverty lines, the absolute poverty line is preferred, and it allows comparisons to be made over time and across countries. For instance, Lanjouw (2001) argues that a majority of countries using poverty lines use absolute poverty lines. The absolute poverty line is preferable to the relative poverty line as it allows for comparisons to be made over time and across different groups (Lanjouw 2001). More so, this approach defines poverty not in comparison to the rest of the society as is the case with relative poverty but according to a given standard of living that makes it advantageous.

3.7.3 Subjective poverty lines

Poverty lines have dominantly been categorised into relative and absolute poverty lines, and the use of subjective poverty lines has not been equally extensive in developing countries. The use of subjective poverty lines is an attempt to incorporate the perceptions of individuals on poverty status (Flik and van Praag 1991). The application of subjective poverty lines assumes that people are best in determining their own poverty status (Flik and van Praag 1991); and allows for a more 'democratic approach' that uses non-experts in the definition of poverty (Noble 2008). Subjective poverty lines are based on a minimal income question where a household barely makes ends meets (Houghton and Khandker 2009). Whereas this seems to show clear distinctions from absolute and relative poverty conceptualization, Hagenaars and van Praag (1985) argue that the application of subjective poverty lines is rather a fusion of absolute and relative poverty concepts. For instance, Frye (2015) argues that the use of subjective poverty measures has components of relative poverty. This is because the minimum income question cannot avoid people comparing themselves to other people surrounding them (Frye 2015).

One of the earliest forms of subjective poverty lines was the Leiden Poverty Line applied in Goedhart et al. (1977) and in van Praag et al. (1980). Ravallion (2012) differentiates between qualitative and quantitative based subjective poverty questions. The Economic Ladder Question (ELQ) and the Satisfied with Life questions are part of the qualitative side of subjective poverty, whilst the Minimum Income Question (MIQ) is quantitative (Ravallion 2012). Previous research has shown that responses to the Minimum Income Question given a specified level of income tends to be an increasing function of the actual income (Ravallion 2012).

In South Africa, the analysis of subjective poverty remains limited and this is a result of fewer studies that have questions relevant to subjective poverty. Posel and Rogan (2013, 2017) constitute part of this body of work that has used subjective poverty analysis in the country. Posel and Rogan (2013) compared the poverty estimates using objective and subjective poverty measures and their analysis found an overlap in subjective and objective poverty estimates of households in South Africa.

Although the use of subjective poverty lines allows the incorporation of objective and self-perceived notions of poverty, subjective welfare cannot be equated with welfare

(Ravallion 2012). Rather the application of subjective poverty seeks to provide poverty related indicators when objective poverty analysis has missing or unreliable data (Ravallion 2012). There is paucity of work that applies subjective poverty analysis in South Africa. Apart from being associated with limited availability of data with the subjective questions, the analysis is also difficult to operationalize when compared to absolute poverty measures.

3.7.4 Poverty lines in South Africa

A brief analysis on the history of poverty lines traces the first poverty line in South Africa as early as the 1940's. Magasela (2005) states that this pioneer poverty line had two components; the Primary Poverty Line and a Secondary Poverty Line and was conceptualized by Professor Batsom. The former consisted of the cost of food and clothing whilst the latter included the accommodation and transport costs (Budlender 1985). In the conceptualization of the Batsom Poverty Line, it sought to analyse poverty by calculating the income level below which a household would be defined as poor and it incorporated six categories, which were considered 'essential for minimal survival' (Budlender 1985).

As expected, disagreements from the weaknesses of the Batsom poverty line resulted in further disagreements and research. What is clear from the academic debates in the use of poverty lines in South Africa is that there is a lack of consensus in the line that can be considered the official poverty line of the country. Existing evidence from poverty literature in South Africa in the pre-1994 period shows that the poverty lines that were used for poverty analysis were the Minimal Living Levels (MLL), the Supplemental Living Level (SLL) and the Household Supplemental Level (HSL). One of the major critics given by Budlender (1985) is that the conceptualization of poverty lines in the pre-1994 period was based on the argument of the *homo-economicus* assumption which he argues to be inappropriate when measuring minimal living standards. The assumption that everyone has perfect knowledge of the market disregards the fact that the poorer segments of the economy have the 'least education' (Budlender 1985:4). The assumption that also dominates the conceptualization of the MLL poverty line is based on the premise of average expenditures of all households (Budlender 1985). Budlender (1985) in her critique of South Africa's poverty lines in the pre-1994 also criticizes the conceptualization of the Batsom poverty line. She argues that the use of poverty lines in the pre-1994 period was also conceptualized on

issues of races such that the lines were constructed differently for the different races in South Africa.

In a report published in 1995, Budlender et al. (2015) state that efforts were made to enumerate the poverty statistics in the country by the Presidency in collaboration with the World Bank. This report dominantly used the above-mentioned poverty lines in the analysis of poverty trends in the country as starting point to the conceptualization of new poverty lines.

The dominance of monetary poverty measurement in South Africa places poverty lines at the core of the poverty measurement literature. May (2001:33) states that many countries use an upper bound and lower bound poverty line to create a distinction between 'the poor' and 'the ultra-poor'. Lipton (1983) indicates that the 'ultra-poor' are more deprived when compared to the 'poor'. In short, the 'ultra-poor' are worse off when compared to the 'poor'.

In South Africa poverty measurement has dominantly been monetary in the post-1994 period (Stats SA 2008a). This is not unexpected; UN (2010) argues that poverty analysis in developing countries is dominantly monetary measured. This is partly explained by the convenience posed by using a predetermined threshold. The World Bank's $1.90 a day measure defines poverty in absolute terms allowing for universal comparisons among countries (UN 2010). Thus, the ease in the application of these monetary measures explains their dominant use in developing countries

The use of poverty lines in South Africa has been deeply contested in the past and the choice of poverty lines has been changing (Bhorat et al. 2011; Budlender et al. 2011). South Africa does not have one official poverty line and part of this is a result of ongoing debates on which poverty line is the best for the country. South Africa utilizes three poverty lines; an Upper Bound Poverty Line, a Lower Bound Poverty Line and a Food Poverty Line (Statistics South Africa 2008).

3.7.4.1 Food Poverty Line (FPL)

The FPL is an absolute poverty and is the value given below which individuals are unable to consume enough food to give them the minimal energy requirement for good health (Stats SA 2018c). The assumption behind this line is that households whose income fall below the line are unable to afford or buy sufficient food that allows them

to consume the minimal caloric intake to meet the minimal energy requirement for a basic lifestyle (Stats SA 2018c).

There are two approaches to the construction of a FPL; the least cost and the expenditure method. In the former approach, a food basket is selected from a range of baskets which is assumed to provide the minimal energy requirements at the least cost (Lanjouw 2001). The use of an FPL constructed using this approach is criticized on the grounds that it does not necessarily reflect the actual consumption patterns of households. On the other hand, the expenditure approach to creating the FPL is designed to reflect the actual consumption patterns of a household as shown by expenditure data (Stats SA 2018c). The fact that this approach takes into account the local conditions and past expenditures of a country makes it preferable over the least cost method (Stats SA 2018c).

In South Africa, the FPL is constructed in three stages using the cost-of-basic needs-approach. Stage 1 establishes the food basket. Due to the diversity of food preferences in South Africa a number of food baskets are selected, and these vary according to location and may change with time (Stats SA 2018c). This allows for a reference basket to be constructed that is representative of the consumption patterns of the country, and the final basket shows the costs of 2 100 Kcal per person which is the minimal daily energy requirement for the reference basket (Stats SA 2018c). The use of international universal standards in the making of FPLs makes it easy to make comparisons across space and time

3.7.4.2 The Upper Bound Poverty Line (UBPL) and Lower Bound Poverty Line (LBPL)

The UBPL and LBPL include a component that increases the absolute line of the FPL to take account of needs beyond food. As such, these lines demonstrate the attributes of a relative line. Both rely on the cost-of-basic-needs approach in which two sets of non-food expenditure derived from two different households are added to the poverty line (Stats SA 2008a). But, unlike the FPL where international standards are used, this is not possible when using the UBPL and the LBPL. It is difficult to come up with a basket acceptable widely and relevant over time (Ravallion 1998; Stats SA 2018a). In fact, Haughton and Khandker (2009:52) argue that the measurement of the non-food component part of the poverty line remains non-uniform and is 'case specific'.

The UBPL and LBPL are two variations of poverty lines that add non-food expenditures to the food poverty line. Whereas the UBPL uses households that have food expenditure that is equivalent or equal to the food poverty line, the LBPL uses a different set of reference households and assumes that such households sacrifice their basic food needs to meet non-food expenditure (Statistics South Africa 2008). The LBPL adds the non-food needs of a reference household whose total expenditure is close to the poverty line (Stats SA 2008a).

3.8 Poverty trends and the relevance of a cohort-based pseudo panel in South Africa

In line with poverty estimates calculated later in this thesis, this section will report Stats SA's official poverty rates based on their application of the LBPL. In the period between 1995 and 2000, few poverty analyses were conducted by Statistics South Africa. Existing household surveys data in this period included the OHSs as well as early versions of the IES conducted in 1995. This section seeks to briefly present the analysis that Stats SA has done on poverty in the period after 1995. The objective is to justify the new possibilities presented by using a cohort-based pseudo panel that will be used later in this thesis. Furthermore, the section by summarising the work that Stats SA has done in poverty measurement in the country creates a foundation for the next chapter that engages key informants on their knowledge on the production and use of poverty related data.

Between 1995 and 2000, the OHSs (1995-1999), the IES of 1995 and the Population Census of 1996 provided data for the analysis of poverty trends during this time period. However, poverty analysis from Stats SA during this period was not comprehensive. Existing surveys during this period were mainly used to explore living conditions of South Africans and not necessarily to provide aggregated analyses of poverty. For instance Stats SA using the OHSs (1995-1999), explored the living conditions of South Africans and the findings showed an increase in the number of households in formal and informal settlements accompanied by a decrease in the access of flush or chemical toilets (Stats SA 2001) This reflected movements of people from rural areas to cities that happened after the end of Apartheid.

More publications within Stats SA on the country's poverty trends expanded in the early 2000s. The 'Measuring poverty in South Africa' by Stats SA (2000) became one

of the early publications from the agency to utilize the OHSs, IES and Population Census to analyse poverty. The study also introduced the use of multidimensional poverty measures by applying poverty mapping exercises as well as a Household Infrastructural Index and a Household Circumstances Index. Using the Household Infrastructural Index the findings showed that Eastern Cape, Northern Province, KwaZulu Natal, North Western, Mpumalanga and Limpopo needed the most infrastructural development when compared to other provinces (Stats SA 2000). However, when the analysis is extended to the use of the Household Circumstances Index, the Eastern Cape, KwaZulu Natal, and Northern Province had the worst indicators. Provinces affected by both indices indicate extreme conditions suggesting the existence of higher poverty rates during this period.

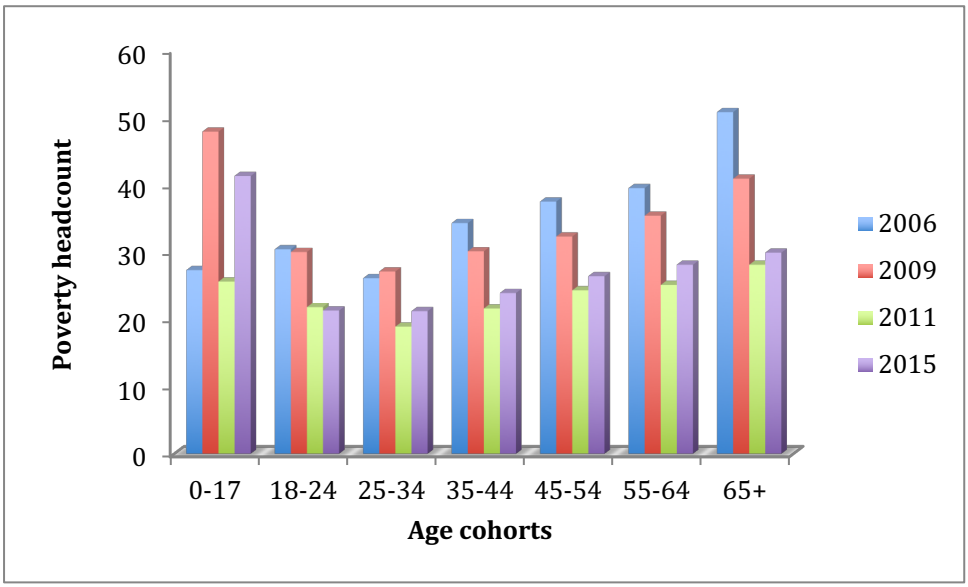
The GHSs have been used in a series of publications themed 'Social profile of vulnerable groups' in South Africa by Stats SA. The poverty profiles targeted youth (Stats SA 2016), children (Stats SA 2019a) and older persons (Stats SA 2017a). The analyses by targeting the most vulnerable groups enabled the household and personal circumstances surrounding them to be explored further.

The Living Conditions Surveys introduced in 2008/9 introduced the analysis of subjective poverty in South Africa by Stats SA and most importantly on how such estimates compare to objective poverty measures (Stats SA 2012a; 2018b). The findings show that applications of subjective poverty measures do not always coincide with monetary dimensions of poverty and where they do, it emphasises the presence of a severe form of poverty.

An analysis of poverty related literature from Stats SA would be incomplete without incorporating their work on the South African Multiple Poverty Index (SAMPI). Stats SA (2014b) used three dimensions of well-being; health, education and living standards and within these dimensions indicators including nutrition, child mortality, years of schooling, school attendance and other variables used as proxies for living standards were used. The findings from Stats SA using these dimensions show that residing in FHH, or in KwaZulu Natal, Eastern Cape and Limpopo provinces as well as in Black African households is strongly associated with poverty. Using both income poverty measures and SAMPI, these households remained poor.

The report on poverty trends by Stats SA (2017b) that utilised the IES of 2005/6; 2010/1 and the LCS of 2008 and 2014 has summary poverty statistics for South Africa during this period. Poverty reduced from 36.0 to 22.9 between 2006 and 2011 before increasing to 25.1 in 2015 for all South Africans (Stats SA 2017b). The following graphs summarise these findings after disaggregating by race, gender and geographic are for all household heads.

Figure 3.1: Poverty headcount for age cohorts (2006-2015).



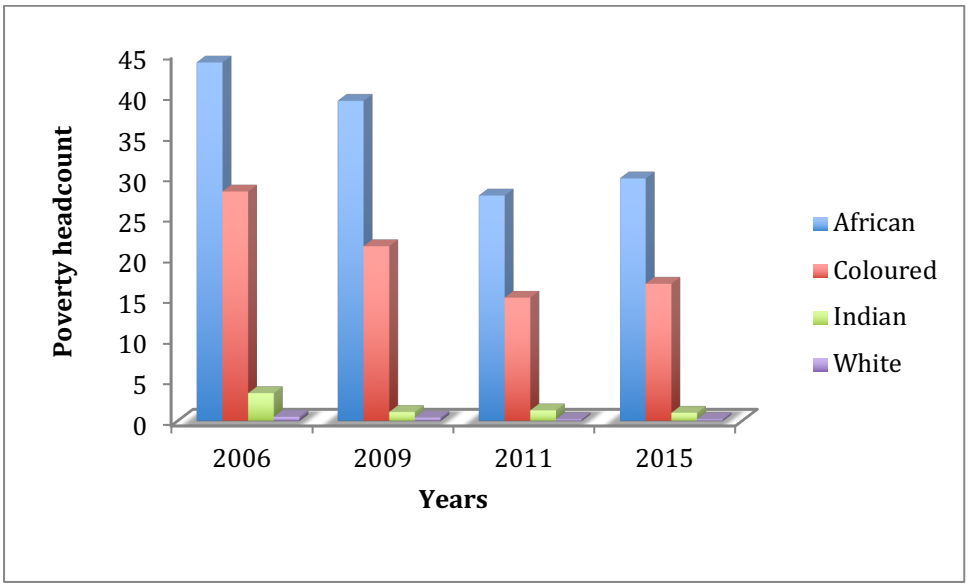
Source: Derived from Stats SA 2017b:92

Figure 3.1 above illustrates the poverty headcount rates for age[7] cohorts applied in Stats SA's poverty trends report (Stats SA (2017b). What is evident from this report is that the youngest age cohort (0-17) and the oldest age cohort (65+) experience the highest rates of poverty compared to other age cohorts. There was also a reduction of poverty between 2006 and 2015 for all birth cohorts except for the 0-17 age cohort and the 25-34 age cohort. In 2011 poverty reduced for all birth cohorts but increased again for all age cohorts except the 18-24 age cohort.

[7] The pseudo panel generated later in Chapter Five uses birth cohorts and not age cohorts to generate a pseudo panel. This decision is based on the fact that age cohorts change over time whilst birth cohorts remain the same. The assumptions applied in the pseudo panel therefore are favoured by this ability of birth cohorts to remain the same over time.

In Figure 3.2 below, between 2006 and 2015 Black South Africans and Coloureds were the most affected with poverty. What is also evident is that between 2006 and 2011, poverty declined for all races but slightly increased for the African racial group from 15.1 to 16.8% in 2015.

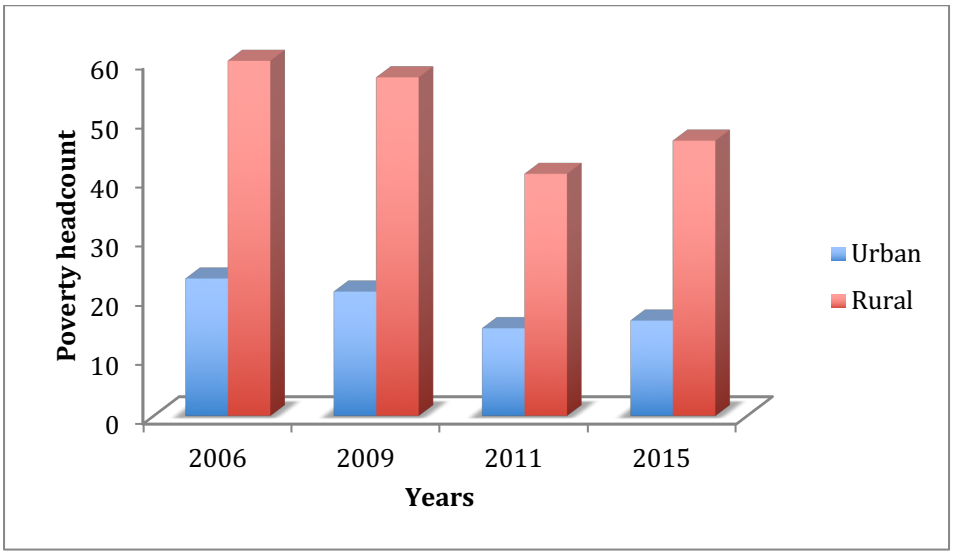
Figure 3.2: Poverty headcount by racial grouping (2006-2015).



Source: Derived from Stats SA 2017b:92

The divide between rural and urban areas is evident in Figure 3.3 below. Rural areas experience poverty rates that are more than twice the poverty rates in urban areas. Overall, poverty rates for both urban and rural areas reduced between 2006 and 2011. In rural areas, poverty reduced from 60.0 to 40.9 in 2011, whilst in urban areas the reduction was from 23.2% to14.8%. In 2015, poverty increased for all households in rural areas (from 40.9% in 2011 to 46.5% in 2015) and in urban areas (14.8% in 2011 to 16.1% in 2015.

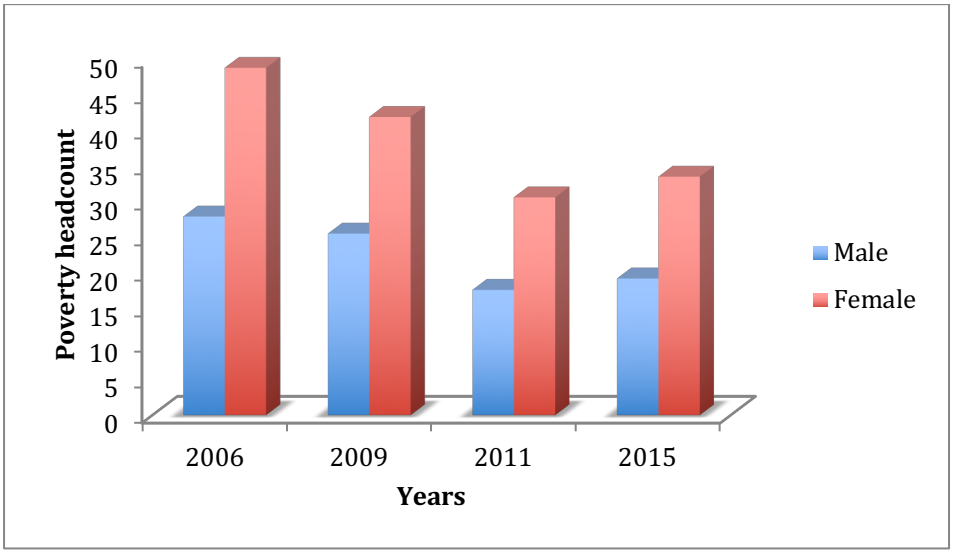
Figure 3.3: Poverty headcount for Urban and Rural households (2006-2015)



Source: Derived from Stats SA 2017b:92

Poverty trends recorded for FHH and MHH between 2006 and 2011 show that for FHH poverty reduced from 48.8% in 2006 to 30.6% in 2011, and from 27.9% in 2006 to 17.6% in 2011 for MHH. However, poverty rates increased between 2011 and 2015 for both MHH (17.6% to 19.2%) and FMM (30.6% to 33.5%) between 2011 and 2015.

Figure 3.4: Poverty headcount for MHH and FHH (2006-2015)



Source: Derived from Stats SA 2017b:92

3.9 Conclusion

This chapter has discussed the conceptualization of dynamic poverty and disaggregated poverty into transient chronic and structural poverty. The distinctions made between these types of poverty enabled the analysis of policy that is often used to deal with these

forms of poverty. This discussion was followed by an exploration of gender, race and geography as some of the dimensions of poverty that have been used in poverty analysis in developing countries, and more specifically in South Africa. What emerged was the existence of different policy regimes in dealing with these dimensions ranging from universal, targeted and a mix of the two policy directives.

The chapter then explored the use of three poverty lines in South Africa. Apart from showing the distinctions that exist between these three lines, the chapter summarised official poverty statistics. The main objective of the chapter in exploring the types and dimensions of poverty sought to highlight the need for more analysis in poverty analysis using South African data.

Chapter Four

Making sense of poverty statistics in South Africa

4.1 Introduction

This chapter extends the analysis of data issues encountered in poverty measurement to include discussions conducted with government statisticians working with official survey data used in poverty analysis in South Africa. The objective of this chapter is to understand what Statistics South Africa (Stats SA) has done and is doing as far as the measurement of poverty is concerned in the post-Apartheid period. The chapter captures the experiences of experts involved in the collection and processing of official poverty statistics and attempts to unpack how the central statistical agency in the country deals with issues relating data availability, timeliness, comparability, and frequency of poverty related data. The broad research interest is on understanding how the use of existing nationally representative surveys can be improved as an attempt to understand official poverty data from the perspectives of the official data producers. Ultimately, the chapter seeks to provide recommendations on how survey conduct can be improved as a way of presenting quality statistics for effective poverty measurement in South Africa.

This chapter is structured as follows; the first section begins by exploring the methodological decisions made in this chapter, and issues on the use of purposive sampling and key informants form part of this discussion. The second section analyses data collected using personal interviews with officials from Stats SA. The last section is a conclusion of this chapter and it notes the efforts of Stats SA and recommends the use of alternative data methodologies to provide new avenues of exploring poverty in South Africa.

4.2 Methodology

In social sciences, the main objective of research enquiry is to analyse questions based on the changes in human experiences and be able to understand these meanings. To achieve this, Jackson et al. (2007) argue that there is need to have a systematic way of analysing life in its various dimensions. Based on this line of reasoning, this first section of the chapter attempts to illustrate the research methodology choices made in this chapter to enable the qualitative research objectives explored in this thesis to be

fulfilled. The choice of a research approach, design, and methods is heavily dependent on the research aims and questions of the study. It is within this line of thinking that the methodology choices in this chapter were conceived.

4.3 Selection of research approach

There are three distinct types of research approaches; qualitative, quantitative, and mixed methods (Creswell 2009a; 2009b; 2014). Research in the past has dominantly been preoccupied with the selection of either qualitative or quantitative research approaches. The distinction between qualitative and quantitative research methods is well documented. Jemna (2016) lucidly defines quantitative research as studies that use figures and statistical analysis and are based on numerical measurements to test causal hypothesis. The fact that qualitative research seeks to give comprehensive descriptions of social phenomenon (Jemna 2016) is an indicator that these two research methods are not conflicting but rather co-exist (Newman and Benz, 1998).

Creswell (2009b) provides a clear distinction between these two approaches. He argues that the difference between the two approaches goes beyond the format of the data; instead it lies in the way the approaches are used and the purposes of such research. Therefore, an understanding of the differences between quantitative and qualitative research goes beyond the use of numbers. At the same time, making distinctions between the three research approaches; qualitative, quantitative and mixed methods, does not translate that the 3 approaches are discrete (Creswell 2014). Rather, Newman and Benz (1998) ague that qualitative and quantitative research approaches are not rigid and polarized. Creswell (2014) suggests that a study can be argued to be more qualitative than it is quantitative or more quantitative than it is qualitative, with mixed methods falling in between.

There have been debates on the superiority of both qualitative and quantitative research approaches spanning for a period exceeding a century (Johnson and Onwuegbuzie 2004). In understanding the differences between the debates that have dominated contestations between qualitative and quantitative enquiry, Johnson and Onwuegbuzie (2004:14) argue the existence of the quantitative purists' school of thought that criticizes qualitative research as lacking objectivity. The conflicting qualitative purists' school of thought criticizes the use of a passive and detached style of writing promoted by the opposing school of thought (Johnson and Onwuegbuzie 2004). This is part of an

extensive debate that surrounds the use of qualitative and quantitative methods in research.

Whereas quantitative and qualitative research methods are well known, Creswell (2009b:14) argues that mixed methods are 'less well known'. The use of mixed methods in research has however been broadly used in various disciplines for more than 20 years (Creswell 2009a, 2009b). This is in comparison to quantitative and qualitative research methods that have been in existence for a period exceeding a century. Starr (2016) defines mixed methods as an approach in research that involves diverse practices that combine qualitative and quantitative methods with the intention to exploit the strengths of each of the research approaches combined. The origin of mixed methods in research has been traced back to the work of Campbell and Frisk in the 1950s when they used multiple methods in understanding psychological traits (Creswell 2009b). Their approach was later extended by other researchers who aimed at neutralizing the limitations of either qualitative or quantitative approaches and the emergence of mixed methods was also rooted in the fact that existing research designs had their own weaknesses and bias (Creswell 2009b). This chapter uses the qualitative research approach in order to better understand the context within which the quantitative component of this thesis can be approached. The chapter thus provides the rationale for constructing a pseudo-panel from existing official statistics.

4.4 Research design rationale

A research design has been defined as 'the plans and procedures for research' (Creswell 2009:3) or the 'blue prints' for the collection, measurement and analysis of data inKothari (2004:32), or simply as the overall plan for connecting the conceptual research problems (van Wyk 2012). This is just of summary from a multitude of definitions, but in common their conceptualization of a research design is that it is a map that shows in detail how a specific research will be conducted.

The choice of a research design allows the integration of the various components of the study to be organized coherently and logically (USC, 2018). This choice in any study is of paramount importance and remains crucial in successful enquiry. Omair (2015) argues that one of the most basic decisions to be made in research is the selection of a suitable design that is appropriate to the aims and objectives of the research. The choice of a research design has been argued to include more than the decision to use

qualitative, quantitative or mixed methods. It also incorporates the selection of the 'strategies of design' and in qualitative research, existing strategies include narrative research, ethnography, grounded theory and lastly case studies (Creswell 2014). It is worth noting that different research designs may be suitable for a given research problem, but Omair (2015) articulates that this choice is mostly influenced by how the research questions are outlined. Kumar (2011) provides two functions of a research design; to identify and develop the procedures and 'logistical arrangements' that formulate a study and to provide a quality assurance that the selected procedures will adequately answer the research questions and objectives.

4.4.1 Key informants

This study used key informants. Parsons (2008:2) defines a key informant as 'a proxy' standing in their personal capacity or on behalf of their associates or organization they work for. In another study, United States Agency for International Aid (USAID) (1996) defines key informants as a group of in-depth qualitative interviews with 15 to 35 people whose selection is based on their knowledge of a subject of interest. Alternatively, Parsons (2008) has also defined key informants a 'non-random group of experts' who are knowledgeable about a given subject (Parsons 2008). The motivation in using key informants in this study was a result of the need of getting an overview of the data related challenges faced by the central statistical office in South Africa specifically. Parsons (2008) in support of the previous statement, states that the use of key informants allows for the collection of information that gives a 'broad and informative view' of a specific subject.

As expected, the use of key informants in qualitative research has its own advantages and drawbacks. The use of key informants allows for the collection of detailed descriptive information that is relevant to the subject. It also allows for the use of small sample sizes by focusing data collection on those assumed possessing particular insight into the object of research. Parsons (2008) states that the selection of key informants is not based on their representativeness of the broader population, but they are selected on their knowledge of a given subject. On the other hand, the use of key informants was without limitations and difficulties in this study. For instance, the scheduling interviews were difficult as the interviewees had busy schedules. There was also the risk of selecting respondents who are not well informed or may have conflicts of interest. Finally, there is a concern about ensuring the anonymity of respondents when

the pool of informants is small. In this chapter, the respondents was made anonymous. Clark (2006) argues that often in qualitative research anonymity is used to respects the privacy and identity of participants and is important in qualitative research.

4.4.2 Case Studies

Central to the research methodological choices made in this chapter, was the decision to use South Africa as a case study. Yin (1989:13) defines a case study, as 'an empirical inquiry that investigates a contemporary phenomenon within its real-life context especially in the boundaries between the phenomenon and the context are not clearly evident'. Merriam (2010) however warns that the definition of a case study is not necessarily an end product of an investigation and is more than the process involved in the conduct of the case study. Instead case studies should be defined as a unit of analysis that makes the case study and may refer to an individual, a group of individuals, or other subjects of enquiry.

There are many types of case studies that have been proposed over the years. Baxter and Jack (2008:547) distinguish between explanatory, exploratory, descriptive, collective, multiple, intrinsic, and instrumental case studies as some of the major types identified over the years. Bogdan and Biklen (2007) on the other hand state that case studies can be classified as historical, organizational, observational or as life histories. The choice on which types of case study to apply is influenced by the research objectives (Baxter and Jack 2008). Yin (1981a); (1981b) distinguishes between exploratory, descriptive or explanatory case studies, but Yin (1989) cautions existing distinctions between these three types of case studies as rather blurred since they all share a common definition.

The use of case studies is sometimes preferred in qualitative research. Jackson et al. (2007) argues that the choice is often ideal if the researcher is trying to understand 'how' and 'why' research questions and there is a general lack of clarity on the boundaries between the context and the phenomenon under examination. In a separate study, Yin (1989) argued that case studies are used when the researcher seeks to make analytic generalizations. In this case, analytic generalizations have been defined as 'a type of generalization where the researcher uses a certain set of circumstances such as a case to refine, dispute, support, or provide deeper detail about a concept, model, or

theory'. This study uses South Africa; a nation regarded as having one of the most efficient agencies in Africa, and thus offers a 'best case'.

4.4.3 Sampling method

Purposive sampling was used in this chapter. The choice to use key informants influenced this choice of sampling method. Purposive sampling is defined as a sampling that identifies and selects 'information-rich' cases for the most effective use of available resources (Patton (2002). Creswell and Clark (2011) further elaborate that purposive sampling identifies and select individuals who have specific knowledge and experience in a subject under study. Qualitative sampling is complex and is not a one step process; instead it involves iterative series of decisions throughout the research process (Guerterman 2015).

When it came to the selection of 10 key informants used in this chapter, the decision was influenced by discussions conducted in a number of studies. Firstly, there exists a rather limited group of studies that have investigated the determination of sample size in qualitative research. Existing work from Creswell (2013) and Morse (1994), have provided recommendations in this regard. Emmel (2013) explored the issue of numbers in qualitative sampling and concluded that the choice of the number of participants in qualitative research has more to do with the 'what for' than it is to do with the 'how many' question. Thus, the consideration is not necessary on the size of the sample but Emmel (2013) further states that there are practical reasons that guide research. The selection of the sample size in qualitative research are concerned with the fact that qualitative data turns to be bulky and after factoring in personal and practical limitations, one eventually draws the line on the size of the sample to work with (Emmel 2013).

The semi-structured interviews that were conducted in this study identified individuals with immense knowledge and experience in dealing with surveys related to this study. The main instrument for data collection was an interview guide.

4.4.4 Data collection summary

This study used open-ended semi structured questions administered face to face to collect information. The use of semi-structured interviews is contrasted to the use of structured interviews, which are more inclined towards quantitative research. The use of semi-structured interviews in this study allowed the researcher to probe on some

questions when responses were either incomplete or needed further enquiry. Edwards and Holland (2006) commended semi-structured interviews on their general flexibility and lack of structure. Edwards and Holland (2006:3) further clarify that semi structured interviews are 'thematic and topic centred' and allow the interactional exchange of information between at least 2 parties.

One of the merits of using semi-structured interviews is that they can be 'loosely structured' allowing for a broader subject matter to be addressed allowing for a 'free flow of idea and information' (USAID 1996:1). Cohen (2006) argues that semi structured interviews allows the informants chosen to express their personal views in their own vocabulary in contrast to selecting predetermined yes/no responses. One of the limitations that should be noted when using semi structured interviews as the sole data source is that generalizability of information to a larger population from key informants can be problematic (Edwards and Holland 2006). This is due to the fact that the conduct of such interviews involves social elements such as location, context and the social space in which they are conducted which cannot be replicated (Edwards and Holland 2006). In short Crow and Pope (2008) argue that the use of interviews involves a complex set of social relationships that cannot be replicated elsewhere. In this study, semi-structured interviews were given to statisticians from Stats SA during one-on-one sessions.

4.4.5 Data coding and analysis

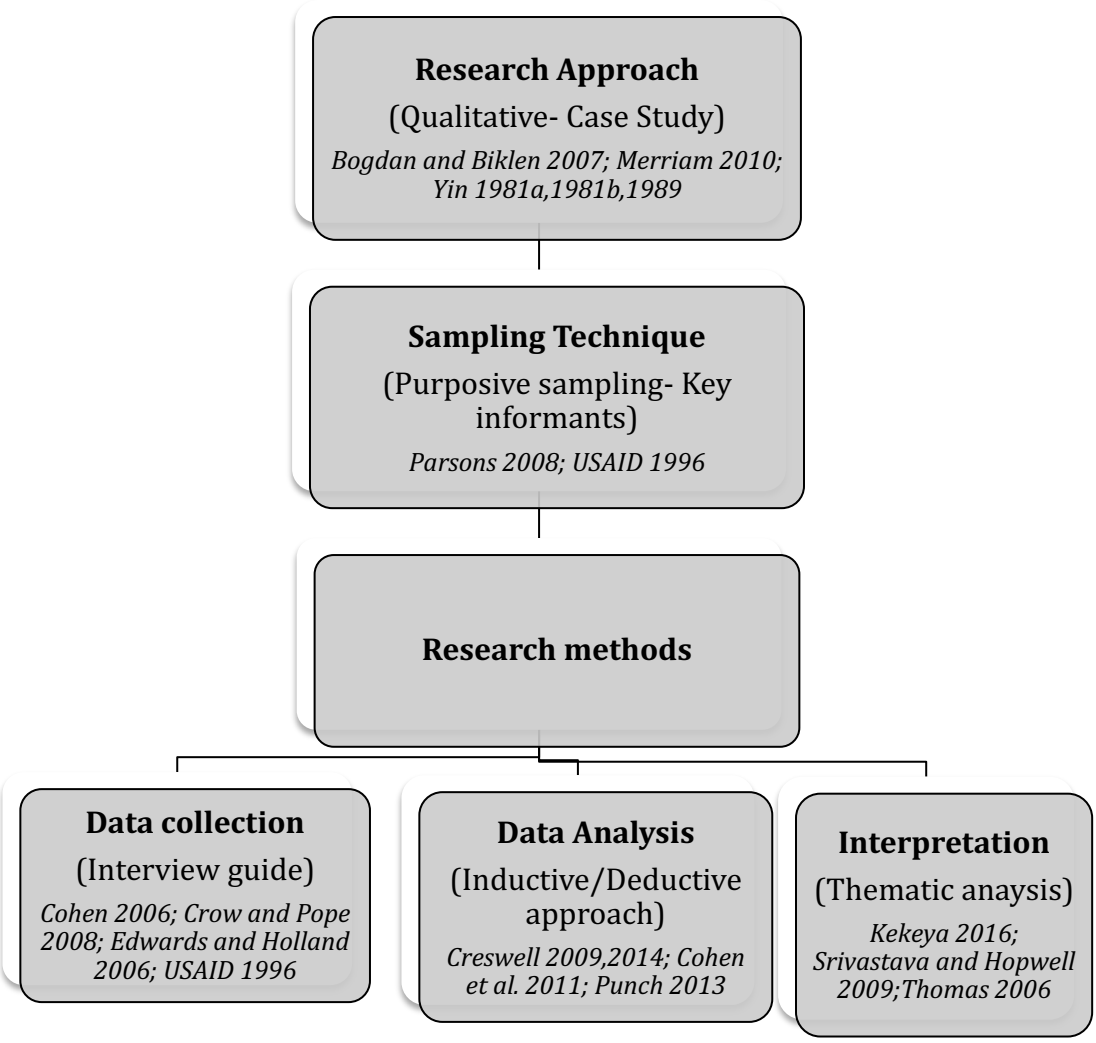
The qualitative data analysis process has been defined as 'a skilful interpretation of raw data, highly dependent on a systematic and rigorous approach' (Srivastava and Hopwood (2009:77). In existence is a multitude of qualitative data analysis techniques. For instance, Kekeya (2016) argues that there is a significant number of qualitative analytic styles in existence. But Srivastava and Hopwood (2009) are of the view that despite the existence of such a plethora of techniques, there exists 'few agreed upon' qualitative approaches where common ground is shared by a majority. In light of this argument, Thomas (2006) justifies that as much as qualitative data has been used for a long time, there exists a rather limited know how on the data analysis techniques. Despite the existence of a multitude of procedures used in qualitative analysis, Thomas (2006) is of the view that such approaches are largely generic and fails to fit in into specific qualitative research traditions.

The qualitative data analysis conducted in this thesis used both the inductive and deductive approach. This approach to data analysis applied in this chapter is discussed in Creswell (2014); Cohen et al. (2011); Punch (2013) and in its application it groups data into a set of themes. Deductively the approach then determines if more data is required or if the existing themes are adequate to answer the given research questions. The inductive approach can be defined as an iterative process that analyses data according to 'categories, patterns and themes' (Cohen et al. 2011; Creswell 2009b; Punch 2013). This study relied on detailed readings of the transcribed interviews and from these themes were derived and used as the mode of interpreting the raw data.

In justification of the choice of this approach, the debate in this research is based on the arguments made by Gabriel (2013). Since the qualitative side of this research was not meant to establish causality, the inductive approach was suitable since it allowed the generation of new understanding on the issues surround poverty related surveys data production in the country. Themes were generated on the changes in survey conduct in terms of the number of surveys, their coverage as well as issues on multidimensional poverty, future survey conduct amongst other themes applied in this chapter.

Figure 4.1 below summarises the methodology choices made in this chapter and it starts with the selection of a qualitative research approach to how the collected data was finally analysed.

Figure 4.1: Summary of Qualitative methodology applied in this chapter.



Source: Derived from cited sources

4.5 A brief history of statistical institutions in South Africa

Prior to discussing the findings from the key informants used in this study, this section takes a historical synopsis on how statistics were conducted in the pre-1994 period. The objective is to try and understand the foundations of Stats SA and get a better understanding of the historical conduct of statistics in the country. It is worth mentioning that the history of statistical institutions in most African countries in the pre-Colonial era remains sketchy and largely undocumented (Gardner 2018). Furthermore, Gardner (2018) is of the view that African statistical history received limited coverage especially when historical comparisons are made with European countries in the early periods. It is argued for instance that information on Gross Domestic Product (GDP) per capita for European countries started being collected as

early as the middle-ages, but in most African countries such data only started being available after the 1950's (Gardner 2018). As a result, limited written records restrict our understanding of African statistical history (Fourie 2016).

The majority of what is documented in African statistical history today is linked to colonial administration. A majority of African countries have a colonial past and it is within this history that their pioneer statistical offices were established. A majority of existing data that documents the statistical history of the country is therefore available mostly starting from the 19th and 20th century and such data were collected by Colonial institutions (Gardner 2018).

Former British colonies for example had statistical systems that replicated their colonial masters. In this regard, Christopher (2011) argues that statistical offices in British colonies were merely an exported version of the British Empire's statistical system and were part of the colonial venture. In line with this view, statistical offices in colonial Africa were a reflection of the colonizers' statistical style and were not designed primarily to provide comprehensive statistical coverage. The design of these statistical offices was more linked to colonial administration and as a result they fail to give accurate reports on African populations and demographics.

It would be inaccurate however to conclude that African states prior to colonialism had no statistical systems in place. Early records have shown the existence of statistical systems prior to the encroachment of colonialism. Such evidence exists in the Kingdoms of Kongo (Gardner 2018) as well as in some early kingdoms of Angola (Heywood and Thornton 1988). Early statistical work in Africa has also been linked to Missionary work but such data remains limited and prejudiced towards who got selected into the missions' stations that were set up (Fourie 2016). Statistical institutions in the pre-colonial era therefore remain sketchy but the information they collected has provided the current knowledge of the pre-colonial statistical systems, as we know them today.

In the Post-Colonial period, statistical systems in African countries had a different agenda. They became systems of accountability of the state and in this new era, statistics became a key element in the pursuit of development related policies in what became known as the 'development decade'. The newly elected governments sought to reverse the inequalities created by colonialism. In light of this argument, Heine and

Oltmanns (2016) state that in the resolutions made by these newly elected governments, official statistics became an indispensable part on the information system of the democratic state. It primarily sought to provide data to the government and the public on the economic, social, and environmental situation (Heine and Oltmanns 2016).

Statistical offices became crucial elements in providing the much-needed information to measure progress. In fact, they became central bodies in the development agenda as the data they provided was the yardstick in the measurement of progress made in attaining development goals. Their relevance was therefore rooted in their ability to provide the much-needed data to report progress to the state and donors that were sponsoring some of these agendas. Kratke and Byiers (2014) argue that at the onset of independence in Africa, state reforms, which included the independence of statistical offices with the objective of regaining ownership of the National Statistical Systems, took place.

Given the fact that statistical offices in most African countries were part of the colonial inheritance, independence exposed these massive gaps in data collection and as such African countries in the post-independence period had to play catch up. Christopher (2011:1) argues that central statistical offices are part of a 'major and powerful governmental enterprise' but has become a subject of inquiry. Therefore, it is not surprising that some of them are still in disarray today. Chapter Two explored in detail the changes that have been recorded in survey conduct in African countries and the developing world at large. Part of the conclusions noted an increasing number of cross-sectional surveys conducted in developing countries as well as a significant number of sub-panels relevant for poverty analysis.

Fourie (2016:202) states that transcribed colonial data continues to inform most of what is known of Africa's past. However, such data remains unreliable as its collection was biased against indigenous populations in favour of Europeans in what has become termed as 'colonial-era bias' (Fourie 2016:202).

In South Africa, statistical systems have a history that goes beyond the colonial era. Gouws (1987); Zuberi and Khalfani (1999) and Zuberi et al. (2005) are part of the limited body of research that documented some of the earliest demographic accounts of South Africa during the Colonial and Apartheid eras. Gouws (1987) and Zuberi and Khalfani (1999) trace the first Europeans to have settled in South Africa at the Cape

around 1652. Despite this early encroachment, the first official population census in the country was only carried out in 1865 (Zuberi et al. 2005).

As a former British colony, South Africa had its first official statistical offices established during the colonial era (Christopher 2011). This platform gave the colonial enumerators an opportunity to impose statistical order on the population using numbers and racial classification systems- an arrangement that became crucial to the colonial process (Christopher 2011). Lehohla (2005:49) states that the statistical system that existed before 1994 in South Africa 'produced a fragmented and deeply divided statistical practice', with statistical exercises divided between research agencies, the Central Statistical Service (CSS), homeland agencies, development agencies and academia. In an infamous quote, J.B. Moffat the then Director of Census in 1911 said,

'*All that is required for redistribution purposes is the number of European male adults in each province'* illustrating the level of bias that characterised survey conduct during this period.

The enumeration of only European population in the Population Censuses of 1918, 1931 and 1941 were therefore not alarming. Later analysis showed that the main concern of the colonial statistical office was to preserve the white population, specifically the European elite and to support this status quo, legislation such as the Population Registration Act of 1950 were passed (Zuberi et al. 2005).

An analysis of existing statistical work conducted during the pre-1994 period in South Africa also shows huge inconsistencies in survey conduct. For instance, Lehohla (2002:16) states that the apartheid era statistical system was deeply fragmented. During the colonial era, prior to 1910, census conduct in South Africa was not coordinated; different states conducted their own censuses under different census commissioners (Christopher 2011). It was only in 1910 when the Union of SA was created that all these states were combined, and issues of statistical record became more evident but still incomprehensive (Christopher 2011). The Census Act of 1910 necessitated the first coordinated efforts in enumeration from a centralized statistical office (Christopher 2011).

An analysis of the population censuses held in the South Africa before 1994 shows that the nominal 'states' of Transkei Bophuthatswana, Venda and Ciskei (TBVC) were

excluded. Census conduct during the colonial era in South Africa showed a pattern of irregularity and unsystematic coverage (Christopher 2011) and focused on a minority of an estimated 5 million white people (Lehohla 2002).

In 1910, a census office was in South Africa and between this period and 1960 a total of 10 censuses were conducted, but race overshadowed all other issues as it determined the inclusion and exclusion of certain areas based on race (Christopher 2011). It should be noted however, that the availability of statistical work during the colonial period in South Africa is marred by the unavailability of adequate paperwork to the public, a factor also noted by Christopher (2011). As a result, little or nothing is known on the actual poverty statistics that prevailed in the country during the colonial and apartheid era. Magasela (2005) states that the official statistical categories that existed in the pre-1994 period were combined with racial attitudes in the collection and representation of data.

In 1991, the infamous Population Registration Act of 1950 was repealed - a cosmetic move by colonial administration that sought to 'appease' the African race, the trade unions and other disgruntled groups rather than addressing the racial bias in existence (Zuberi et al. 2005). The 1991 Population Census as much as it was not comprehensive, had better coverage than previous censuses. Specifically, it managed to capture 64% of the South African population, but, only 54% of the African population were counted (Zuberi et al. 2005;15).

It is against this background that Stats SA was launched in 1995, an office that inherited a system aligned along issues of race and segregation. In 1999 with the new Statistics Act, Stats SA mandated to conduct comprehensive statistical work that supported the country's development trajectory. In the post-independence period, Lehohla (2008) states that similarly, African countries after regaining their independence started building statistical infrastructure and towards the end of the 21st century most of them showed evidence of rapid gains in statistical development.

The Population Census of 1996 therefore was one of the first mandates of Stats SA in the post-Apartheid era. The TBVC states that had been excluded entirely in the 1980, 1985 and 1991 Population Censuses were enumerated in 1996, however language barriers, accessibility issues and other issues resulted in some South Africans not being counted (Zuberi et al. 2005). However, the improvements in enumeration translated to

93% of the South African population being counted (Zuberi et al. 2005), marking a significant improvement in the conduct of a population in South Africa.

4.6 The evolution of poverty measurement in South Africa

In the period after 1994, Stats SA has made meaningful changes in poverty measurement. An acknowledgement of these institutional changes that occurred within the organization is a good starting point for this section. Annual reports from Statistics South Africa showcase these changes witnessed in the past 24 years. Institutional capacity is one of the most evident changes within Stats SA. The 2006/2007 Annual Report from Statistics South Africa shows that the Population and Social Division only had three sub-divisions; Population Statistics, Social Statistics and Health and Vital Statistics.

As surveys expanded and became more subject specific, these divisions expanded. In 2017, the Population and Social Statistics Division had six departments as new departments such as the Poverty and Inequality Division and a Labour specific division was added to the department. During an interview conducted on the 17th of May 2017, one of the key statisticians in the Poverty and Inequality Statistics division summarized how the existing Poverty and Inequality Statistics Division was created. In his narrative, what started as a project team to conduct the IES survey in 2005/6 ended up becoming a poverty specific team that became housed under the Poverty and Inequality Division. Currently, the Poverty and Inequality Division at Stats SA is a fully-fledged unit and focuses on the collection of data that allows the multidimensional measurement of living conditions including poverty and inequality.

Altogether, Stats SA has seven programs; National accounts, Social statistics, Poverty and inequality statistics, Demography, Health and vital statistics, and Labour statistics (Stats SA 2018b). The Population and Social statistics programme under which the Poverty and inequality division is housed had 174 staff members from a total of 3 023 permanent staff members for the entire organization during the 2017/18 reporting period (Stats SA 201b8).

The Poverty and Inequality division provides statistics on poverty, inequality and expenditure trends of South Africans (Stats SA 2018b). The division also conducts analysis and to date have, apart from the statistical releases associated with the release

of every household survey and amongst their achievements, the construction of poverty maps and the South African Multidimensional Index (SAMPI) using South African household surveys and census data.

4.6.1 Increase in the number of household surveys

One of the most noticeable indicators in the evolution of poverty statistics in South Africa is the increase in the number of household surveys relevant for poverty measurement in the period after 1994. In one of its earliest official poverty reports, Stats SA acknowledged the existence of limited data for the measurement of poverty in the country between 1994 and 1999 (Stats SA 2000). At this period, the datasets available for poverty analysis included the IES of 1995, a Population census of 1996 and the OHSs of 1994, 1995, 1996, 1997 and 1998 were the only data points available in before 2000.

In another interview, another senior member of staff summarized this increase in the conduct of household surveys in the post-2000 period to include the IES, GHS and LCS household surveys and stated this surge as an important cornerstone in the agency's ability to gauge poverty in the post-1994 period (Respondent Two 2017, personal communication, 22 May). In response to this increasing number of data sources in poverty analysis, there has also been an increase in the number of programs whose main objective is to analyse poverty trends in the country. For example, Oloo (2011) states that in the post-apartheid period several seminal reports were produced as part of poverty reduction efforts from different commissions. These included the Key Indicators of Poverty in South Africa in 1995, the Poverty and Inequality Report in 1996, the Participatory Poverty Assessment South Africa Report of 1998 as well as an increasing number of policies and frameworks on poverty reduction from government (Oloo 2011). Table 4.1 below summarises some of the key poverty related publication from Stats SA.

Table 4.1: Summary of Stats SA publications on poverty in South Africa

Year of publication	Title of publication	Survey used
2000	Measuring poverty in South Africa	IES 1995, Census 1996
2001	South Africa in transition: Selected findings from the OHS of 1999 and the changes that have occurred between 1995 and 1999	OHS
2002	Poverty Report: Measuring poverty in South Africa	
2006	Deprivation: The provincial indices of multiple deprivation for South Africa 2001	IES, OHS, LFS, KIDS
2007	A discussion note: Constructing comparable household survey data for the analysis of poverty analysis in South Africa (1995-2000)	OHS 1995, 2000, LFS 2000, IES 1995, 2000
2008	Measuring poverty in South Africa: Methodology report on the development of the poverty lines for statistical reporting	IES 2005/6
2010	Social profile of South Africa (2002-2009)	GHS 2002-2009
2011	Social profile of vulnerable groups in South Africa (2002-2010)	GHS 2002-2010
2012	Subjective poverty in South Africa: Findings from the Living Conditions Survey 2008/9	LCS 2008/9
2012	Poverty profile of South Africa: Application of poverty lines on the Living Conditions Survey 2008/9	LCS 2008/9
2013	Social profile of vulnerable groups 2002-2012	GHS 2002-2012
2013	Men, women and children: Findings from the Living Conditions Survey 2008/9	LCS 2008/9
2014	The South Africa Multidimensional Poverty Index: Creating a multidimensional poverty index using Census data	Population Census 2011
2015	Income dynamics in poverty status of households in South Africa	Population Census 2001, 2011
2015	Methodological report on rebasing of national poverty lines and development of pilot provincial poverty lines: Technical Report	IES 2010/11
2016	Vulnerable Group Series 1: Social profile of youth 2009-2014	GHS 2009-2014
2017	Poverty trends in South Africa: An examination of absolute poverty between 2006-2015	IES, 2005/6, LCS 2008/9, IES 2010/11, LCS 2014/15

2017	Living conditions of households in South Africa: An analysis of the household expenditure and income data using the Living Conditions Survey of 2014/15: Statistical Release P0310	LCS 2014/15
2017	Social profile of older persons	GHS 2011-2015
2017	Vulnerable Group Series II: The social profile of older persons 201-2015	GHS 2011-2015
2018	National poverty lines: Statistical release P0310.1	IES
2018	Subjective poverty in South Africa: Findings from the Living Conditions Survey of 2008/9 and 2014/15	LCS 2008/9, 2014/15
2018	Men, women, and Children: Findings of the LCS of 2014/15	LCS 2014/15
2018	Poverty mapping in South Africa: Applying Small Area Estimation techniques using the IES 2010/11 and Census 2011	IES 2010/11, Population Census 2011
2018	Vulnerable Groups Series III: The social profile of children aged 7-17	GHS 2002-2016

Source: Stats SA website

Table 4.1 above shows that before 2000, there are no/few publications specifically on poverty from Stats SA. Moreover, in support of its efforts in poverty reduction in the post-apartheid period the South African government is a signatory to several poverty reduction commitments such as the 1995 Copenhagen Declaration, the recently ended MDGs and the current SDG commitments in poverty reduction.

In an illustration on the challenges experienced in providing data for the MDG indicators, Respondent One said the following,

'Stats SA only managed to come with an official poverty estimate in 2005 since 1994. During the MDG era, we only managed to provide data at 3 data points only for poverty analysis, 2006, 2009 and 2011. The IES and the LCS of 2008/9 provided about 70-72% of the data required for the poverty related indicators.'

An analysis of the surveys used on MDG Goal 1a also showed that there were comparability problems over the years. The Director of Poverty and Inequality Statistics at Stats SA highlighted the following,

'During the MDG period there was a lack of a solid base line for poverty analysis and this only came in 2005/6. The methodologies we were using before 2005/6 were not directly comparable but we have been working on providing more comparable poverty surveys after 2005.'

Stats SA (2015c) in its Final Report of Goal 1 of the MDGs concurs to the fact that they experienced data comparability problems in providing data for the indicators of Goal 1. In understanding this challenge, Stats SA (2015c) states that the main causes of this challenge were rooted in the use of different methodologies of the surveys used as well as a change in the way questions were asked in the different years of similar surveys. Incomparability issues also emanated from the fact that different adjustments were made in weighting, in dealing with missing values and in the treatment of unspecified values.

Therefore, the increasing number of poverty related surveys in the post-1994 period in South Africa is more evident in the post-2000 period as shown in Table 4.1 above. This has also been accompanied by a corresponding increase in the usage of such data by both researchers and policy makers and is an indication of an evolution in poverty measurement in the country. The post-1994 period marked a clear change in survey conduct and methods and may not have solved all the poverty related enquiry as existing household survey data still have issues but marks a move towards better poverty measurement.

4.6.2 Evolving data collection methods

Data collection methods at Statistics South Africa have also recorded significant improvements in the post-1994 period. One of such changes is the expansion in sampling frames. Household surveys that were conducted in the country using the Population Census of 1991 had incomprehensive sampling frames. Orkin (1996:2) notes that the Population Census of 1991 excluded the TBVC states and it also excluded other 'self-governing territories' and the Population Census of 1991 also had Enumeration Areas that were not clearly defined.

In a personal interaction with Respondent One, he indicated that in the post-1994 period one of the biggest achievements for Statistics South Africa was an improvement of the sampling frames used in household surveys. Specifically, he said the following;

'*When it comes to issues like poverty for example, we know that the problem predominantly affects the black African majority and old sampling frames that were used were not very representative of the black areas'* (Respondent One 2017, personal communication, 17 May).

The resulting effect of incomplete/incomprehensive sampling frames is that they fail to capture poverty accurately. Ultimately such data are ineffective in informing poverty reduction policy. The sampling frames used in later surveys improved on their representativeness of the country's demographic make-up.

The methodology related improvements were also recorded in the data collection methods used at Statistics South Africa. In a detailed account, Respondent One highlighted that the move away from the use of recall methods only to collect expenditure and income data to the inclusion of diaries was an important milestone in the collection of data in poverty related household surveys (Respondent One 2017, personal communication, 17 May). Respondent One also mentioned that in 2005/6 the diary method was incorporated into the IES survey conducted in a similar time frame and was later adopted for later household surveys (Respondent One 2017, personal communication, 17 May).

Linking this improvement in data collection methodologies at Statistics South Africa to data issues, a general underreporting of income or alternatively its overestimation, is expected to significantly reduce in the case where it was caused by the use of recall methods only in the collection of expenditure or income data. If we assume that households that underreported or overestimated income were affected by recall bias, then the incorporation of diaries significantly alleviates this problem. Research is therefore still to be done on how the introduction of diaries affected the problem of non-response in South African household surveys.

Close to this line of argument, Respondent One further explained in detail how the agency further reduced the diary keeping periods from one month to two weeks, in a move to try and minimize item non-response and a general underreporting of income (Respondent One 2017, personal communication, 17 May). We therefore also expect item non-response to reduce even more if respondents were leaving out sections in the diaries as a result of fatigue. Respondent One stated that one of the immediate challenges encountered when working with income and expenditure related surveys was the fact that they were demanding in terms of time needed by households to furnish all the required details.

'Currently, households participate for about six weeks with a two-week diary keeping period. The diary keeping is the tough part and prior to this we used to have a one-

month diary-keeping period. After looking at our statistics from the diaries the responses declined in quality after the first weeks. We said how can we streamline the data collection process to reduce the burden on households without compromising the methodology that we had established? That change definitely helped with responses it was productive and more balanced tool that allowed us to still maintain the information we had to get from the people, but it was now less demanding' (Respondent One 2017, personal communication, 17 May).

Respondent One further highlighted that the adoption of the 2-week diary-keeping period was cost effective. For example, he went on and elaborated that one-month diaries would require an estimated 500 field workers to collect data. The reduction to 2 weeks translated to a reduction of fieldworkers to about 360 and this was a significant reduction in costs incurred during data collection.

The fact that diaries are a self-enumeration tool also has the drawback that it depends entirely on the ability of respondents to accurately report their expenditures and income accurately (Respondent One 2017, personal communication, 17 May). The reduction from monthly diary keeping periods to 2 weekly diaries apart from reducing the burden on households also made it easy to input data as respondents would only focus on a shorter period of diary keeping and give better statistics resultantly (Respondent One 2017, personal communication, 17 May).

4.6.3 Embracing multidimensional poverty

Another significant milestone in the measurement of poverty in South Africa was witnessed by the gradual shift from the use of pre-dominantly money-metric based approaches to incorporate multidimensional forms of the concept. One such move was the introduction of the Living Conditions Survey (LCS) in 2008. Respondent Three argued that the introduction of the LCS in 2008/9 marked a shift in poverty statistics; it extensively incorporated the multidimensionality of poverty and moved beyond the previously dominant monetary survey approach (Respondent Three 2018, personal communication, May 29). The introduction of the LCS enabled the analysis of other multidimensional forms of poverty with existing modules capturing relative and subjective poverty. To date the LCS has been conducted twice, in 2008/9 and in 2015/16

Respondent Four acknowledged that Statistics South Africa was more biased towards one-dimensional approaches to poverty in the early years (Respondent Four, personal communication, 16 May 2017). In fact, she argued that social statistics released from the organization in the early days were not always capturing multidimensional aspects of poverty. Instead the focus was more on the production of yearly specific thematic reports.

In a similar argument, Respondent Two highlighted that one of the changes witnessed in poverty measurement was the movement from absolute poverty measurement which was more money based to incorporate the multidimensional aspects of poverty measurement. These decisions included the introduction of the LCS and the introduction of multidimensional poverty dimension indices (Respondent Two 2017, personal communication, May 22).

Respondent One also supported this shift and argued that whereas surveys such as the IES allowed for the collection of detailed income and expenditure surveys, it could not give an adequate poverty profile of the affected households (Respondent One 2017, personal communication, 17 May).

Apart from the introduction of the LCS, there have been other efforts at Stats SA to incorporate poverty in its multi-dimensions. One of the earliest efforts in embracing multidimensional poverty was the adoption of the Multidimensional Poverty Index (MPI) to the South African Multidimensional Poverty Index (SAMPI) (Respondent One

one, two, and three). Respondent One specifically mentioned that these efforts have been funded by UNISA to train staff to incorporate the MPI as proposed by Alkire and Foster (2007).

Whereas the Global MPI is made up of three major dimensions; health, education and standard of living (OPHI 2015), the SAMPHI was tailor made to suit the South African context to include health, education, standard of living and economic activities. Respondent One argued that the SAMPHI team at Statistics South Africa after deliberations added another dimension to the MPI that would capture unemployment. The four dimensions encompassed in the SAMPI had 11 indicators. The indicators captured income, type of floor, years of schooling, number of unemployed adults as

part of the key dimensions of the SAMPI and they arguably captured better the poverty dimensions in the country (Respondent One 2017, personal communication, 17 May).

4.6.4 Evolution of poverty lines in South Africa

Understanding the changes in poverty analysis in South Africa is incomplete without factoring in how the use of poverty lines has evolved. Chapter Three discussed the arguments on how the existing three poverty lines in South Africa came into place and how they differ. There seems to be a consensus at Statistics South Africa that the Lower Bound Poverty Line (LBPL) is the preferred line in poverty measurement by policy makers. In a discussion with the second Respondent, she stated that,

'The Lower Bound Poverty Line is the preferred poverty line by Statistics South Africa since it gives a better sense of who is in poverty' (Respondent Three 2017, personal communication, 26 May).

In support of this line of thinking, the fourth Respondent also argued that the Lower Bound Poverty Line is more sensitive to policy and as such, interventions are better informed using this line since it identifies households most in need of assistance (Respondent Four 2017, personal communication, 16 May). This use of the LBPL as the preferred measure for poverty in South Africa aligns with the country's National Development Plan (NDP) that has also adopted this line in its official reports.

Despite the agreement in policy circles that the LBPL is the preferred line, the Director of Poverty and inequality statistics pointed out that the use of three poverty lines caters for the different data users of poverty statistics in South Africa,

'The use of the Upper Bound Poverty Line makes poverty figures look really big and as a result politically the preference is to use the lower bound poverty line which gives more preferable figures. Therefore, the use of three lines caters for the different users of poverty statistics ranging from academics to policy makers' (Respondent One 2017, personal communication, 17 May).

Respondent One stated that other countries on the other hand use only one poverty line. Instead the use of three poverty lines in South Africa allows for poverty to be measured differently and gives a good range on the different profiles of poverty experienced in the country (Respondent One 2017, personal communication, 17 May). Respondent

Two also supported that other countries use one poverty line and the use of three lines allow for a more comprehensive understanding of poverty in the country.

Outside Stats SA, the usage of the LBPL as the 'preferred' poverty line has received criticism. Budlender et al. (2015) are of the view that the LBPL is not convincing- as households are likely to sacrifice basic food for non-food purchases at different degrees until their expenditure is on the upper poverty line. On the upper poverty line, households are able to meet both their food and non-food basic needs without sacrificing the food component (Budlender et al. 2015).

Bearing the critic from Budlender et al. (2015) in mind, the LBPL is not altogether void. The choice to use poverty lines using the Basic Needs Approach when constructing poverty lines already involves arbitrary elements in poverty measurement. Allowing distinctions to be made on the various degrees of deprivation allows for the different depths of poverty to be identified. Though it is difficult to provide a sound theory on the compromises a household has to make on the purchase of food and non-food items, the minimal amount pegged at the food poverty line remains theoretically unjustified. This study argues that the LBPL shows different degrees of poverty and allows poverty analysis to show these differentiations.

4.6.5 Analysis of poverty dynamics

Poverty measurement in South Africa has also evolved beyond the analysis of poverty trends and has incorporated the analysis of poverty dynamics. The existence of the KwaZulu Natal Dynamic Study (KIDS) and National Income Dynamic Study (NIDS) is a clear indication of changes in this direction. However, the KIDS and NIDS are not conducted by Statistics South Africa, yet our knowledge of poverty dynamics in the country has dominantly come from these two panel surveys. The consensus from the respondents from Statistics South Africa acknowledged the fact that apart from the LFS and the QLFS that have a panel component, they have not conducted any panel survey that is relevant to poverty analysis. But they acknowledged the existence of the KIDS and NIDS as meaningful and useful efforts in the analysis of poverty in the country.

Respondents one and two acknowledged that there were no plans from Statistics South Africa to conduct a panel survey in poverty analysis. In an interview conducted on the 16th of May 2017, Respondent Five, the Chief Director of NSSD, highlighted that the GHS and QLFS remain the closest that the agency has come when it comes to panel

data. The GHS and QLFS have rotating samples and allow similar households to be visited every years (Respondent Five 2017, personal communication, 16 May).

Similar arguments were also maintained by Respondent Two, and she further suggested that the plan was to bring the NIDS to Stats SA in terms of data collection so that it enhances the instruments available to the agency for data collection in terms of poverty. But she indicated that this was part of an on-going discussion. (Respondent Two 2017, personal communication, May 22).

What was clear from the personal interviews I conducted at the agency was the acknowledgement on the importance of NIDS in poverty analysis in South Africa. The Director of Poverty and Inequality Statistics indicated that the organization supported the Presidency in the NIDS projects and acknowledged that panel survey data was relatively more expensive to conduct when compared to cross sectional household surveys (Respondent1 2017, personal communication, May 17 May). Respondent One also stressed that it was not necessarily appropriate for the organization to have a competing panel survey that goes up against NIDS (Respondent One 2017, personal communication, 17 May). Instead he suggested that efforts would rather focus on strengthening that tool and working on improving existing efforts. In a more conclusive discussion with Respondent Four on the issue of a panel survey for poverty analysis, she indicated that this was not feasible given existing budgets and there were no plans in the near future for such a survey to be conducted.

Existing poverty literature in South Africa concurs that there exist transient, chronic and structural forms of poverty in the country. In South Africa, the analysis of dynamic poverty has confirmed the findings from other developing countries, but some of the outcomes are context specific to South Africa. For instance, due to the high dependency on paid employment in the country, causes of transient poverty are often linked to the availability of employment for individuals in a household. For instance, Du Toit (2005a) found that households that escape transient poverty are more likely to have secured employment. Similar conclusions were also drawn by Cichello et al. (2003) who found that the mobility of households in and out of transitory forms of poverty in South Africa are not a result of changes in demographic variables but are related to the earning volatility of households

The existence of structural forms of unemployment in the country has also been found to be a cause of chronic poverty (Woolard and Klasen 2005; Schotte et al. 2018). Du Toit and Neves (2008:1) also found that in South Africa, 'social exclusion and adverse incorporation' are the key drivers of chronic and structural poverty. The most vulnerable groups include the rural dwellers with no land or the unemployed urban poor. Given the fact that land continues to be controlled by a few and the fact that South Africa has been experiencing a jobless growth, it then becomes apparent that there is a serious problem in designing policy to effectively deal with chronic and structural poverty in the country.

More specifically, du Toit (2005a) found that in South Africa, transitory poverty was associated with the labour market marginality as well as a lack of social and cash transfers. This brings out an immediate weakness in our understanding of transient poverty, a household that secures a temporary job may exit poverty for a brief period, but this is not an accurate reflection of dynamic poverty. In the case where the job is seasonal, the same household will slide back into poverty after the job ends.

In South Africa, colonialism and apartheid were systems designed to restrict mobility especially on the African population and chronic poverty was a result of the lack of opportunity to accumulate human and other forms of capital compounded by the lack of access to basic services and infrastructure (Aliber 2003). Despite the fact that Apartheid ended in 1994 the vestiges of the system can still be felt. Findings using panel data such as the KwaZulu Natal Dynamic Survey (KIDS) found chronic poverty to exist amongst the rural poor. Aliber (2003) states that Africans and Coloureds living in rural areas are more likely to be chronically poor in South Africa than other races. The incidences of chronic poverty are also more pronounced for female-headed households and the elderly, a direct result of the land dispossession that occurred in the colonial and apartheid period (Aliber 2003).

4.7 Challenges

The discussions in the second half of this chapter have attempted to show the changes that have affected how poverty is measured in the country. The findings have illuminated a general improvement in survey conduct over the years, especially if comparisons are made on existing household surveys in the pre-1994 with those conducted in the post-1994 period. Yet gaps still exist here and there and various issues

continue to be associated with available data. To partly understand these gaps and weaknesses of existing data, this section discusses the challenges faced by Statistics South Africa and an analysis of these issues may illuminate the roots of current data problems.

4.7.1 Funding and infrequency of survey data in South Africa

Issues associated with infrequent survey conduct in the country are tied to funding constraints. The Director of Poverty and Inequality Statistics supported the fact that the conduct of household surveys is an expensive enterprise;

'*Household surveys are expensive to administer, each with close to 30 000 households and a cost amounting to a couple of hundred million. Our reduction to use the diary to two weekly was a cost saving* initiative' (Respondent Three, 2017, personal communication, 29 May).

In fact, Respondent Three indicated that,

'Statistics South Africa would benefit from conducting the IES and LCS more frequently preferably twice a year instead of the usual 3 to 4 years. Funding would go a long way in improving the frequency of existing surveys as well as improving the depth of the statistics collected' (Respondent Three, 2017, personal communication, 29 May).

The LCSs seem to be at the receiving end of these funding constraints. For starters, Respondent Three said that the LCS at first was to be conducted every 3 years.

'The LCS was supposed to have been conducted in 2012/13 but due to budget constraints in funding we only were only able to conduct it in 2014/15' (Respondent Three, 2017, personal communication, 29 May).

In a separate interview, the first Respondent Two also echoed similar thoughts, she argued that the unavailability of adequate funding often results in the periodicity of surveys being affected. This explains why sometimes surveys are not released on the expected dates (Respondent Two 2017, personal communication, 22 May).

The second Respondent also indicated that the LCS did not have its own funding allocation as other surveys and the 2008 and 2014 LCSs relied entirely on savings from

funds allocated to other surveys (Respondent Two 2017, personal communication 22 May).

Funding issues have also affected the conduct of Demographic Household Surveys and this has resulted in delays in their release (Respondent Two 2017, personal communication 22 May).

To attempt to understand how the agency funded itself, Respondent Two mentioned that Statistics South Africa received its funding from the National Treasury and the same department also is responsible for funding other government departments (Respondent Two, personal communication 22 May). Therefore, funding for household surveys depends on budget allocations from the National Treasury, and whilst established surveys such as the IES and GHS might have timely budgets, it is clear that the introduction of new household surveys is problematic.

4.7.2 Issues with the Population Census data

At the heart of household surveys, are sampling frames and ultimately the Population Censuses where they are derived from. The Mo Ibrahim Foundation (MIF) (2016) argues that population censuses form the basis of any statistical analysis in a country but, are often heavily politicized. This is due to the fact that in some cases they are used as the basis of budget allocations and may reveal sensitive issues on delicate issues related to ethnic or national minority (MIF 2016). In South Africa, funding issues have mainly affected the conduct of the Population Censuses. In the post 1994 period, the Census has been conducted in 1995, 2001 and 2011.

In line with the stipulations of the Statistical Act of 1999, after 2011, the next Population Census was expected to have been conducted in 2016, but following a cost-benefit analysis conducted by Spencer et al. (2016), the South African government decided not to conduct a population census in 2016 as expected. Rather a decision was made 'to improve data and capacity for producing post-censal estimates' (Spencer et al. 2016:36).

In another analysis May and Lehohla (2005) found that the Population Census conducted in South Africa in 2001 was perhaps the most expensive census in the history of the country. Part of the reason on the high costs of the 2001 population census was a result of an increase in costs of census taking exercises on a growing population,

coupled with smaller and more complex family structures among other reasons (May and Lehohla 2005). Ultimately, the 2016 Population Census in South Africa was not conducted and was replaced by the Community Survey.

But despite this, there have been improvements in the conduct of Population Censuses in the country. An analysis of Census data in the country shows some advances especially as far as coverage is concerned. The 2001 and 2011 Population Censuses showed huge gains in enumeration as they counted people in all areas in the country when compared to the 1996 Census. Moreover, Oloo (2011) notes that the 1996 and 2001 Censuses have contributed significantly to poverty analysis in the country and have been used extensively in poverty reduction programs and policies. An analysis of the sampling frames of the household surveys conducted in the 2000s in South Africa show a significant improvement in the coverage and as such became more relevant in poverty analysis. Currently, Stats SA is preparing to conduct the 2021 Population Census.

Despite these improvements, previous chapters highlighted issues associated with Census data in South Africa especially on the magnitude of zero income households. In response to a question on the conduct of Census data, the fourth respondent acknowledged the existence of coverage, high refusals, and a general underreporting of income in Population Censuses in South Africa (Respondent Four 2017, personal communication, 16 May). It is therefore imperative that analyses using the Population Censuses in the country require vigorous checks and various imputation methods to allow for meaningful poverty analysis.

4.7.3 Data disaggregation

The discussions with Respondent Four also showed that one of the challenges faced by the agency was that existing major household surveys cannot be used to make analysis at lower levels of disaggregation, i.e. districts, and wards where they are more likely to provide more meaningful analysis. She argued that we need to move beyond the fact that one province is poorer than the other to more specific geography-based statistics. But she also indicated the poverty mapping has been used by the agency to locate poverty at such low levels of disaggregation. She stated that,

'One of the weaknesses of existing surveys is that they do not provide information disaggregated to district and ward level, which are more meaningful units for policy

formulation and implementation' (Respondent Four, 2017, personal communication, 16 May).

Statements from Respondent Five and Six also supported the lack of properly disaggregated data for poverty analysis (Respondent Five 2017, personal communication, 16 May; Respondent Six 2017, personal communication, 16 May). Respondent Three,

'*One of the weaknesses of existing surveys is that they do not provide information disaggregated to district and ward level, which are more meaningful units for policy formulation and implementation'.*

Respondent Two also stated the importance of the agency producing more disaggregated data meaningful to policy,

'*We now are faced with a need of developing statistics that are relevant to SA's development related challenges. We need to go beyond simple enumeration of the unemployed or poor but produce data that is disaggregated at lower levels that is more useful to policy' (*Respondent Two 2017, personal communication, 22 May).

The importance of issues related to data disaggregation is highlighted in the 2030 Agenda. UN (2016) states that in this Agenda, countries committed to meet specified disaggregation elements as stipulated in the SDG indicators. In the agreement reached by the Inter-Agency and Expert Group on SDG Indicators, they agreed that, '*The SDG indicators should be disaggregated, where relevant by income, sex, age, ethnicity, migratory status, disability and geographic location, or other characteristics, in accordance with the Fundamental Principles of Official Statistics'*. (UN 2016:7).

One of the earliest poverty mapping research in South Africa emerged in the early 2000s in Alderman et al. (2001) where they combined the Population Census of 1996 and the IES of 1995.

4.7.4 Quality and comparability of survey data

The production of quality statistics is at the forefront of Agenda 2030 and as such ANSOs are facing a new challenge. The Africa Union in its Agenda 2063 acknowledges the need for improvements in data availability. The Strategy for the Harmonization of Statistics in Africa (SHaSA) acknowledges the importance of data in African National

Statistical Offices (NSOs). In light of this, SHaSA seeks to promote the promote the production of timely, reliable and harmonised statistics as well as better coordination in the production of statistics among African countries.

In South Africa, Statistics South Africa uses the South African Statistical Quality Assessment Framework (SASQAF) as the basis for the production of quality statistics. Statistics South Africa (2010) stipulates that this framework 'provides clear criteria and transparent procedures for evaluation of official statistics and other data'.

Respondent Six mentioned 'Data *quality in Statistics South Africa is measured using the SASQAF framework. All the dimensions of data produced is supposed to conform to the dimensions of the SASQAF'.*

SASQAF was first released in 2008 and was a response to the increasing data requirements in the country. The framework has also been used as a self-assessment tool by data producers' and is used in certifying official statistics as indicated in the Statistics Act on 1999 (Statistics South Africa 2010). The SASQAF maximises on the following data dimensions; accuracy, relevance, quality, timeliness, accessibility, interpretability, comparability, methodological soundness, and integrity as the cornerstones for the production of official statistics in South Africa.

However, one of the challenges faced in the production of poverty statistics in South Africa is that not all data producers adhere to the requirements of the framework.

'*But one of the most immediate challenges we are facing is other government agencies that also produce statistics for the country. Of all the government departments that are providing statistics of the country, the Department of Environmental Affairs and the Department of Science and technology have gone through quality assessments of their output. Most of the data sources in this country do not have the sufficient infrastructure in place to ensure data quality but it is something we are working on currently.*' (Respondent Six 2017, personal communication, 16 May).

In support of the same arguments, Respondent Five argued that:

'We need to improve the implementation of our NSS to bring to par all data collection and processing methodologies of all agencies that produce official statistics. The main challenge is that for data produced outside Statistics South Africa *we can only*

recommend that the data conform to SASQAF. But this is something we are changing in the Statistics Act and we hope to provide adequate infrastructure necessary to ensure data quality' (Respondent Five 2017, personal communication, 16 May).

4.8 Future survey conduct

The future of poverty measurement at Statistics South Africa is anchored in the introduction of the Continuous Population Survey (CPS), a new survey that will incorporate various modules from existing household surveys. Respondent One for instance argued that the CPS is part of the organization's strategy to improve frequency of household survey conduct and its creation is argued to be a long-term strategy that would increase the frequency of poverty statistics as well as the depth of poverty analysis:

"We are planning to improve the existing surveys we want to merge the IES, LCS, GHS and the Community survey into one tool which we call the Continuous Population Survey. It would run on a two-year cycle' (Respondent One 2017, personal communication, 17 May)

Respondent Four elaborated the operationalization of the CPS to be two-way; in the first year, the CPS will collect municipal level data and in the other year, the data would be collected for the LCS and IES. The introduction of the CPS also suggests that household surveys such as the LCS and IES would be discontinued. However, the CPS was expected to begin in 2018 as a new vehicle for poverty measurement and this is yet to be officialised (Respondent Four 2017, personal communication).

The CPS is argued by Stats SA to be a better survey to measure poverty. Dr Schmidt indicated that its use would allow disaggregation to be conducted at municipal levels and would also be conducted more frequently, after every two years. Documentation for the CPS continues to be largely unavailable on the public domain, statistical releases on the cps are still to be released into the public domain by Stats SA. Stats SA (2018b) states that the CPS incorporates the GHS, LCS and IES, with changes being made in the GHS of 2014 to incorporate part of the sections in the LCS.

Respondents from Statistics South Africa also indicated the move away from paper-based data collection tools for household surveys, to the use of computer-based systems. Respondent Two indicated that the organizations had invested in training of

its staff members for the usage of the Computer-Assisted Personal Interviewing (CAPI). She estimated that the CAPI would be used in the 2020 round of Census in South Africa. The use of CAPI seeks to minimize human error in data collection and also reduce fieldwork related costs.

Stats SA (2017d) used the CAPI to collect data for the Community Survey, this reduced data collection costs from R2000 for every household if a paper-based system had been used to R200 using the CAPI. The use of the CAPI also reduced the turnaround time, the results for the CS for 2016 were released two months after data collection with the results argued to be more accurate (Stats SA 2017d).

4.9 Conclusion

This chapter began by discussing on the methodological choices made to pursue the qualitative research questions in this thesis. The discussion included the use of a qualitative research approach as well as the use of South Africa as a case study. The use of key informants in this chapter also justified the use of purposive sampling and the use of inductive and deductive data analysis methods. In the second section, the findings from the qualitative enquiry were presented and the developments that have been recorded at Statistics South Africa as far as poverty measurement is concerned were discussed. The last paragraphs discussed some of the challenges encountered by the organization as well and within these challenges, some of the issues identified in Chapter Four to be affecting poverty analysis were revisited and their causes investigated.

Poverty measurement in South Africa during the post-1994 period has evolved considerably, and the changes observed include an increase in survey conduct as well as a general improvement of survey design and methodologies. This does not mean that available official cross-sectional household data that has been used in poverty measurement are without issues. Instead, this chapter acknowledges existing efforts from the country's central statistical office. This chapter drew on my interactions with statisticians involved in the production of poverty related household surveys from Statistics South Africa. A look at existing household surveys and their methodologies shows that poverty measurement has improved and more surveys relevant to poverty measurement have been produced. This study acknowledges that we know more about

poverty trends and dynamics now than we did in the pre-1994 period. But, available poverty relevant data in South Africa are not adequate and are not without criticism.

This chapter therefore acknowledges existing efforts from Stats SA and at the same time also accepts that gaps do exist in poverty measurement. To support an argument made by the respondents used in this study, there is need for more analysis using existing data to uncover new research in poverty analysis. Bearing this in mind, this chapter argues that there indeed is a place for synthetic panels in poverty analysis in South Africa to attempt to complement what we already know in poverty dynamics in South Africa. It is therefore the contribution of this thesis to analyse poverty dynamics using a pseudo panel. The following chapter gives the methodology used in the generation of this pseudo panel.

Based on the qualitative data analysed in this chapter, it is evident that the colonial and apartheid era statistical conduct did not provide a solid base for statistical conduct in South Africa. The Colonial and apartheid era statistics did not take account of the socio-economic situation of the bulk of the country and left the post-apartheid era with the task of rebuilding a credible statistics system.

Much has been done with more and better data, and more and better poverty statistics. The improvements made in the post-1994 period especially in the period after 2000, showcase an improvement in survey conduct especially as far as comprehensive statistics are concerned. Issues such as a lack of comparability of household surveys still remain, but the introduction of CAPI and the CPS show efforts in trying to mitigate human error in data handling, moves that are relevant in the 2030 and Agenda 2063 era.

Stats SA is keenly aware of the importance of poverty statistics and has steadily grown their capacity and the coverage and frequency of data collection. However there remain significant challenges and shortcomings, especially related to the availability of sufficient funding. Although new initiatives are planned, these will not provide data on the dynamics of poverty, leaving the country dependent on a single panel study, the NIDS. As a result, innovative use of existing cross-sectional data is required, thus justifying the use of alternative data sources in this thesis.

Chapter Five

Constructing A Cohort Based Pseudo Panel For Poverty Analysis: Data Issues And Methodology

5.1 Introduction

The primary contention of this thesis is that the analysis of poverty dynamics in developing countries is hampered by a general paucity of panel data. Chapter Two discussed in detail the inadequacies of both cross sectional and panel survey data in poverty analysis in Sub-Saharan Africa and specifically in South Africa. It is also important to note that due to the multidimensionality of poverty, it would be costly to have data that allows for the dynamic analysis of poverty in all its diverse forms comprehensively. Thus, where data requirements are not being met, data imputations, and simulations provide possible options to achieving poverty analysis research objectives.

Synthetic/pseudo panel methods are one option that allows the analysis of poverty dynamics in the absence of genuine panel data. It is within this line of thinking that this study creates a pseudo panel for the analysis of expenditure patterns and poverty rates of birth cohorts using cross sectional household survey data from South Africa. A new 'panel' data source for poverty analysis in South Africa is thus created using three rounds of cross-sectional household survey data from the Living Conditions Surveys of 2008 and 2014 and the Income Expenditure Survey of 2010.

The previous chapter concluded that Statistics South Africa (Stats SA) has significantly increased the number of household surveys relevant for poverty analysis in the country in the post 2000 period. However, researchers, policy-makers, and Stats SA staff agree that more needs to be done in terms of analysis of existing survey data. It was also argued earlier that despite this significant increase in the number of surveys available for poverty analysis, incomparability and infrequency issues affect the analysis of existing household surveys. In light of the need to prioritize national budgets and to allocate resources between sub-national structures, there is increasing demand for high quality and frequent data and the need for extensive analysis of existing surveys to meet existing research gaps. This chapter is an attempt to move away from the response that 'it is out of the scope of this study'/'more research is needed' frequently encountered

in policy input to a solution-orientated approach in dealing with data issues associated with poverty research in social sciences. This chapter explains how the pseudo panel used in this study was constructed and the discussion also illustrates the decisions made in the choice of the cross-sectional surveys used in the generation of the pseudo panel.

5.2 Constructing a cohort-based pseudo panel

Pseudo panels have enabled economic analysis to be conducted at cohort level (Collado 1993; Deaton 1985; Deaton and Paxson 1994; Guillerm 2017; McKenzie 2004; Moffit 1993; Rentz et al. 1983) and at household/individual level (Bourguignon and Moreno 2015; Bourguignon et al. 2004; Dang et al 2014; Dang and Dabalen 2017; Guell and Hu 2006). In this study I generate a pseudo panel that would enable the 'tracking' of birth cohorts in three time periods using household heads born between 1924 and 1988..

The use of cohorts instead of genuine panel data has been merited in pseudo panel literature. For example, cohort distributions in pseudo panels are argued to give more accurate summary statistics when compared to individual data from genuine panels and are sometimes preferred in longitudinal studies since they are not affected by attrition and non-response (Deaton 1985; Deaton and Paxson 1994). In another dimension, the use of birth cohorts also allows for the pseudo panel to have a relatively large sample size in the rounds where the birth cohorts are followed when compared to genuine panels. This is a considerable attribute in South Africa given the small sample problems associated with the use of studies such as the NIDS data in poverty analysis. Riley et al. (1972) also attributed cohort-based pseudo panels on their ability to simultaneously enable ageing and within cohort differences as part of the merits of this type of data.

The LCS of 2008 and 2014 and the IES of 2010, used the diary and recall methods to collect expenditure data as an attempt to offset the shortcomings of using a single approach. Stats SA (2017), argues that the introduction of the diary and recall methods in household surveys such as the Living Conditions Survey of 2014/15 reduced non-response rates. Furthermore, the reduction in the diary-keeping period from one month to two weeks introduces in the IES of 2010 and also used in the LCS of 2014 further minimized respondent fatigue and 'improved diary-keeping' Statistics South Africa (2015:1).

5.2.1 Description of pseudo panel approach

The approach used to generate the pseudo panel in this chapter is derived from the methods applied in Deaton (1985). That 'pseudo panels observe cohorts' as argued in Guillerm (2017:109), influenced the choice of the pseudo panel methodology approach applied. Birth cohorts are generated for household heads born between the years 1924 and 1988. In total, 13 birth cohorts were created with a five-year band and these are 'tracked' in three time periods.

5.2.2 Assumptions and matching techniques

The definition of a cohort as applied previously in pseudo panel literature has been followed. Cohorts have been defined as 'a group with fixed membership, individuals who can be identified as they show up in the surveys' (Deaton (1985:109) and as 'a set of observed characteristics that are stable over time' (Guillerm 2017:110). In this study, a cohort refers to groups of individuals who are born in similar years and are spaced with five-year intervals that are followed in three time periods 2008, 2010 and 2014. Following this definition, the following assumptions were applied;

i. The immediate assumption emanating from the definitions given above is that households headed with individuals born in the same birth cohort are identical[8] in terms of their aggregate expenditure. Therefore, a birth cohort can be observed in three time periods and is assumed to be the same in all rounds.

ii. Secondly, the approach used follows the use of group means as cohort observations. This follows the approach adopted by Attanasio and Weber, (1993); Blundell et al. (1994); Browning et al. (1985); Collado (1997) and others in which sample means are treated as consistent estimators of the true cohort population mean and are presented as observations of a genuine panel.

[8] This assumption assumes intracohort homogeneity, as similarly applied in other pseudo-panel literature. Hsaio (2003: 284) lists intra-cohort homogeneity as perhaps one of the key assumptions in the generation of synthetic panels, yet this might be problematic in some settings especially if there are factors that can violate this assumption. For example, in South Africa, the issue of HIIV/AIDS could have affected cohorts differently over the years and could have introduced intra-cohort heterogeneity. Stats SA (2018g) for instance shows that HIV/AIDS tends to affect the 15-49 age cohort more than younger (15-24 age cohort) or older cohort (above 49). There has been an increase in the prevalence of HIV/AIDS in South Africa, the population of persons infected with HIV/AIDS rose from 4.25 million in 2001 to 7,25 million in 2018 (Stats SA 2018g). HIV/AIDS also tends to affect women with a majority within the Black African population group. This study notes the possible heterogeneity that is likely to be introduced within the birth cohorts used in this study that can be a result of the HIV/AIDS trends in the country, but intra-cohort homogeneity assumption continues to hold.

iii. In this study the average expenditure and poverty rates observed for a specific cohort applies to all 'members' in that specific cohort and is assumed to be different from the mean of the next birth cohort. Whereas in a genuine panel, individuals or households are identifiable using unique identification numbers, in a cohort, parameters are chosen that represent similar characteristics; in this case the pseudo panel identifies 13 birth cohorts separated by five-year bands.

iv. Another assumption applied in this pseudo panel is that each household head can only participate in one cohort. This assumption was also applied in Ryder (1965:844), in (Guillerm 2017:113) and in (Reintz et al. 1983). In their application of this assumption, Reintz et al. (1983:12) argue that 'members of any cohort are entitled to participate in only one slice of life - their unique location in the stream of history'. Using data from the LCSs and IES of 2010, it is possible that this assumption may get violated in the case where the same household heads were interviewed in more than one time period. It is possible, though highly unlikely that a household was interviewed in 2008 as part of the LCS in that year and either re-interviewed in 2010/1 and/or in 2014. The LCS of 2008, 2014 and the IES in 2010 used a Stats SA's Master Sampling Frame derived from the 2001 Population Census. The methodology applied in this thesis therefore assumes that there are no such households and all households in the 2008/9, 2010/11 and 2014/15 time periods are not the same.

5.3 Weighting the pseudo panel

The pseudo panel applied household level population weights in the analysis of expenditure patterns and poverty status of the birth cohorts. Since the analysis done in the pseudo panel seeks to maintain the distribution of the genders and racial groups of household heads as well as the spread of birth cohorts between urban and rural areas, household weights and not individual weights were appropriate. The household weights specified by Stats SA for each of the 3 surveys are thus used. The birth cohorts generated in this thesis follows a group of household heads born within the specified five-year bands.

5.4 Application of poverty lines

The main purpose of the pseudo panel generated in this chapter is to conduct an analysis of poverty transitions of birth cohorts in South Africa between 2008 and 2014. The sampling strategies used in the three survey rounds used in this chapter have large enough cohorts that are representative at cohort level. Figures 5.1 and 5.2 earlier in this thesis illustrated the proximities in cohort size between the birth cohorts constructed in this thesis and those from the latest mid-population estimates from Stats SA. In light of this broad objective this study uses the Lower Bound Poverty Line (LBPL) from South Africa and this will be applied on the Foster Greer Thornbecke (FGT) poverty indices.

This study acknowledges the existence of three poverty lines in South Africa; the Food Poverty Line (FPL), the Lower Bound Poverty Line (LBPL) and the Upper Poverty Line (UBPL). Whereas both the UBPL and the LBPL incorporate costs associated with a basic living standard, the UBPL goes beyond basic costs to include more living costs making the LBPL the preferred poverty line in some scenarios especially if the analysis seeks to incorporate a basic living standard. This study adopted the values of the LBPL for the years 2009, 2011 and 2015 given in a recent Stats SA poverty lines update in Stats SA (2019b:3-4). The poverty lines are all inflated to 2014 prices.

The fact that Statistics South Africa acknowledges that that LBPL is preferred in policy making is justification to use it in this thesis. The use of the LBPL is also in line with the country's National Development Policy (NDP), the National Planning Commission (NPC) applies the same poverty line as its official poverty line to report poverty in the country.

5.5 Age restriction of birth cohorts

The pseudo panel generated in this chapter restricted the age of the birth cohorts to try and capture dominantly the working age population. An allowance was made to accommodate household heads that had just joined or were about to join the labour market (born in 1984-1988) as well as those that had left the labour market (born in 1924-1928) to allow the analysis to capture how entry and exit into the labour market affects expenditure and poverty transitions of the birth cohorts. Moreover, despite the small size of the informal sector in South Africa, its existence in the country makes it difficult to cut off the active labour force of the birth cohorts to coincide with the

country's official retirement age. Additionally, a significant portion of households headed by older individuals is supported by social grants.

Different age restrictions have been applied in other pseudo panel analyses. In a recent study, Dang and Dabalen (2017) argue that the 25-55 age range restriction has become the 'traditional' age range applied in pseudo panels. To translate this age restriction to birth cohorts, the pseudo panel generated in this chapter would have to generate birth cohorts for household heads born between 1953 and1983 (to capture the 25-55 age range in the first year of analysis). The argument is that in this age range, the household heads' educational attainment remains relatively stable over time so that the changes observed in labour market earnings are not necessarily affected by an increase in educational attainment. However, the acknowledgment by Dang and Dabalen (2017) that age restrictions are highly dependent on cultural and are context specific justifies the choices made in the pseudo panel generated in this chapter.

5.6 Selecting a household survey for the pseudo panel

One of the key issues discussed in this chapter not only concerns the generation of the pseudo panel, but also deals with the decisions leading to the selection of household surveys used in the construction of the pseudo panel used in this study. The selection of a household survey perhaps marked the beginning of a succession of key decisions that had to be made before the construction of the pseudo panel. The choice, however, was highly dependent on the availability of at least two household surveys that were relatively comparable. Dang and Dabalen (2017:2) indicate that two 'comparable' cross sectional surveys are the minimal requirement in the generation of a pseudo panel.

Apart from the consideration of existing cross-sectional household surveys, pseudo panels have also been created by treating genuine panel data as independent cross sections under a set of given assumptions; Bourguignon and Moreno (2015) and Cruces et al. (2013) pursued pseudo panel approaches to this end. In this study I however only considered cross sectional household surveys for the pseudo panel and genuine panel[9] data were not considered.

[9] Genuine panel in this study refers to the National Income Dynamic Study Survey (NIDS) or the KwaZulu Natal Income Dynamics Study (KIDS).

Table 5.1 below summarizes existing income and expenditure surveys that were considered for the creation of the pseudo panel in this study. The table also provides key information on the conduct of the mentioned surveys in the post-Apartheid period in South Africa.[10] The table summarizes the years of availability of the surveys, their coverage, and their types.

Table 5.1: Survey availability for poverty analysis in South Africa

Survey	Years of availability	Type of survey	Coverage
Income and Expenditure Survey (IES)	1995 2000 2005 2010/11	Cross sectional household survey	Nationally representative
General Household Survey (GHS)	2002-2017 (Annual)	Cross sectional household survey	Nationally representative
Community Surveys (CS)	2007 2016	Cross sectional household survey	Nationally representative
KwaZulu Natal Income Dynamics Survey (KIDS)	1993 1998 2004	Panel survey	Covers KwaZulu Natal province only
Living Conditions Survey (LCS)	2008 2014/15	Cross sectional household survey	Nationally representative
Labor Force Survey (LFS)	2000-2007 (Bi-annual)	Rotating panel household survey developed to replace the OHS	Nationally representative
National Income Dynamics Survey (NIDS)	2008 2010 2012 2014 2016	Panel survey	Nationally representative
October Household Survey (OHS)	1993-1999	Cross sectional household survey	Nationally representative

[10] This study excludes surveys conducted in the pre-1994 period in South Africa due to their incomplete sample frames. They may be mentioned but no detailed discussions will be conducted. Surveys that are not nationally representative were also not considered for this study.

Survey	Years of availability	Type of survey	Coverage
Quarterly Labor Force Survey (QLFS)	2008-2017 (Quarterly)	Rotating panel household survey	Nationally representative
Cape Area Panel Study (CAPS)	2002 2003/4 2005 2006 2009	Panel	The sample is restricted to the metropolitan area of Cape Town
All Media Products Survey (AMPS)	1974-2015 (Annual and semi-annual)	Cross sectional	Data collected only from metropolitan areas

Source: Datafirst (2019)

As evident in Table 5.1, South Africa has a significant number of expenditure household surveys that have been used over the years in poverty analysis that could be used in the construction of a pseudo panel. Yu (2013) concurs that in the post-apartheid period in South Africa, more surveys relevant to poverty enquiry have been available than in the preceding periods. This makes the generation of a pseudo panel promising, as the choice of the household survey to be used is not necessarily limited by data availability.

The General Household Surveys (GHSs), Income Expenditure Surveys (IESs), Community Surveys (CSs) and Living Conditions Surveys (LCSs) were considered for the construction of the pseudo panel from the onset. GHSs were an option given the fact that they have been conducted annually since 2002, they collect income and expenditure data and they are nationally representative. A pseudo panel using the GHSs would arguably provide a pseudo panel that had coverage for a longer period when compared to other income and expenditure surveys available.

A closer analysis on the key variables of interest within the GHSs shows that using the available income and expenditure variables would make such a pseudo panel problematic. The GHSs do not use one-shot questions to collect income and expenditure data. Instead income bands with small intervals are used and this reduces the usefulness of any distributional analysis that emanates from such bands. Furthermore, the overrepresentation of households in the lowest expenditure categories

worsens this scenario (van der Berg 2007).

Van der Burg et al. (2008) used the All Media Product Survey (AMPS) as an alternative data source in poverty analysis. Their findings show consistency within this survey in capturing aggregate income and to be more consistent than existing income and expenditure surveys, as reflected in the country's national accounts. However, the AMPS are not designed to measure poverty and as such the surveys do not capture effectively individual and household characteristics, which are key in designing a pseudo panel for poverty analysis. Therefore, the AMPS could not be used for the generation of the pseudo panel in this study.

Moving on to the Community Surveys (CSs) of 2007 and 2016, these were conducted to fill in data gaps between the conduct of the Population Censuses of 2001 and 2011. The CSs could not be used in the creation of the pseudo panel since they did not collect any consumption data and have limited information on income, which hampers the ability to make a meaningful money metric poverty analysis (Stats SA 2012b).

5.7 Selection of the LCS surveys and IES of 2010

This study uses three household survey rounds, the LCS of 2008 and 2014 and the IES of 2010. Whereas other household surveys capture extensively the income and expenditures of households, the LCS was designed primarily to capture poverty trends in the country; this was the one of the key reasons why the LCSs were selected in this study. The introduction of the LCSs was an acknowledgement by Stats SA that there existed a gap in poverty relevant survey data in the country and its introduction was an attempt to cover this gap. In fact, Stats SA (2012b) concedes that prior to the introduction of the LCS in 2008 poverty analysis was dominantly reliant on surveys whose design were not specifically intended to analyse poverty. The agency maintains that despite the availability of the above-mentioned surveys in South Africa in the pre-2008 period, there was a 'serious gap' in poverty statistics since none of them had been specifically designed to analyse poverty (Stats SA 2012b:47). In 2008, the LCS became the first household survey in South Africa to collect data that is comprehensively specific to poverty analysis (Stats SA 2012b).

It therefore makes for a meaningful argument to attempt to create a pseudo panel to analyse poverty transitions of households using a household survey specifically

designed to measure poverty in the first place. The LCS in comparison to similar expenditure and income surveys such as the GHS and CS also has relatively more modules on income and expenditure. The LCS of 2008 and 2014 were used in the pseudo panel. Additionally, the IES of 2010 was used as another round in the pseudo panel.

To justify the inclusion of the IES of 2010 into the pseudo panel, the availability of the LCS in two time periods meant the pseudo panel would be limited in terms of time coverage. Recent pseudo panel literature cited earlier in this chapter concludes that two comparable cross-sectional surveys can be used to generate a pseudo panel. Despite this, the existence of a comparable round of the IES in 2010 provided an opportunity to use another time period in the analysis.

The IESs were primarily designed to provide information to update the country's Consumer Price Index (Stats SA 2012b), and not necessarily to monitor poverty. But the presence of a detailed income and expenditure modules in the surveys have also contributed to what we know about income poverty in South Africa in the post-1994 period and makes them ideal for monetary poverty measurement. The IESs unlike other household surveys conducted to date in South Africa have extensive income and expenditure data. Van der Berg et al. (2008) found that the IESs have a large enough sample size and can produce point estimates on income and expenditure more commendably.

The IES to date has been conducted in 1995, 2000, 2005/6 and in 2011. Whilst this would have enabled a pseudo panel that allows for birth cohort to be tracked between 1995 and 2010/11, a reasonable time frame to understand poverty dynamics, IES data are marred with comparability problems. For instance, the IES of 1995 and 2000 used the recall method only to collect data and the diary method only started being used together with the recall method in 2005 and later in 2010.

The first IES was conducted in 1990 and it only collected information from 12 urban areas in South Africa making it incomprehensive (Datafirst 2019). In 1995 the survey was conducted again, and it collected information from the same 30 000 households that were visited in the October Household Survey (OHS) of 1995 and it used the 1991 Population Census to draw its sampling frame (Stats SA 1997). However, in the IES report for 1995, Stats SA acknowledged that when the IES of 1995 was conducted the

actual numbers of households in South Africa were unknown; instead they only knew the number of questionnaires that had been completed in the Population Census of 1996 (Stats SA1997). The income and expenditure estimates derived from the IES of 1995 therefore remained mere indicators of the patterns of expenditure and not representative of the country's expenditure (Stats SA 1997).

Further restrictions within the household questionnaire of the IES of 1995 make it more problematic. The IES of 1995 only allowed a maximum of 10 people in a household to participate in the surveys and information collected on income was restricted to only five people per household (Yu 2008). Therefore the 1990 and 1995 IESs are not directly comparable; the IES in 1990 was only conducted in 12 main urban areas whereas the 1990 IES collected data from the whole country (Stats SA1997).

There also exist major incomparability issues with the IESs of 1995 and 2000; Bhorat and Kanbur (2005) cited huge gaps between the total household income estimates and the national accounts data. Simkins (2005) concluded that the property income estimates derived from the IES of 2000 were too low. Lastly Hoogeven and Ozler (2006) show that the IES of 2000 under-reported the white population group whilst over-reporting the black population group.

Stats SA (2008b) found that comparability of the consumption expenditure of the IES of 1995, 2000 and 2005 is difficult. In 2005/6, for instance, the IES for the first time used an 'imputed rent' variable. Previously, mortgage payments had been used to estimate housing expenditure and this new change makes housing expenditure not directly comparable. The introduction of the diary and recall methods in the IES of 2005/6 further compounds on existing comparability challenges. Despite the fact that the IES of 2005/6 also used the diary and recall method to collect information, it was excluded from this study because it used the diary method to collect information on the purchase of non-durable items only.

Bearing these limitations associated with IES data conducted in the pre-2010 period in mind, only the IES of 2010/11 could be used in the pseudo panel. The IES of 2010 was added to the pseudo panel based on the fact that it drew its sample from a similar master frame as the LCSs and its use of similar data methodologies made it an alternative

addition. More significantly, the IES of 2010 has been used in poverty related enquiry[11] in South Africa and this makes a good argument on its addition in the pseudo panel.

5.8 Comparability of the IES 2010 and the LCS of 2008 and 2014

The sampling frames of the three surveys used to construct the pseudo panel were all obtained from the same Master Sample derived from the 2001 Population Census of South Africa. This partly accommodates the comparability requirement suggested in Dang and Dabalen (2017:2). It is not out of place to reiterate that one of the most important assumptions in the generation of a pseudo panel as stated earlier in this chapter is that the cross-sectional household surveys used should be comparable. There are 25 075 households in the LCS of 2008, 23 380 households in 2014 and 25 328 in the IES of 2010. Although the number of households in each of these surveys is not identical and represent differing rates of sample attainment, they are sufficient to yield comparable poverty estimates. The number of households in each of the cross sections is therefore within an acceptable range to generate birth cohorts that can be 'tracked' in three time periods and give comparable results. Table 6.2 below summarizes key information on the LCS of 2008 and 2014 as well as the IES of 2010.

Table 5.2 below shows that the IES of 2010 and the LCSs of 2008 and 2014 all used the diary and recall methods in collecting expenditure data from households. However. In 2008, four household diaries were used and households would keep them for one week. In 2010 and in 2014, the number of diaries was reduced to two, and to compensate for this, households kept the diaries for two weeks. Whilst a one-week diary was kept in 2008, giving it a shorter recall period, this also meant a household had to enter expenditure information weekly. On the other hand, two diaries kept over a period of two weeks is argued to have reduced respondent fatigue (Stats SA 2017c). Most importantly Stats SA (2017c) argues that the collection of expenditure data collected data through the diaries was for the same time period. The decision also was meant to deal with respondent fatigue (Stats SA2017c:2)

To date there has been two LCSs conducted in South Africa, in 2008 and in 2014 whilst the IES has been conducted in 1995, 2000, 2005 and 2010. This was shown earlier in

[11] Budlender, Leibbrandt and Woolard (2015) used the IES of 2010 to review poverty lines in South Africa. Finn, Leibbrandt and Ouisthuizen (2014) also used IES data to analyse poverty and inequality in the post-apartheid period in South Africa.

Table 6.1 The three surveys have a similar data collection period of 12 months and they consistently used the acquisition, payment and consumptions approaches in data collection.

Table 5.2: Data collection for the LCSs and IES of 2010

Description	LCS 2008	IES 2010	LCS 2014
Household Questionnaire Modules	i. House composition and structure ii. Health disability, education and employment iii. Welfare, assets and information on dwellings and services iv. Consumption expenditure v. Consumption expenditure vi. Savings investment, debt, remittances and income vii. Anthropometric measurements (7 modules)	i. Household composition Detailed income and expenditure data from a comprehensive list	i. House composition and structure, education and employment ii. Expenditure iii. Expenditure iv. Subsistence, living circumstances food security, finance, banking, income (4 modules)
Data collection method	Diary and recall	Diary and recall	Diary and recall
Number of households	25 075	25 328	23 380
% Male (weighted)	48.18*	48.55*	48.89*
%Female (weighted)	51.82*	51.55*	51.11*
Number of diaries	4	2	2
Diary keeping period	One week	2 weeks	2 weeks
Survey periods	12	26	26
Data collection period	September 2008- August 2009	September 2010- August 2011	October 2014- October 2015
Visits per household	6	4	4
Zero income households	552	164	68[12]

[12] The number of zero income households significantly reduced between 2008 and 2014, but this has not been properly documented by Statistics South Africa. They applied hot deck imputation on missing demographic variables, but it remains unclear how they imputed for zero income households in the LCS of 2008 and 2014 as well as the IES of 2010.

Description	LCS 2008	IES 2010	LCS 2014
Classification of expenditure	COICOP	COICOP	COICOP
Population Census	2001	2001	2001

Source: Own calculations; Statistics South Africa (2011,2012, 2017)

* Own calculations, Data are weighted

Data collection in the LCS in both 2008 and 2014 and IES of 2010 was done using household diaries, a household questionnaire, and a summary questionnaire. In 2008, four weekly diaries were given to households and these were reduced to two in 2010 and 2014. Stats SA (2017c:1) justified this change in the diary-keeping period as a move to minimize respondent fatigue and 'improve diary keeping'. Furthermore, whereas a household was interviewed six times as was the case with the LCS of 2008, the number of household visits reduced to four in 2010 and 2014 (Stats SA 2012b; 2017c). Despite these changes, the three surveys used a similar annualization technique.

As previously indicated, the IES of 2010 similar to the LCS of 2008 and 2014 drew its sample of dwelling units from the Statistics South Africa Master Sample based on information from the 2001 Population Census. A survey period of 12 months is also applied in the three surveys and this gives both surveys a similar reference time frame.

The variable of interest in this chapter is the expenditure variable and the LCS of 2008 and 2014 as well as the IES of 2010 collected comparable expenditure data. A summary of the expenditure categories and Rand values in the three cross sections is shown in Table 5.3 below. The table also shows the percentage contributions of various expenditure categories.

Table 5.3: Average Expenditure in the LCS and IES surveys (Nominal values)

Main expenditure category	2008* (Rand)	2010** (Rand)	2014*** (Rand)	2008 (%)	2010 (%)	2014 (%)
Food and non-alcoholic beverages	13 914	12 200	13 292	19.3	12.8	12.9
Alcoholic beverages and tobacco	699	1 045	911	1.0	1.1	0.9
Clothing and footwear	3 474	4 284	4 939	4,8	4.5	4.8
Housing, electricity, water and other utilities	17 922	30 505	33 625	24.9	32.0	32.6
Furniture, furnishings and other household equipment	3 860	4 877	5 321	5.4	5.1	5.2
Health	950	1 357	935	1.3	1.4	0.9
Transport	10 978	16 319	16 826	15.3	17.1	16.3
Communication	2 428	2 702	3 509	3.4	2.8	3.4
Recreation and culture	3 069	2 900	3 933	4.3	3.0	3.8
Education	2 002	2 544	2 531	2.8	2.7	2.5
Restaurants and hotels	1 697	2 313	2 181	2.4	2.4	2.1
Miscellaneous goods and services	10 715	14 003	15 167	14.9	14.7	14.7
Unclassified items	201	134	55	0.3	0.1	0.1

Source: Statistics South Africa (2012b; 2012c; 2017b)

*LCS 2008; **IES 2010; ***LCS 2014

As shown in Table 5.3 above, the three surveys used in this study have similar main categories of expenditure. Of note as well is the fact that the major drivers of expenditure were mostly from expenditure on housing, electricity, water and other utilities; Transport; Food and non-alcoholic beverages; and miscellaneous goods and services categories. This is consistent in 2008, 2010, and 2014.

It is also worth mentioning that Stats SA applied a similar definition in 2008, 2010 and 2014 of expenditure. To be more specific, Stats SA in the IES and LCSs defines consumption expenditure as; '*Expenditure of goods and services acquired and privately used by household members including imputed values for items produced and consumed by the household itself' (*Stats SA 2012b:61; 2012c:31; 2017b:66).

The IES and LCS surveys also have the advantage that they were all conducted by the same agency; Stats SA and this means similar resources were used in their conduct and

this improves chances of comparability of the surveys. Furthermore, the IES of 2010 and the LCSs of 2008 and 2014 used a similar definition of households and household heads. A household for instance is defined as '*a group of persons who live together and provide themselves jointly with food and or other essentials for living, or a single person who lives alone*' (Stats SA 2012c1:31; 2012b:31; 2017c:66). On the other hand a household head is defined as '*a person recognized as such by the household usually the main decision maker or the person who owns or rents the dwelling, or the person who is the main breadwinner'* (Stats SA 2012b:31; 2012c:31; 2017c:66). The three surveys as a result were mostly comparable as far as household expenditure and household headship were concerned, the key variables of analysis in this study.

In general, household surveys that collect information on expenditure and income are characterized by a relatively higher number of refusals due to their demanding nature when compared to other types of household surveys (Stats SA 2017c). The LCSs seemingly were more affected with non-response than other income and expenditure surveys in South Africa. Table 5.4 below shows the response rates of the LCSs by province.

It is clear from the table below that the response rate of the LCS is lower in 2014 than in 2008. Green (2017) illustrated the 'high wall' bias associated with upmarket suburb areas as the main justification why response rates are lower in provinces with significantly large urban areas. The argument in Green (2017) suggests that the existence of 'high walls' in rich neighbourhoods makes data collection difficult. Part of this is partly attributed to the fact that high-end households tend to be 'busy' and not always available for surveys.

Table 5.4: Response rates for the IES and LCSs

Province	Response Rate % LCS (2008)	Response Rate % IES (2010)	Response Rate % LCS (2014)
Western Cape	85.2	94.3	79.1
Eastern Cape	94.2	95.8	92.7
Northern Cape	90.4	95.6	90.4
Free State	95.9	94.7	86.9
Kwa-Zulu Natal	84.8	92.3	89.6
North West	89.3	91.6	90.0
Gauteng	79.7	82.9	65.3
Mpumalanga	88.5	93.5	91.6
Limpopo	94.9	90.3	95.6
South Africa	88.0	91.6	84.9

Source: Statistics South Africa (2012b; 2012c; 2017b)

Table 5.4 above indicates that response rate was the lowest in 2014 and Western Cape and Gauteng provinces consistently have the lowest response rates on average. To attempt to link the 'high wall' bias discussed earlier, Table 5.5 below illustrates the differences in the make-up of South Africa's provinces based on the urban and rural break down. In a bid to understand the response rate patterns of households in South Africa, it is safe to conclude the existence of lower response rates in high-end suburbs when compared to other residential areas that are arguably 'more accessible'. It would also seem reasonable if not ambitious to attempt to argue that provinces that are dominantly more urban such as Gauteng and Cape Town are also more likely to have lower response rates when compared to provinces that have a significant rural population such as Limpopo and Eastern Cape provinces. Statistics South Africa (2003) using the Population Census of 2001, states that 90.4% and 9.6% of the Western Cape is urban and rural respectively, Gauteng is also shown to be dominantly urban with close to 90% of the province classified as urban compared to 3% that is rural.

Table 5.5: Urban and rural make-up of South African provinces

Province	Statistics South Africa (2003) EA Type	
	Urban (%)	Rural (%)
Western Cape	90.4	9.6
Eastern Cape	38.8	61.2
Northern Cape	82.7	17.3
Free State	75.8	24.2
KwaZulu-Natal	46.0	54.0
North West	41.8	58.2
Gauteng	97.2	2.8
Mpumalanga	41.3	58.7
Limpopo	13.3	86.7

Source: Stats SA (2003:71)

Using the LCS of 2014, Gauteng and the Western Cape Province had the lowest response rates of 65.3% and 79.1% respectively. On the other hand, Provinces such as Limpopo with an 86% rural area and 13% urban and Eastern Cape with a rural population making up 61.2% and 38.8% urban have higher response rates in the LCS (95.6% in Limpopo and 92.7% in the Eastern Cape.

5.9 Real and nominal expenditure values

The expenditure variable used in this study is expressed in 2014 prices. The use of nominal prices distorts the actual changes in expenditure that occurs over the years. For example, an observed increase in nominal expenditure may not necessarily show that households' expenditure increased in real terms during the observed period. Instead it could be a case of price inflation. Therefore, real expenditure by adjusting for this defect is able to capture the actual increase of expenditure in the different years under analysis.

5.10 Descriptive analysis of pseudo panel

This section gives a preliminary analysis of the birth cohorts especially on their sizes and composition by gender and geography. Table 5.6 below summarizes the size of the 13 birth cohorts used in the pseudo panel between 2008 and 2014.

Table 5.6: Number of Birth cohorts in pseudo panel

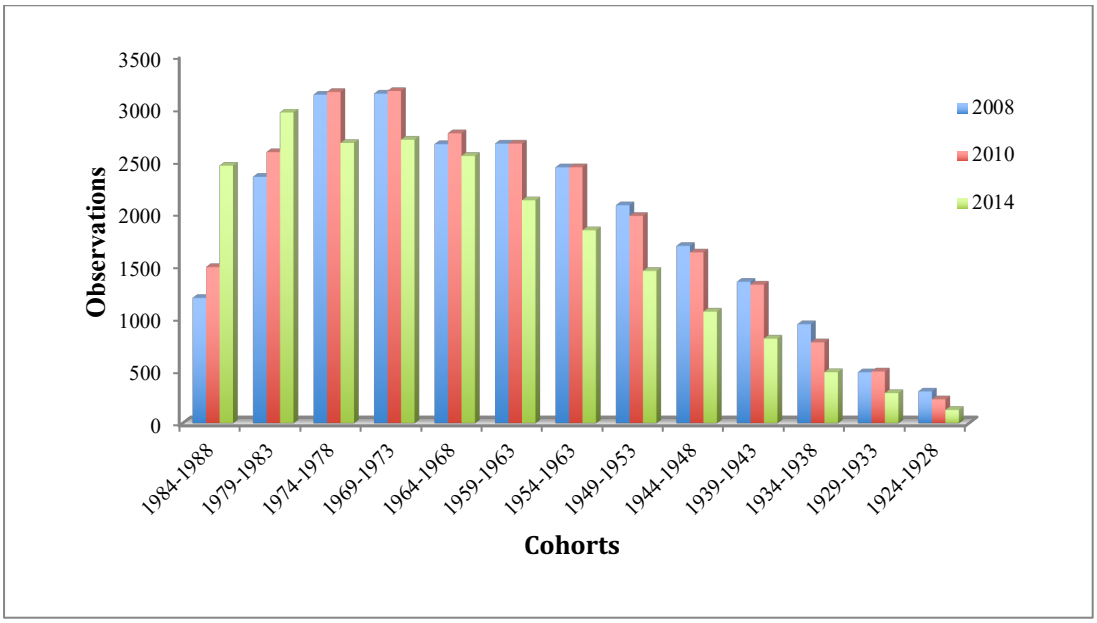
Year of Birth Cohorts	(2008) Observations	(2010) Observations	(2014) Observations
1984-1988	1191	1485	2449
1979-1983	2343	2578	2955
1974-1978	3124	3150	2665
1969-1973	3134	3159	2696
1964-1968	2654	2757	2542
1959-1963	2658	2659	2120
1954-1958	2433	2434	1835
1949-1953	2073	1971	1448
1944-1948	1686	1624	1061
1939-1943	1345	1318	804
1934-1938	941	769	485
1929-1933	484	492	286
1924-1928	302	227	126

Source: Author's own calculations

What is evident from the table above is that older cohorts tend to be smaller in size when compared to younger birth cohorts. To illustrate the differences in sizes of the birth cohorts more visibly, Figure 5.1 below graphically shows the distribution of the birth cohorts in the three time periods. The figure almost proved the existence of a bell-shaped population distribution, but rather the shape is more of a left skew suggesting that the younger population is higher than the older birth cohorts. This is to be expected in a youthful demography and remains therefore not surprising.

Figure 5.1 below additionally shows three major distinctions in terms of the sizes of the birth cohorts. The youngest birth cohorts are relatively smaller than the 'middle aged' birth cohorts and the sizes of the birth cohorts reduce when they are older. This could partly be a result of increasing household sizes as younger people start their own families and have children and break away to start new households. The sizes start reducing as they get older and this is partly because older household heads have passed the childbearing age and therefore the old aged birth cohorts become smaller.

Figure 5.1: Distribution of birth cohorts (2008-2014).

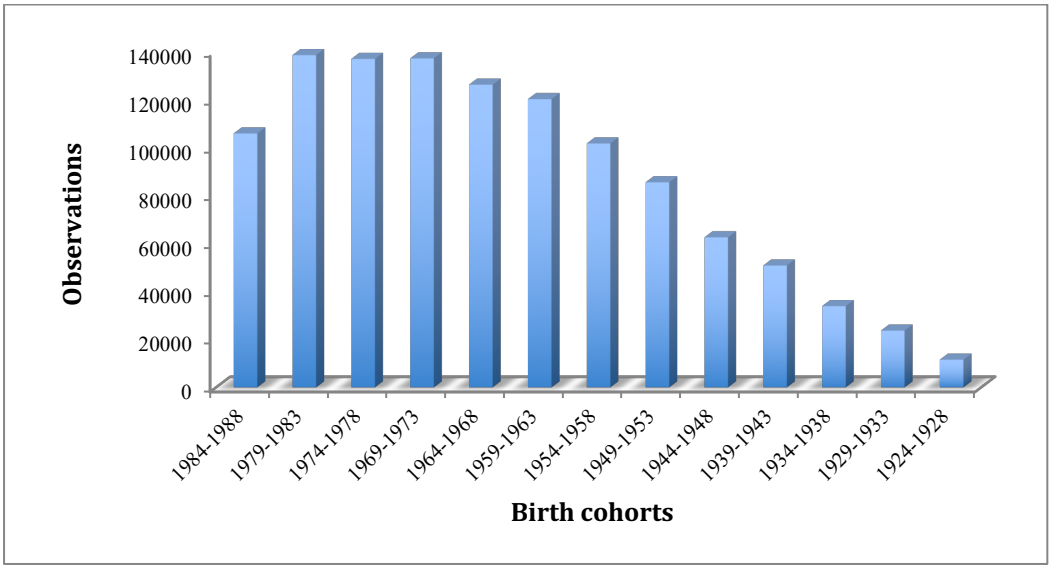


Source: Author's own calculations

Figure 5.1 above also shows that the youngest birth cohorts (1984-1988 and 1979-1978) grew in size between 2008 and 2014. However, for the older cohorts they either grew in size between 2008 and 2014, maintained similar sizes, or in some cases reduced in size in the same period.

Using Stats SA's Population Census 10% sample, this study generated birth cohorts with similar restrictions. Figure 5.2 below shows a similar distribution of birth cohorts based on household headship. The sizes of the birth cohorts increase and then start declining, validating the distribution of birth cohorts illustrated in Figure 5.1.

Figure 5.2: Distribution of birth cohorts from the 2011 Population Census 10% sample.

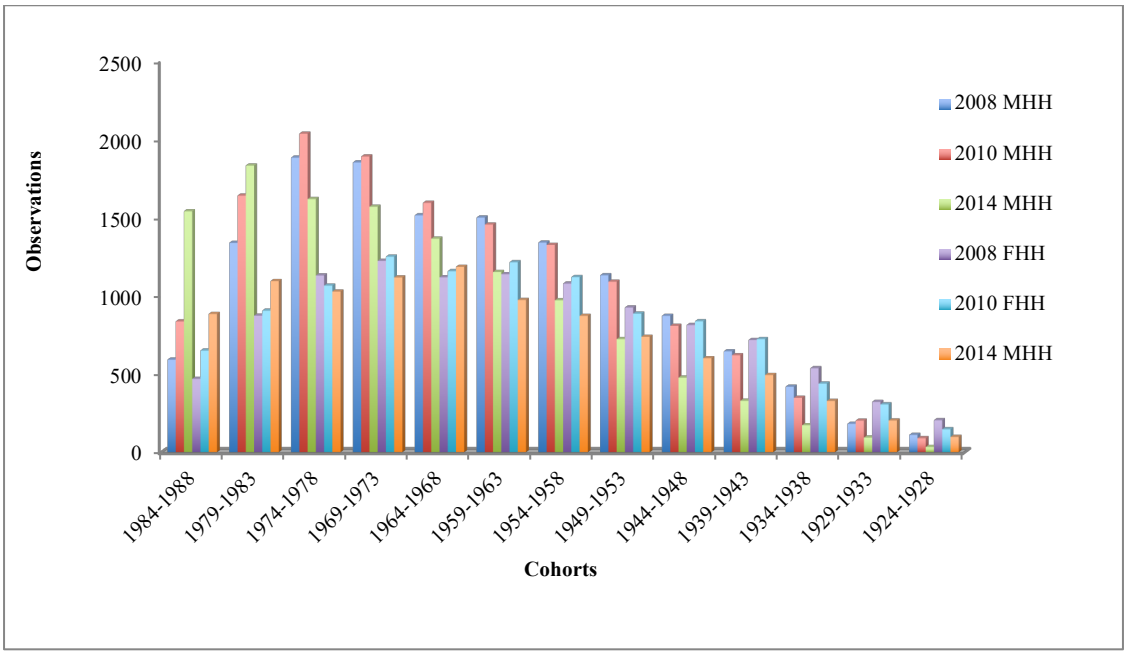


Source: Author's own calculations using the 10% sample of the 2011 Population Census

Figure 5.2 below disaggregates the sizes of the birth cohorts generated using the LCSs and IES of 2010, by gender of the household head. The illustration shows the changes in sizes of the birth cohorts for Male Headed Households (MHH) and Female Headed Households (FHH) in 2008, 2010 and 2014. For each birth cohort, the illustration shows how the sizes of the MHH and FHH differ in size over time.

In 2008, 2010 and 2014, MHH are more than FHH and this is more pronounced for younger birth cohorts. This reverses however, for older birth cohorts. For instance, starting from the 10th birth cohort (1939-1943) to the oldest birth cohort (1924-1928), FHH have larger sizes then MHH. This is expected given the fact that women have a higher life expectancy then men in South Africa. Stats SA (2019c:1) using the latest Population Census data found that women have a higher life expectancy (67.3 years) than men (61.1 years).

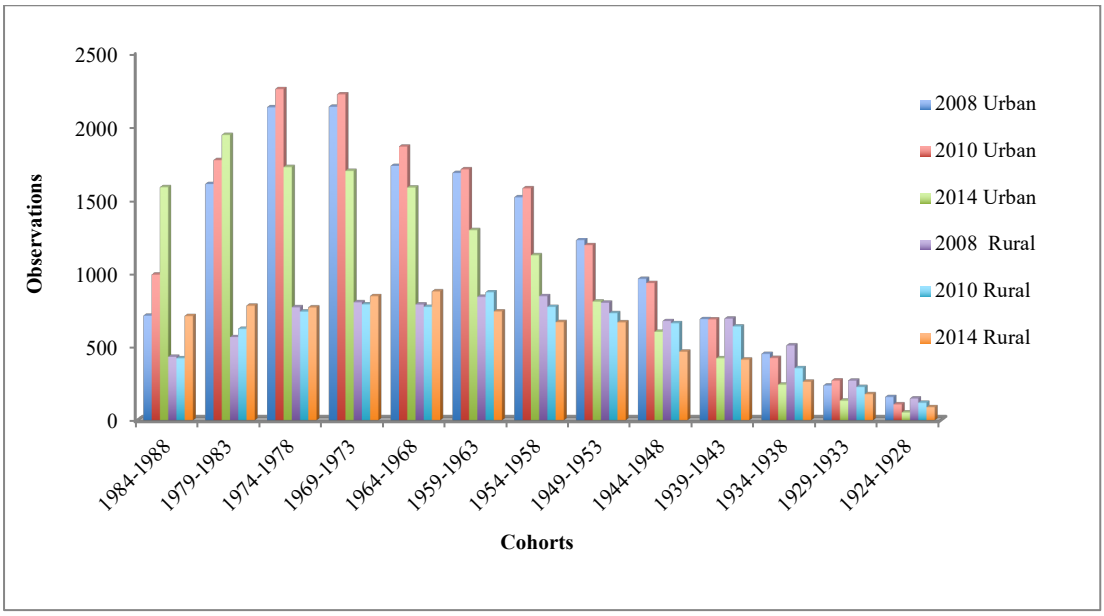
Figure 5.3: Distribution of birth cohorts by gender.



Source: Author's own calculations

Figure 5.3 below shows the distribution of the birth cohorts in rural and urban areas. A geographic analysis of the size of the birth cohorts in the Figure 5.4 below shows that the sizes of birth cohorts are larger in urban areas when compared to birth cohorts in rural areas. This is consistently observed in 2008, 2010 and 2014 for the youngest 10 birth cohorts. The difference in the sizes of the first 10 birth cohorts shows a concentration of a significantly large working population. This is partly fuelled by the likelihood of rural to urban migration in younger cohorts as they seek employment in urban areas. But the last four birth cohorts show that older household heads are more concentrated in rural areas, reversing the observed trend. This is also not unforeseen, as older people are less likely to migrate to urban area. Instead migration pattern might be an urban-to-rural as older cohorts are retired and seeking rest.

Figure 5.4: Distribution of birth cohorts by geography.

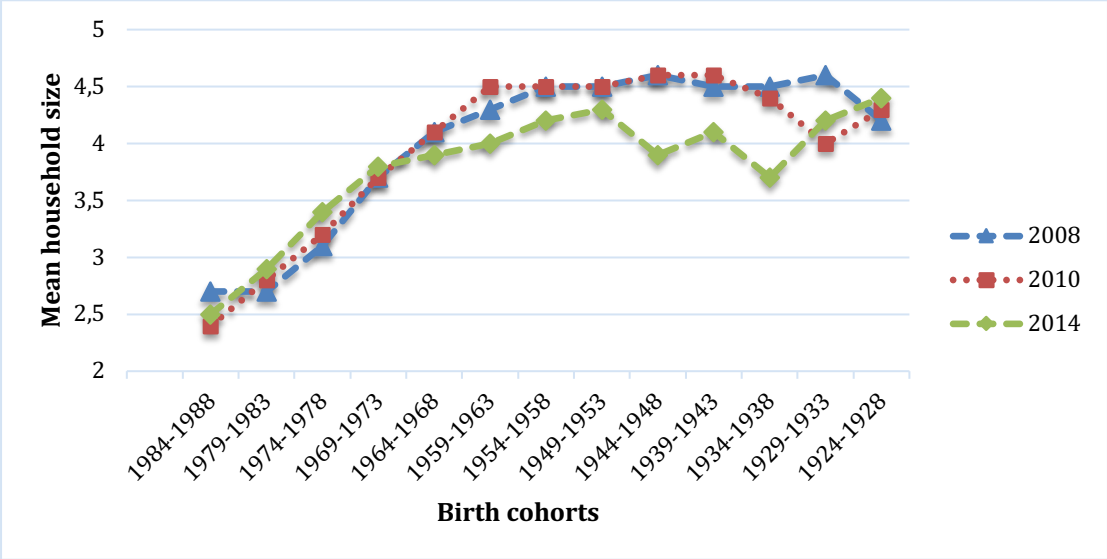


Source: Author's own calculations

The analysis of the sizes and distribution of the birth cohorts in the pseudo panel generated in this study would be incomplete without factoring in the household sizes of the birth cohorts. The mean household sizes of the birth cohorts in 2008, 2010 and 2014 ranges between 2.4 and 4.4 and a median of between two and four. As expected, younger birth cohorts have a relatively smaller mean household size when compared to older birth cohorts. The household sizes also tend to reduce slightly in size for the oldest birth cohorts. This is more evident in 2008 but slightly indistinct in 2010 and 2014.

The existence of smaller mean household sizes for younger cohorts is anticipated. They are more likely to be single, or just starting their families and as such they remain with a smaller household size. The mean and median household sizes then increase as the cohorts get older and remain large but constant for old aged birth cohorts. Figure 5.5 below attempts to illustrate this argument visually.

Figure 5.5: Mean household size of birth cohorts.



Source: Author's own calculations

Figure 5.5 above illustrates that household sizes increase, as cohorts get older before either declining or stabilizing.

5.11 Conclusion

This study created a birth cohort-based pseudo panel as an alternative panel data source for poverty analysis in South Africa. Existing analysis of poverty dynamics in South Africa to date have been limited to using the National Income Dynamic Study (NIDS) and the KwaZulu Income Dynamic Study (KIDS) as some of the key panel data that enable the analysis of poverty over time in South Africa. On the other hand, the use of synthetic panels has also been growing. Existing research also alludes to the existence of different approaches in creating synthetic panel data and each method is dependent on the type of data available as well as the assumptions and research questions that one seeks to answer. In this chapter, the approach used was a cohort-based pseudo panel methodology.

The approach in the generation of the pseudo panel entailed the usage of three successive rounds of cross-sectional data in 2008, 2010 and 2014. Within these three cross sections, birth cohorts of the household head separated by five-year bands born between 1924 and 1988 were generated. In total, 13 cohorts were generated and tracked in three time periods. Whereas a genuine panel might use household identification to follow similar households over time, this pseudo panel assumed that household heads

with a similar birth year are the same over time, an assumption that allows this study to 'track' their expenditure and poverty status in three periods.

The chapter also included a detailed discussion on how the cross-sectional surveys used in this chapter were selected. Considerations mostly highlighted the importance of comparability of data as well as the existence of large sample surveys. The fact that South Africa has a commendable number of household surveys that have been used in poverty analysis was a great starting point. The chapter justified the choice of the LCSs of 2008 and 2014 as well as the IES of 2010. The comparability of these three household surveys established that they were adequate and relevant to generate a birth cohort-based pseudo panel for poverty analysis. The last section conducted a descriptive analysis of the pseudo panel. In particular, the sizes of the birth cohorts as well as the mean and median of their household sizes were explored. The analysis was extended for FHH and MHH birth cohorts as well as for birth cohorts in rural and urban areas.

Chapter Six

Estimating Expenditure Patterns and Poverty Dynamics Using a Cohort Based Pseudo Panel in South Africa

6.1 Introduction

This chapter analyses the patterns of expenditure and poverty rates of a cohort-based pseudo panel. The chapter is divided into three sections. The first part presents descriptive statistics on the differences in mean expenditure per capita for 13 birth cohorts. The descriptive statistics help to show the extent of heterogeneity across different characteristics, such as gender, race, and geographic variables specific to the household heads. Section two discusses results from a multivariate regression estimation. In particular, the section outlines the trends and patterns of conditional (residuals from the regression) means and cross-sectional variance. The conditional variables help to control for observable characteristics across cohorts such as cohort, time and age effects, or other household characteristics. Lastly, the third section presents estimates for the changes in poverty rates for the birth cohorts in 2008, 2010, and 2014.

6.2 Estimating the changes in mean expenditure per capita for birth cohorts in the pseudo panel

This section gives the changes in mean expenditure per capita for birth cohorts[13] in 2008, 2010 and 2014. The expenditure values used in this chapter use the 95% Confidence Interval, (CI). The mean expenditure per capita values are displayed graphically and the accompanying tables with CIs are also included. The use of the CI allows an inclusion of a margin of error in the given estimates. The CI gives us the range of possible values for the true population mean, and therefore represents the values that we can be 95% certain that they contain the true population mean. Increasing the sample size would increase the width of the 95% CI (Sedgwick, 2014), and in this case we can say with more precision that the range of values contain the true population

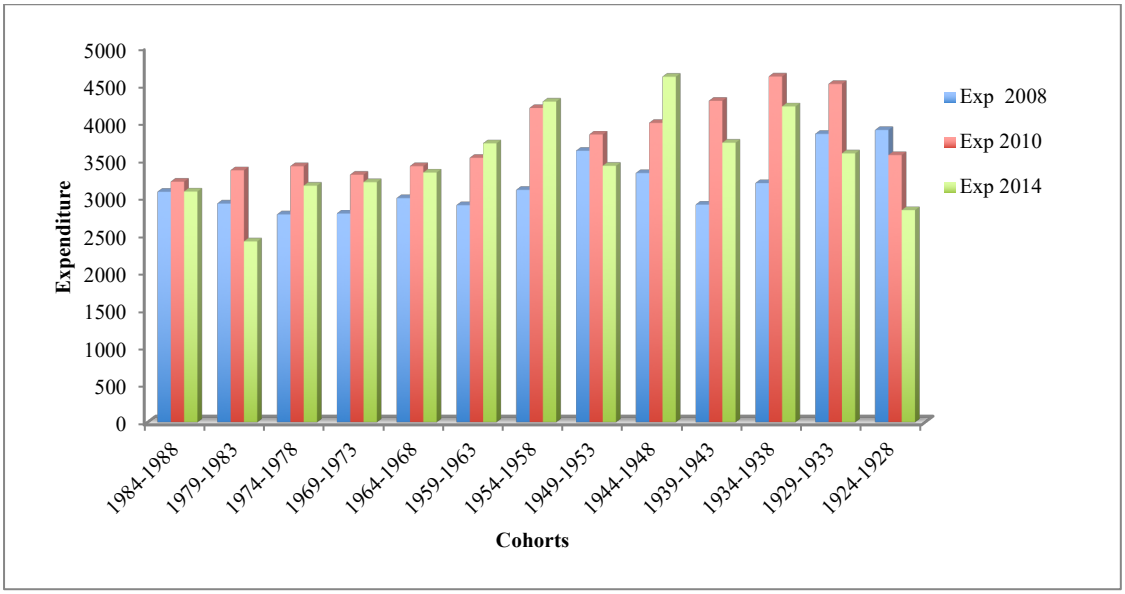
[13] The analysis on expenditure patterns and later on poverty done in this chapter uses the household as a unit of analysis. The analysis uses individual characteristics of the household head such as labour market participation and social grants to try and understand the sources of income of the he household head. It is possible that within the same household, there is more than one individual receiving income from the labour market or from social grants. A household head may also not be the one receiving such income.

than if we used a small sample. The inclusion of the tables with CI ranges therefore enables the analyses to include the level of uncertainty and express how the given mean values are likely to contain the true population mean.

The confidence intervals calculated for all the mean values discussed in this chapter captures the range within which the mean falls. This is except the oldest Indian birth cohort which as a result of the small samples has mean values that are not within the CI range. An analysis of the Confidence Intervals (CI) for the mean expenditure per capita for all birth cohorts during the period under analysis confirms the increases and decreases for a majority of the observed mean expenditure per capita values at the 95% CI for all birth cohorts. The true population mean is a specific number and with a majority of the estimate made in this section, the true population means falls within the given CI ranges.

Figure 6.1 below shows the changes in per capita mean expenditure per capita for all birth cohorts in 2008, 2010 and 2014.

Figure 6.1: Mean expenditure per capita for all birth cohorts.



Source: Author's own calculations

For a majority of birth cohorts in the pseudo panel displayed in Figure 6.1 above, mean expenditure per capita was lower in 2008 when compared to 2010 and 2014. However, mean expenditure per capita in 2014 was lower for most birth cohorts than in 2010. Figure 6.1 also shows an improvement in expenditure from 2008 to 2014 for a majority of the birth cohorts, but this is less than would have been expected had the income

growth 2008-2010 been sustained. Younger birth cohorts also have lower mean expenditure per capita when compared to older birth cohorts. And middle cohorts (1933-59) have experienced the greatest rates of change. The 1979-1983 seem to have been the most disadvantaged with a large fall in their incomes in 2014. The only other group in this position is the very old (1924-1928).

Table 6.1 below illustrates the graph represented in Figure 6.1 above and most importantly includes the CI ranges. The patterns displayed in Figure 6.1 above remain unchanged after adding the CI. We are 95% confident that the sample mean expenditure values for all birth cohorts displayed in Figure 6.1 above fall within the CI range at the 95 CI.

Ttests were conducted to determine if there existed statistical significance between the expenditure means recorded for all birth cohorts for the years under analysis. Specifically, an unpaired ttest was run on data that pooled all the 13 birth cohorts together. Three mean expenditure values were estimated for 2008, 2010 and 2014 for all birth cohorts and an unpaired ttest was conducted. The objective was to test if the changes in expenditure for 2008, 2010 and 2014 were statistically significant. The difference between the mean expenditure value for all 13 birth cohorts in 2008 and 2010 were statistically significant with a p-value below 0.05. Similar ttests were conducted with variations between gender, geography and racial categories and comparisons for the years 2008 and 2010 remained statistically significant. But the 2010 and 2014 expenditure values for all birth cohorts are not statistically significant, with a p-value above 0.05. Table AP14 in the Appendix has the results of the ttests conducted.

Table 6.1: Mean Expenditure per Capita (Overal)

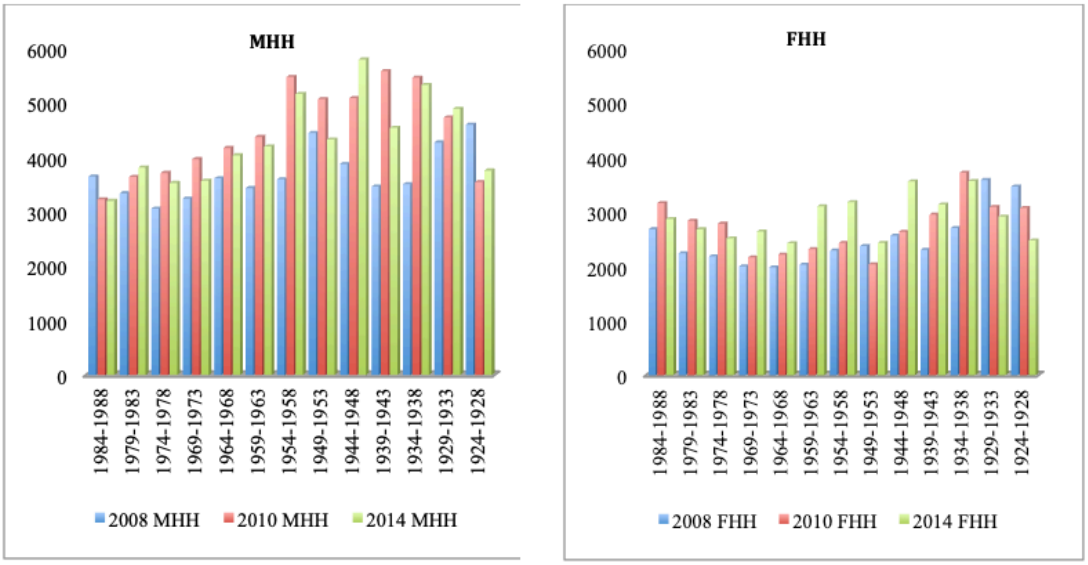
Birth cohorts	2008	CI 95%	2010	CI 95%	2014	CI 95%
1984-1988	3071	2424-3718	3206	2835-3577	3075	2693-3457
1979-1983	2914	2622-3207	3358	3050-3665	2410	3031-3789
1974-1978	2770	2556-2984	3412	3042-3782	3153	2882-3425
1969-1973	2782	2559-3004	3299	3010-3588	3200	2929-3471
1964-1968	2987	2756-3218	3413	3079-3748	3328	2986-3670
1959-1963	2893	2667-3119	3524	3215-3832	3717	3308-4125
1954-1963	3097	2795-3399	4188	3460-4916	4274	3733-4815
1949-1953	3618	3168-4068	3835	3281-4390	3418	2980-3856
1944-1948	3321	2912-3731	3989	3348-4630	4605	3805-5405
1939-1943	2900	2508-3293	4284	3453-5116	3725	3018-4432
1934-1938	3187	2652-3722	4608	3493-5722	4209	3325-5092
1929-1933	3843	2429-5258	4509	3009-6009	3583	2570-4597
1924-1928	3895	2833-4958	3561	2386-4735	2827	1728-3926

Source: Author's own calculations

6.2.1 A gendered analysis of per capita expenditure

Extending the analysis of mean expenditure per capita of birth cohorts by gender of household heads, Figure 6.2 below illustrates that Male Headed Households (MHH) have higher mean expenditure per capita than Female Headed Households (FHH) in all the years under analysis. The differences in mean expenditure per capita for these two categories are higher for older birth cohorts. To be more specific, the differences in expenditure for MHH and FHH birth cohorts are higher for birth cohorts born between 1934 and 1958, but these starts declining for the 1929-1933 and 1924-1928 birth cohorts.

Figure 6.2: Mean expenditure per capita (in Rands) by gender of household head.



Source: Author's own calculations

For MHH birth cohorts, mean expenditure per capita was the lowest in 2008 and it increased in 2010 and 2014. The increase in per capita expenditure in 2014 however is experienced by fewer MHH birth cohorts as most cohorts did not gain in expenditure between 2010 and 2014. Figure 6.2 above also shows the birth cohorts with birth years 1934-1958 having the highest mean expenditure per capita and this declines with the two oldest birth cohorts. FHH cohorts (1948-1969) had their incomes grow consistently from 2008-2014 while male heads did not, albeit from a lower base.

An analysis of the differences in mean expenditure per capita of the FHH, shows a slightly different trend of mean expenditure per capita of this group of birth cohorts when compared to MHH. Mean expenditure per capita for FHH starts by decreasing for younger birth cohorts. The decline continues for a majority of the birth cohorts with few birth cohorts experiencing an increase in mean expenditure per capita in 2014. Mean expenditure per capita for older birth cohorts increases especially for the 1944-1948, 1934-1938 and 1929-1933 FHH birth cohorts

Table 6.2: Total monthly household expenditure per capita for cohorts (MHH)

Birth cohorts	2008	CI 95%	2010	CI 95%	2014	CI 95%
1984-1988	3630	2499-4762	3215	2766-3665	3185	2768-3602
1979-1983	3325	2893-3758	3625	3224-4027	3797	3277-4316
1974-1978	3045	2773-3316	3697	3203-4190	3515	3188-3842
1969-1973	3228	2919-3538	3951	3530-4373	3557	3229-3886
1964-1968	3598	3276-3921	4153	3648-4659	4021	3591-4452
1959-1963	3420	3113-3728	4356	3948-4765	4180	3590-4769
1954-1963	3579	3189-3970	5452	4249-6654	5142	4396-5889
1949-1953	4429	3746-5112	5046	4160-5932	4307	3676-4939
1944-1948	3861	3287-4436	5067	3975-6159	5771	4665-6877
1939-1943	3448	2843-4054	5560	4098-7022	4521	3454-5588
1934-1938	3491	2716-4266	5439	3880-6999	5308	3811-6806
1929-1933	4255	2813-5698	4714	3722-5706	4871	3083-6660
1924-1928	4582	2657-6507	3529	1715-5343	3742	1755-5728

Source: Author's own calculations

Table 6.2 above shows the figures represented in Figure 6.2. The CI indicates that 95% of the time, the mean expenditure values illustrated in Figure 6.2 fall within the CI and conforms to the observed patterns for both MHH and FHH in Table 6.3 below.

Table 6.3: Total monthly household expenditure per capita for cohorts (FHH)

Birth cohorts	2008	CI 95%	2010	CI 95%	2014	CI 95%
1984-1988	2667	2102-3232	3147	2485-3809	2855	2065-3645
1979-1983	2229	1912-2546	2824	2380-3287	2667	2180-3153
1974-1978	2169	1824-2515	2768	2275-3260	2499	2013-2985
1969-1973	1989	1713-2266	2154	1885-2422	2624	2153-3094
1964-1968	1971	1688-2254	2205	1931-2178	2411	1859-2963
1959-1963	2021	1725-2317	2304	1854-2755	3088	2548-3628
1954-1963	2277	1844-2709	2417	2014-2821	3164	2392-3935
1949-1953	2360	1908-2812	2028	1681-2375	2417	1827-3007
1944-1948	2549	1992-3106	2621	2095-3146	3541	2395-4687
1939-1943	2289	1801-2777	2936	2225-3647	3116	2170-4062
1934-1938	2690	2014-3366	3700	2049-5350	3546	2446-4646
1929-1933	3570	1395-5745	3077	2193-3961	2897	1665-4129
1924-1928	3448	2212-4684	3053	1656-4450	2463	1116-3811

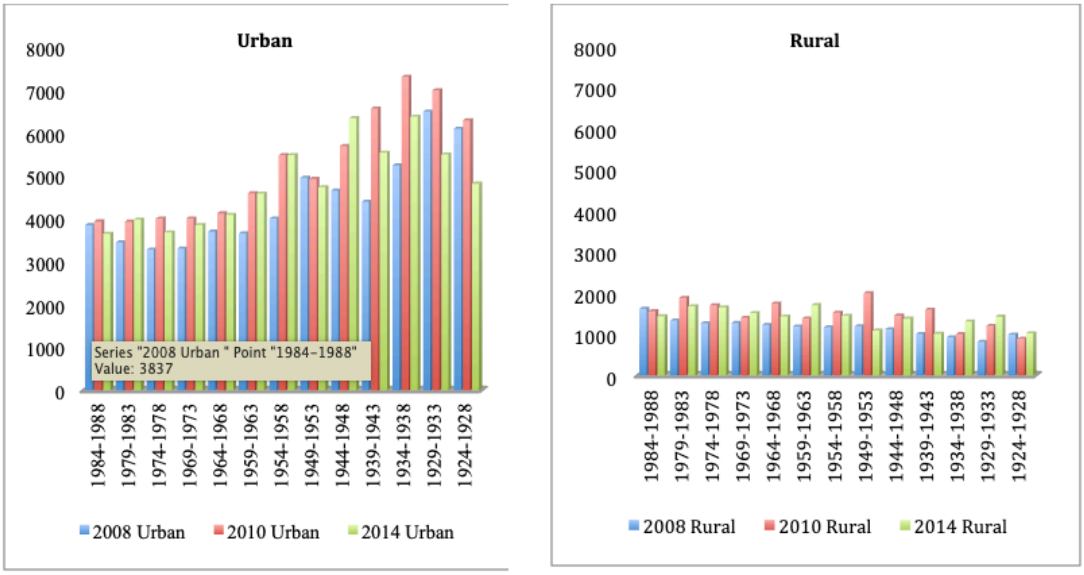
Source: Author's own calculations

6.2.2 A geographical analysis of mean per capita expenditure

The analysis of expenditure for birth cohorts was further disaggregated geographically. In Figure 6.3 below, what is evident are the stark differences in expenditure that exists between birth cohorts in urban and rural areas. The results show that birth cohorts in urban areas have higher expenditures when compared to birth cohorts in rural areas. This finding is consistent in 2008, 2010 and 2014.

The differences in mean expenditure per capita for birth cohorts in rural and urban areas are more prominent for older birth cohorts. Figure 6.3 below shows that mean expenditure per capita differences increases especially for birth cohorts born between 1929 and 1944. There is a decline in mean expenditure per capita for the last birth cohort.

Figure 6.3: Mean expenditure per capita by geography.



Source: Author's own calculations

What is also clear in Figure 6.3 is that the differences in mean per capita expenditure are more pronounced for older birth cohorts. There is also a decline in expenditure inequalities between birth cohorts in urban and rural areas that comes in with the oldest birth cohort.

The mean expenditure per capita for birth cohorts in urban areas increases as birth cohorts get older before declining for the 1929-1933 and 1924-1929 birth cohorts. The increase in mean expenditure per capita for the younger birth cohorts is less noticeable. However, for older birth cohorts born before 1950, mean income increased in 2010 and 2014. Birth cohorts born before 1939, also experienced a decline in their mean expenditure per capita in 2014 and this was a reduction from the mean expenditure per capita gained between 2008 and 2010. What is evident from Figure 6.3 above is that differences in expenditure in urban areas are more pronounced for older birth cohorts. Younger birth cohorts in urban areas seem to have similar level of mean expenditure per capita. This is shown for all birth cohorts born after 1968.

In rural areas, younger birth cohorts tend to be more unequal in expenditure patterns and the differences reduce as they reach middle ages to old age. Figure 6.3indiates that despite the fluctuations in mean expenditure per capita for the birth cohorts in rural areas, the general trend shows declining mean expenditure per capita. As the birth cohorts get older the decline is more evident for the last three birth cohorts. What is

also hard to miss are the increases in mean expenditure per capita experienced in 2010 for a majority of the birth cohorts in rural areas. Figure 6.3 also reflects the influence of some outliers. The 1949-1953 rural birth cohort and the urban 1934-1943 birth cohort, for instance, have a sharp increase in mean per capita income in 2010 and this declines in 2010.

Tables 6.4 and 6.5 below show the total monthly household expenditure per capita as illustrated in Figure 6.3 above. An analysis of the CIs between 2008 and 2014 shows the mean per capita expenditures for birth cohorts in urban and rural areas are within the upper and lower limits of the CI and therefore 95% of the time we are confident the sample mean contains the true population mean.

Table 6.4: Total monthly household expenditure per capita for cohorts (Urban)

Birth cohorts	2008	CI 95%	2010	CI 95%	2014	CI 95%
1984-1988	3837	2914-4761	3924	3414-4434	3637	3138-4136
1979-1983	3438	3058-3817	3917	3522-4312	3969	3480-4458
1974-1978	3270	3006-3552	3991	3513-4469	3670	3322-4018
1969-1973	3296	3012-3580	3992	3610-4374	3843	3484-4201
1964-1968	3689	3689-3994	4112	3672-4553	4078	3616-4540
1959-1963	3646	3343-3950	4578	4149-5008	4571	4017-5125
1954-1963	3990	3570-4411	5462	4418-6506	5466	4729-6202
1949-1953	4938	4279-5597	4912	4236-5587	4718	4074-5363
1944-1948	4641	4030-5252	5678	4686-6671	6323	5160-7487
1939-1943	4382	3753-5011	6546	5265-7827	5518	4412-6625
1934-1938	5226	4305-6147	7282	5486-9078	6358	5001-7715
1929-1933	6477	3901-9053	6969	4489-9449	5481	3878-7084
1924-1928	6078	4431-7725	6274	4192-8355	4801	2604-6997

Source: Author's own calculations

Table 6.5: Total monthly household expenditure per capita for cohorts (Rural)

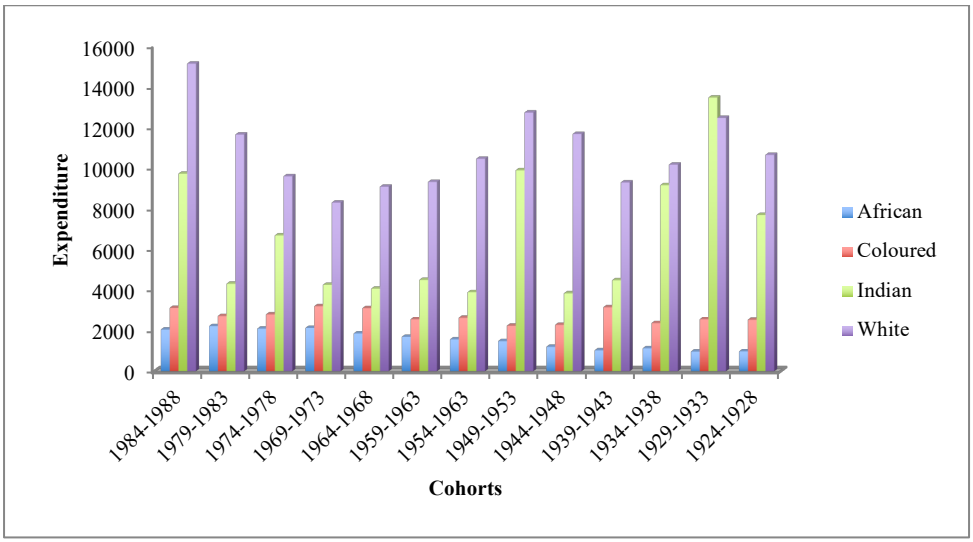
Birth cohorts	2008	CI 95%	2010	CI 95%	2014	CI 95%
1984-1988	1622	1014-2230	1561	1322-1800	1438	1240-1636
1979-1983	1334	1187-1481	1888	1525-2250	1682	1438-1906
1974-1978	1273	1106-1440	1708	1378-2048	1658	1388-1927
1969-1973	1275	1083-1467	1402	1252-1551	1519	1316-1722
1964-1968	1233	1105-1360	1748	1371-2124	1436	1246-1626
1959-1963	1191	1037-1345	1383	155-1611	1709	1396-2022
1954-1963	1173	1034-1311	1527	1160-1895	1449	1203-1694
1949-1953	1196	1034-1359	1998	1046-2949	1096	982-1209
1944-1948	1126	925-1327	1456	1146-1767	1382	1104-1661
1939-1943	1012	851-1173	1600	831-2369	1014	897-1131
1934-1938	925	813-1037	1005	894-1116	1315	635-1994
1929-1933	821	734-907	1207	995-1420	1432	565-2300
1924-1928	993	676-1310	893	751-1034	1027	848-1205

Source: Author's own calculations

6.2.3 An analysis of mean per capita expenditure by race

An analysis of mean per capita expenditure would be incomplete without factoring in race. Figure 6.4 illustrates the differences in mean expenditure per capita for the four racial groups in the pseudo panel. The African and Coloured birth cohorts consistently have less mean expenditure per capita than the Indian and White birth cohorts in 2008. Furthermore, a comparison of the mean expenditure per capita for the low expenditure birth cohorts also consistently shows that the African birth cohorts have the least mean per capita expenditure. The White birth cohorts on the other hand have the highest mean per capita expenditure in 2008.

Figure 6.4: Mean per capita expenditure by race of head (2008).



Source: Author's own calculations

What is also shown in Figure 6.4 above is that for African birth cohorts, mean per capita expenditure is higher for the younger birth cohorts and it declines for older birth cohorts. The mean per capita expenditure trend for the white birth cohorts shows high mean per capita for the youngest birth cohort and it declines before starting to rise again. The peak is reached with the 1944-1948 white birth cohort before declining and rising again with the two oldest white birth cohorts. A similar trend is also observed with the Indian birth cohorts.

Table 6.6 below shows that the mean per capita expenditure of the oldest Indian birth cohort (1924-1928) does not fall within the CI range. Therefore, we cannot say with confidence that 95% of the time the coefficient for the expenditure for this birth cohort is captured within the CI. This indicates a small sample problem for this birth cohort. It is therefore unlikely that the mean value given for this birth cohort is reflective of the true population mean.

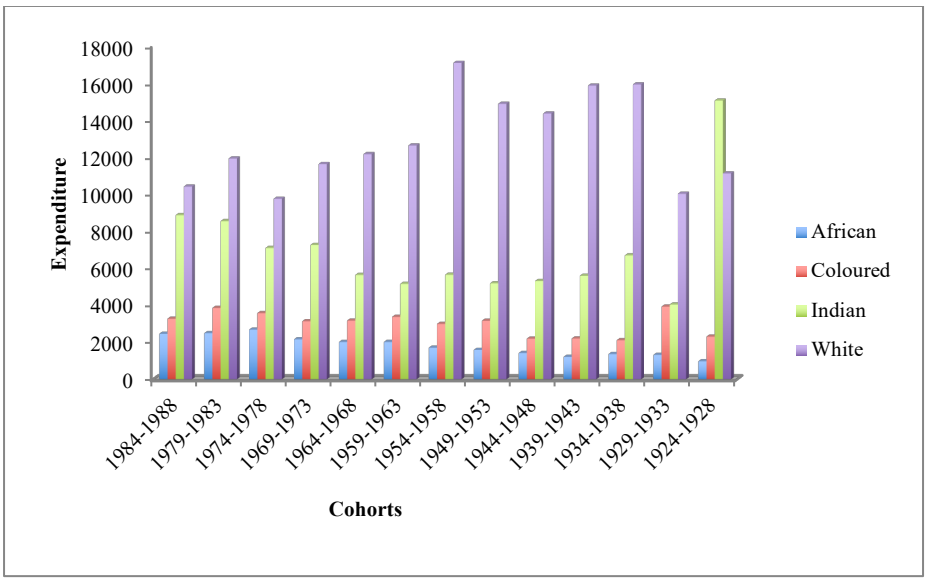
Table 6.6: Mean per capita expenditure by race with CI (2008)

Birth Cohorts	African	CI	Coloured	CI	Indian	CI	White	CI
1984-1988	2045	1772-2319	3110	2018-4202	9712	3314-16110	15128	7765-22491
1979-1983	2214	2017-2410	2710	2104-3307	4299	1759-6839	11639	9063-14215
1974-1978	2089	1921-2258	2790	2349-3231	6676	4185-9168	9586	7843-11328
1969-1973	2125	1896-2355	3192	2560-3824	4245	3082-5409	8291	7182-9400
1964-1968	1852	1710-1994	3096	2463-3730	4055	2948-5163	9073	8085-10062
1959-1963	1688	1581-1795	2553	2147-2960	4488	3124-5852	9311	8241-10382
1954-1963	1560	1432-1689	2625	2283-2967	3881	2412-5351	10448	9090-11806
1949-1953	1476	1359-1593	2234	1901-2567	9881	3606-16157	12723	10914-14532
1944-1948	1196	1106-1286	2275	1931-2619	3825	2860-4791	11667	10087-13248
1939-1943	1022	941-1103	3141	2111-4171	4471	1476-7466	9283	8031-10535
1934-1938	1122	931-1314	2362	1797-2926	9148	568-17728	10160	8483-11836
1929-1933	963	872-1055	2551	1796-3306	13459	-15021-41940	12458	6623-18292
1924-1928	972	866-1078	2528	1347-3708	7682	-90.8-5454	10640	7608-13672

Source: Author's own calculations

Figure 6.5 below extends the analysis of mean expenditure per capita to the year 2010. The African and Coloured birth cohorts continue to have less mean per capita expenditure than the Indian and White birth cohorts, with the former having the least mean expenditure per capita.

Figure 6.5: Mean per capita expenditure by race of head 2010.



Source: Author's own calculation

The trend observed in 2010 for the Indian and White birth cohorts changes in 2010. Mean per capita expenditure for the White birth cohorts increases as the birth cohorts get older and starts declining after the 1954-1958 birth cohort. However, for the Indian birth cohorts mean per capita expenditure is higher for the younger cohorts and declines as the birth cohorts age. A low mean per capita expenditure for the Indian birth cohorts is recorded for the 1959-1963 birth cohort and mean per capita expenditure rises again and is the highest for the older Indian birth cohort. The 1929-1933 Indian birth cohort records the least mean per capita expenditure for this racial category.

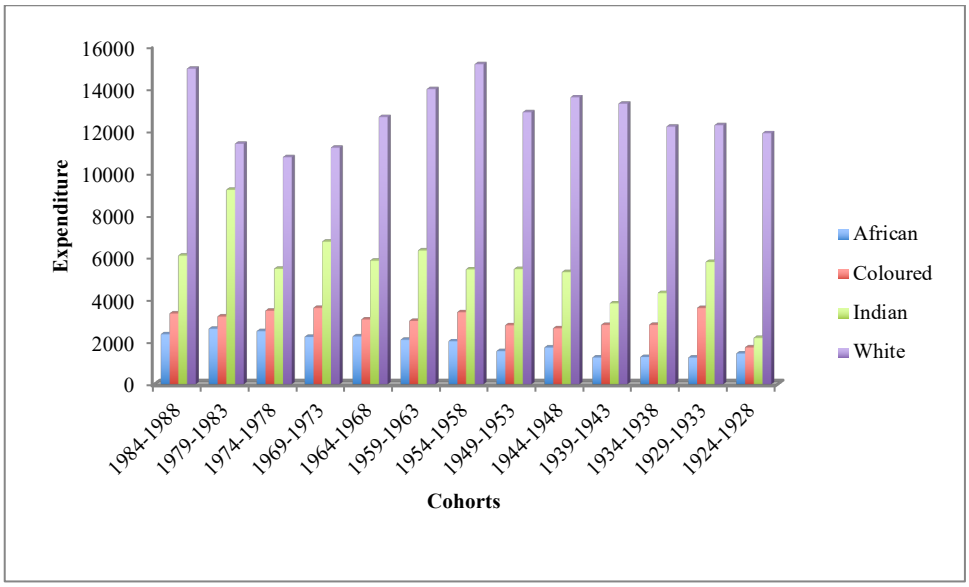
Table 6.7: Mean per capita expenditure by race of head CI (2010)

Birth Cohorts	African	CI	Coloured	CI	Indian	CI	White	CI
1984-1988	2449	2110-2787	3262	2590-3935	8851	1876-15825	10396	7796-12996
1979-1983	2482	2248-2717	3843	2865-4821	8529	4774-12285	11911	9930-13892
1974-1978	2674	2273-3075	3564	2952-4175	7081	5050-9112	9738	8115-11361
1969-1973	2150	1916-2385	3125	2603-3646	7240	4970-9510	11611	10183-13039
1964-1968	2004	1866- 2142	3161	2734-3588	5627	4350-6904	12145	10326-13964
1959-1963	2007	1839-2174	3361	2786-3936	5143	3975-6312	12614	11028-14200
1954-1963	1699	1585-1813	2981	2495-3467	5636	4119-7152	17071	12869-21273
1949-1953	1571	1454-1688	3153	2486-3820	5173	3984-6361	14863	12130-17596
1944-1948	1409	1270-1549	2187	1869-2506	5295	3660-6930	14339	11821-16854
1939-1943	1204	1088-1320	2196	1789-2603	5580	1640-9520	15846	13229-18464
1934-1938	1355	1054-1657	2099	1561-2637	6681	2058-11304	15910	12055-19764
1929-1933	1310	1130-1491	3918	2059-5777	4031	2554- 5509	10008	8444- 11572
1924-1928	966	849-1083	2296	1143-3450	**15039**	-	11105	7656- 14555

Source: Author's own calculations

The oldest Indian birth cohort in 2010 again has a mean per capita expenditure value that is not within the upper bound and lower bound CI also indicating the small sample of the Indian population at this age in the surveys. As was observed in 2008, the mean value is not reflective of the true population mean of this birth cohort in the true population.

Figure 6.6: Mean per capita expenditure by race of head (2014).



Source: Author's own calculations

In 2014, the mean expenditure per capita for all the races is lower than in 2010. The same hierarchy of the least to highest mean per capita expenditure by race observed in 2008 and 2010 remains in 2014. The youngest white birth cohort has the highest mean per capita expenditure in 2014 and this indicates a recovery from 2010 where they recorded a lower mean per capita expenditure value. The 1949-1953 white birth cohort also made a similar recovery. Younger African birth cohorts have higher mean per capita expenditure compared to older birth cohort within the same race.

Table 6.8: Mean per capita expenditure by race of head CI (2014)

Birth Cohorts	African	CI	Coloured	CI	Indian	CI	White	CI
1984-1988	2359	2096-2623	3340	2402-4279	6078	3743-8413	14924	11013-18835
1979-1983	2619	2253-2985	3198	2343-4054	9193	5938-12448	11371	9594-13149
1974-1978	2504	2271-2738	3473	2351-4596	5455	3155-7756	10735	9090-12380
1969-1973	2233	2032-2434	3605	2946- 4265	6738	3954-9522	11185	9496-12874
1964-1968	2253	2034.274 2473.508	3060	2513- 3607	5840	4042-7639	12636	10163- 15109
1959-1963	2098	1921-2275	2992	2565-3419	6320	4175-8464	13955	11602-16307
1954-1963	2030	1826-2235	3402	2518- 4285	5428	4127-6729	15136	12653-17619
1949-1953	1560	1425-1694	2787	2017-3558	5442	3506-7377	12861	10902-14820
1944-1948	1730	1486-1974	2643	2082-3203	5297	3284-7311	13572	10813-16332
1939-1943	1258	1133-1383	2800	2096-3504	3824	1877-5771	13266	10481-16050

1934-1938	1282	1092-1472	2803	1913-3693	4308	1804-6812	12182	9691-14673
1929-1933	1256	977-1535	3600	1071- 6289	5778	3956-7599	12250	8971-15529
1924-1928	1449	1073-1826	1734	783-2685	**2185**	-	11872	4332-1941

Source: Author's own calculations

Again, in 2014 the mean value for expenditure per capita for the 1924-1928 Indian birth cohort does not have CI, this cohort only has one household in this year. The mean for this cohort as argued earlier is not plausible for the mean of the true population mean

6.2.4 Discussion: Distribution of mean per capita expenditure for birth cohorts

To my knowledge there are no studies in South Africa on expenditure patterns for birth cohorts. The analysis of the differences in the mean expenditure per capita of birth cohorts done in the previous section is a contribution to this end. In Figure 6.1 earlier in this chapter, the analysis of the differences in expenditure for all birth cohorts showed that mean expenditure per capita increased for younger birth cohorts, with the highest mean expenditure per capita observed for birth cohorts born between 1933 and 1948 before declining with the oldest birth cohort.

This increase in mean expenditure per capita observed for younger birth cohorts could be an indication of an improvement in labour market earnings as birth cohorts get more working experience as they age. The sudden decline witnessed with the oldest birth cohort could partly be a result of the small sizes of older cohorts discussed earlier in Chapter Six, and is thus error, but could also a result of the absence of labour market derived income that has worsened over time for this group.

An analysis of the LCS of 2008, and 2014 and the IES of 2010, indicates that the major sources on income within these surveys is dominantly from salaries/wages and social grants. To illustrate this Table 6.1 below summarises the percentage contributions made by the two dominant sources of income captured in the LCS of 2008 and 2010 and the IES of 2010.

Table 6.9: Percentage contribution of salary/wages and social grants to total income

Description	Source of income	2008 (%)	2010 (%)	2014 (%)
All	Salary	44.21	45.09	42.19
	Social grants	26.72	27.57	32.51
MHH	Salary	55.70	55.31	52.16
	Social grants	14.59	15.58	18.65
FHH	Salary	29.98	31.92	29.86
	Social grants	41.86	43.17	49.62
Urban	Salary	53.85	53.72	50.80
	Social grants	18.48	19.59	24.05
Rural	Salary	28.61	29.53	28.79
	Social grants	40.05	41.97	45.66
African	Salary	42.90	42.75	40.08
	Social grants	29.30	30.44	34.99
Coloured	Salary	54.53	53.29	51.57
	Social grants	22.19	24.27	29.84
Indian	Salary	50.77	53.22	45.70
	Social grants	18.88	17.07	22.40
White	Salary	50.87	55.98	52.58
	Social grants	5.41	6.15	7.71

Source: Author's own calculations using the LCS of 2008, 2014 and the IES of 2010

The LCSs have a variable that capture the sector of employment and it distinguishes between formal, informal and work in private households. There also exists another variable that captures the existence of verbal contracts and written aggreements. However, similar variables are absent in the IES of 2010. Instead the IES has a variable that captures whether an individual is employed as well as involvement in self employment. Due to the incomparabilty of the employment related questions between these surveys, the analysis relating to the employment of the birth cohorts will not be conducted.

Based on the findings in Table 6.1, MHH birth cohorts have a higher percentage contribution of their income from salaries/ wages (55.68%, 55.31% and 52.16% in 2008, 2010 and 2014 respectively). This partly drives their mean expenditure per capitas higher in all the years under analysis as FHH birth cohorts have a lower percentage contribution of their income from a similar source (29.97%, 31.92% and 29.86% in 2008, 2010 and 2014).

Heightening the expenditure inequality between MHH and FHH birth cohorts, is the fact that FHH tend to rely more on social grants than they do on salaries/wages. For instance, Table 6.1 above shows that FHH have more than 40% of their income coming from social grants, whilst salaries/wages contribute less than 30% in 2008 and 2014 and close to 32% in 2010. On the other hand for MHH birth cohorts, social grants contribute 14.56%, 15.58% and 18.65% in 2008, 2010 and 2014.

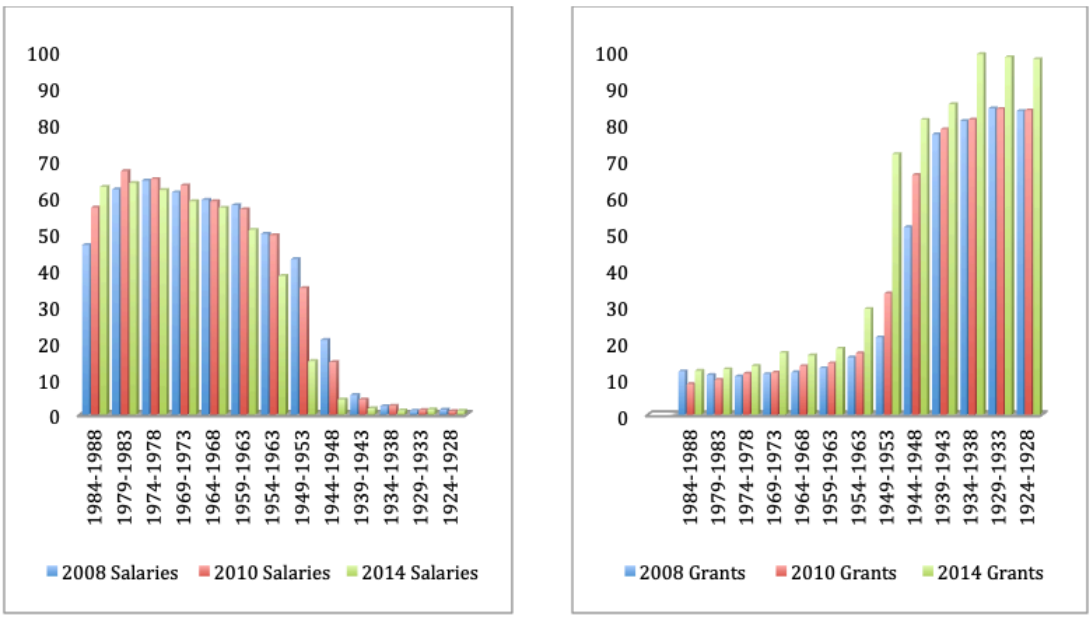
Earlier it was observed that the mean expenditure per capita for MHH birth cohorts increases as the birth cohorts get older. This suggests that for MHH, the major driver of expenditure differences is labour market income. Younger birth cohorts for MHH earn less than older MHH birth cohorts and part of this could be attributed to less working experience. The decline observed for older birth cohorts could be attributed to the absence of labour market related income and reliance on pensions and old age grants.

Extending a similar analysis to FHH birth cohorts, their mean expenditure per capita declines for younger cohorts, and almost flattens out before increasing at older birth cohorts. It is likely that as FHH start their own families, they pay a wage penalty; childbearing might result in them losing hours of work and therefore earning less resulting in less mean expenditure per capita. Once they reach old age, as this burden is lifted, their expenditures also start increasing. Despite the increases, the mean expenditure per capita at old age remain lower than MHH birth cohorts.

Birth cohorts in rural areas depend more on social grants (40.05%, 41.97% and 45.66% in 2008, 2010 and 2014), and salaries and wages contribute less to their income (28.61%, 29.53% and 28.79%) in similar years. On the other hand, birth cohorts in urban areas cite salaries and wages as the major source of income (53.85%, 53.72% and 50.80% in 2008, 2010 and 2014), whilst social grants contribute 18.48%, 19.59%

and 24.05% within the same period. Figure 6.7 below extends the analysis on the contribution of salaries of salaries and social grants on income for all birth cohort

Figure 6.7: Percentage contribution of Salaries and social grants to income for all birth cohorts.



Source: Author's own calculations

Figure 6.7 shows that younger birth cohorts are more likely to depend on income from salaries and wages more than what they receive from social grants. What is also discernible is that income from salaries and wages starts by increasing as birth cohorts age. This indicates that younger birth cohorts as new entrants into the labour market gain in wages as they gain more experience. However, these gains start reducing at a certain age and stagnate before declining. Older birth cohorts have the least contribution to their main income from salaries In fact, they rely mostly on social grants.

Based on the percentage contribution of salaries/wages and social grants in the LCSs and the IES, the discussion has found that these two drivers of income are key in understanding the differences in the mean expenditure per capita of the birth cohorts in the pseudo panel. Whilst the low amounts of income associated with social grants drive down the mean expenditure per capita of birth cohorts in rural areas as well as FHH birth cohorts, expenditures are also driven up for urban areas and MHH with salaries and wages and this explains the contrasts in these categories of birth cohorts in the pseudo panel.

It should be noted that expenditure can also be driven by other variables beyond income from salaries and social grants. For instance, households can spend income from borrowing or savings, but the savings variables in the LCSs are not comprehensive and in the IES, savings data are absent. The question on the main source of income in the LCSs has a section that captures income from savings. However, income from this category remains very low. For instance, only 0.42% and 1.39% of households had income from this category in 2008 and 2014 respectively. A look at the savings culture of South Africa also supports this lack of savings in South African households. Cronje and Roux (2010) found that in South Africa, there exist very low levels of personal savings. Instead, South has been found to have a 'debt culture' than a savings one and part of this is attributed to the huge income inequalities in the country which results in a more immediate need to survive and a reduction in the ability to save (Cronje and Roux 2010). Based on the discussion from the above argument, it is expected that a larger portion of households are more likely to borrow than they are to save. The IES does not have questions on savings and borrowing and therefore it is impossible to use related arguments to explain the differences in mean expenditure per capita for the birth cohorts.

6.3 Cohort analysis: Age, cohort and time effects

Estimating cohort, age and time effects is a difficult task because all three are linearly correlated (Deaton 1997; Guillerm 2017). The previous sections analysed changes in expenditure patterns observed for all the birth cohorts in the pseudo panel. The findings from the previous section largely support the evidence that mean expenditure per capita starts by increasing when the birth cohorts are younger and then starts declining as they get older when analysis is done for all birth cohorts. However, this increase in mean expenditure per capita cannot be entirely due to ageing. When analysing cohorts, Rentz et al. (1983) argue that longitudinal designs confound age and period effects. In this case, we cannot attribute the changes in expenditure patterns observed in the previous section to the effect of ageing alone. The reason for the challenge is that it is difficult to isolate the age variable alone experimentally and then attribute all the observed changes in expenditure to ageing (Baltes 1968; Glen 1979; Riley 1973; Schaie 1965 as cited in Rentz, Reynolds and Stout 1983). Rentz et al. (1983:13) argue that cohort analysis 'is a group of methods designed to separate cohort, age and period effects'.

Modigliani (1966) argued that consumption by households or individuals is influenced by current and future income and this became the basis of the Life cycle hypothesis, a theory important in cohort analysis. The theory assumes that households/individuals seek to smooth their consumption in a lifetime, and they tend to save when they have high levels of income and borrow when income levels are low, to smoothen consumption (Modigliani 1966).

Using consumption theory, the cross-sectional variance in expenditure is supposed to increase over time as the different cohorts are affected by economic shocks differently. In addition, the income profiles within each cohort tend to fan out as well, signalling an increase in variance, which should reflect in an increase in variance in consumption. Thus, to provide a more accurate pattern of expenditure, there is need to control for age, cohort, and time effects.

i. Age effects

In the pseudo panel, as the household heads get older, different job opportunities become available to them and to other members of their households within similar households. As a result, this difference affects consumption choices and creates dispersion on their consumption paths. The extent of this dispersion points to the level of within cohort inequality. Therefore, there is need to control for age effects in order to understand what is driving the expenditure changes, (the surrogate for well-being) that were observed in the previous section.

Guillerm (2017) argue that different generations are likely to experience different economic opportunities due to changes in policies or economic conditions. In South Africa for instance there is a clear difference in terms of the job opportunities available in the Apartheid period and the post-Apartheid era. Wiehhahn et al. (1982) and Crankshaw (1997) found that during the apartheid era, legislation such as the Bantu Labour Amendment Act of 1970, the Regulation of Mines and Works Act of 1911, and the Industrial Reconciliation Act of 1970 reserved skilled and other better paying jobs to white employees. However, such legislation was later reversed in the post-1994 period with legislation such as the Employment Equity Act of 1998 (Martin and Durrheim 2006). This transition in labour legislation in South Africa clearly marks a shift in job opportunities for all South Africans.

Thus, lumping all the cohorts together when analysing the trends in the previous

sections may lead to misleading conclusions. In particular, there is likely to be less inequality among the younger cohort who are just joining the labour market as compared to the older generation about to retire. In this case, aggregating the cohorts can hide some of these within cohort dynamics. To control for age effects, the study uses age dummies for household heads aged 25-75 years and then uses age "25" as the reference group.

ii. Cohort effects

Schaie (1965) attribute cohort effects to the use of the time-lag design of pseudo panels that confounds the effects of cohort membership over time. Cohort effects are associated with the differences that are observed between people as a result of different behavioural characteristics related with the time they were born. These behavioural characteristics in turn affect their consumption and spending patterns. This study by generating birth cohorts that are separated by five-year band attempted to control for this effect. The assumption here therefore is that household heads that are in the same birth cohort are more likely to have similar expenditure behavioural patterns. But this does not adequately control for cohort effects. In the regression, yearly birth-cohorts were temporarily created to adequately cover all household heads aged between 25 and 75 years old. The regression model therefore has 50 yearly birth cohorts to attempt to control for cohort effects.

iii. Time effects

Rentz et al. (1983) state that time effects are caused by environmental influences that are associated with a given time period. Individuals are affected differently by economic shocks and business cycles in an economy. For instance, a recession might result in some individuals losing their jobs, but some individuals may actually advance their careers in a similar time period. Whereas cohort effects are more likely to affect people's behavioural patterns in what they consume, time effects are more inclined towards non- behavioural conditions such as those that are a result of the prevailing economic environment. The result is that their labour market related income is affected differently and this impacts on their expenditures. To control for time effects, the regression includes dummies that represent the three time periods under analysis in this study (year category variables i.e. 2008, 2010, and 2014).

6.3.1 Regression Analysis: Exploring the Changes in Expenditure for Different Birth Cohorts (2008- 2014)

This section using the Consumption theory attempts to explain the changes observed in expenditure patterns for the birth cohorts in the previous section, by analysing the conditional mean and variance of the residuals from a multivariate regression. The analysis in this section follows the work of Aguiar and Hurst (2013) who analysed the hump shape of expenditure over the life cycle as well as the increase in cross sectional consumption inequality as individuals age. This section analyses these two main concepts of the life cycle consumption expenditures using expenditure variables. However, this study will not separate expenditures into durable and non-durable expenditures as was demonstrated in Aguiar and Hurst (2013).

In the multiple regression conducted[14] in this chapter the dependent variable is the log of expenditure while the regressors include dummies for age, cohort, time and other household characteristics related to expenditure. The regression model is as follows:

$$\ln \boldsymbol{E_{it}} = \boldsymbol{\beta_0} + \boldsymbol{\beta_{age}Age_{it}} + \boldsymbol{\beta_c Cohort_{it}} + \boldsymbol{\beta_t D_t} + \boldsymbol{\beta_t Household_{it}} + \boldsymbol{\varepsilon_{it}} \qquad (1)$$

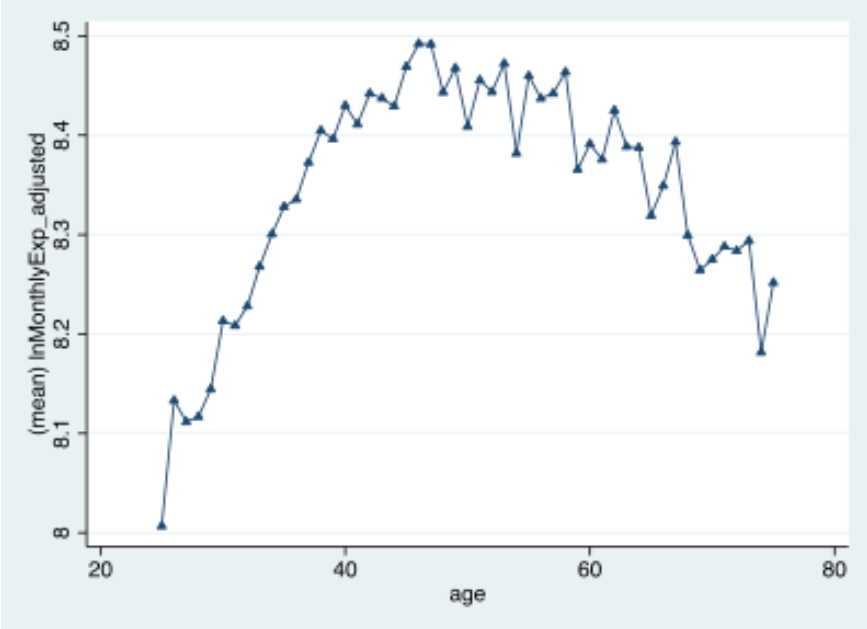
where:

i. lnE_{it} is the log of expenditure of household *i in* a time period *t*
ii. Age_{it} represents the vector of one-year age dummies, in this model 50 age dummies are included for household heads aged 25-75 years.
iii. $Household_{it}$ is a set of dummy household variables which are; type of settlement, level of education, province, gender of household head and race of household head.

The objective is to purge out the observables characteristics that have an effect on expenditure among South African households. Because the regression does not give the variance, I extract the residuals in the regressions. For each age category we calculate the mean by adding up all the residuals. We then use the mean for each age category to get the graph in Figure 6.8 below.

[14] The regression results are attached as Appendix 1.

Figure 6.8: Conditional Mean of Log Expenditure by age.



Source: Author's own calculations

Figure 6.8 above depicts the conditional mean of the log of monthly household expenditure for the ages 25-75 and 25 is the omitted group. Each point shows the conditional mean for the specified age group in relation to 25-year olds. The selection of the age 25 allows a youthful population that is just entering the labour market or has just joined the labour force to be captured. Figure 6.8 above confirms that expenditures are hump shaped over the lifecycle. This is expected and supports similar findings in Aguiar and Hurst (2013).

Using the Life Cycle Hypothesis to analyse expenditure patterns of birth cohorts, the conditional variance of the log of monthly expenditure is extracted for the birth cohorts from the regression. The variance is a measure of dispersion. In this section, it is conditional on the variables included in the regression mentioned earlier. In the case where there is no expenditure inequality between birth cohorts, we expect to see a straight line.

Equation (2) below estimates the variance of the log of monthly expenditure;

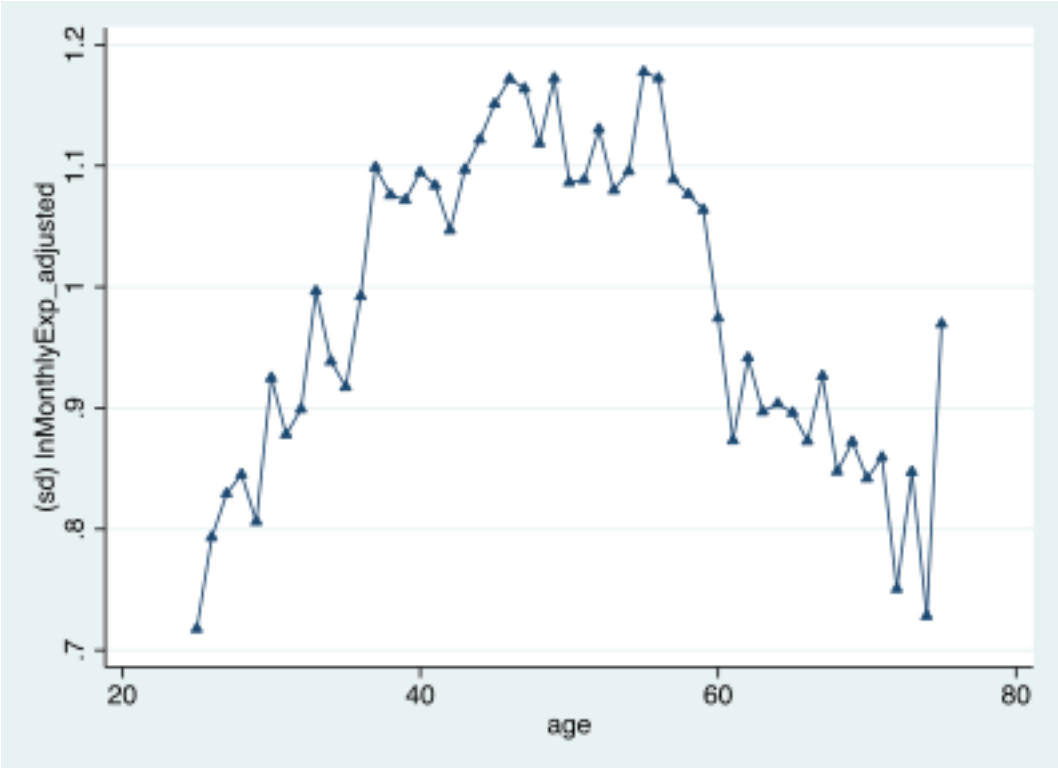
$$(\boldsymbol{\sigma^2})_{it} = \boldsymbol{\sigma_0} + \boldsymbol{\sigma_{age}Age_{it}} + \boldsymbol{\sigma_{cohort}Cohort_{it}} + \boldsymbol{\eta_{it}} \qquad (2)$$

In Equation (2) above, $(\boldsymbol{\sigma^2})_{it}^{k}$ is the variance of $\boldsymbol{\varepsilon_{it}}$calculated in equation (1) previously. As in Aguiar and Hurst (2013:451) and Deaton and Paxson (1994),$\boldsymbol{\sigma_0}$ represents the

vector of age coefficients for expenditure, shows the changes in variance over a lifecycle. Figure 6.9 below shows this conditional variance of the log of monthly expenditure of the residuals that were extracted from equation (1).

What is illustrated below shows that variance increases until the late 50s. The assumption is that labour market earnings drive these expenditure inequalities and as individuals start to retire, they lose labour market earnings, and this starts to reduce their expenditure inequalities.

Figure 6.9: Conditional variance of the log of expenditure for yearly birth cohorts over time.



Source: Author's own calculations

Moreover, the variance in the log of expenditure is expected to increase as household heads get older. This is because we expect income from the labour market to increase as people gain more work experience. As the birth cohorts get older, the variance is expected to start falling after people retire and start losing their earnings from the labour market. Therefore, as people move past the 50s age mark, the variance should drop more sharply so people are more equal as they are relying on pensions that are not large unless they have savings. But most people are likely to be relying on pension after they reach retirement age. Aguiar and Hurst (2013) also observed similar results. Using the

permanent income hypothesis, they reported an increase in the variance of consumption within cohorts over time as the cohorts age.

6.4 Changes in poverty status of birth cohorts in pseudo panel

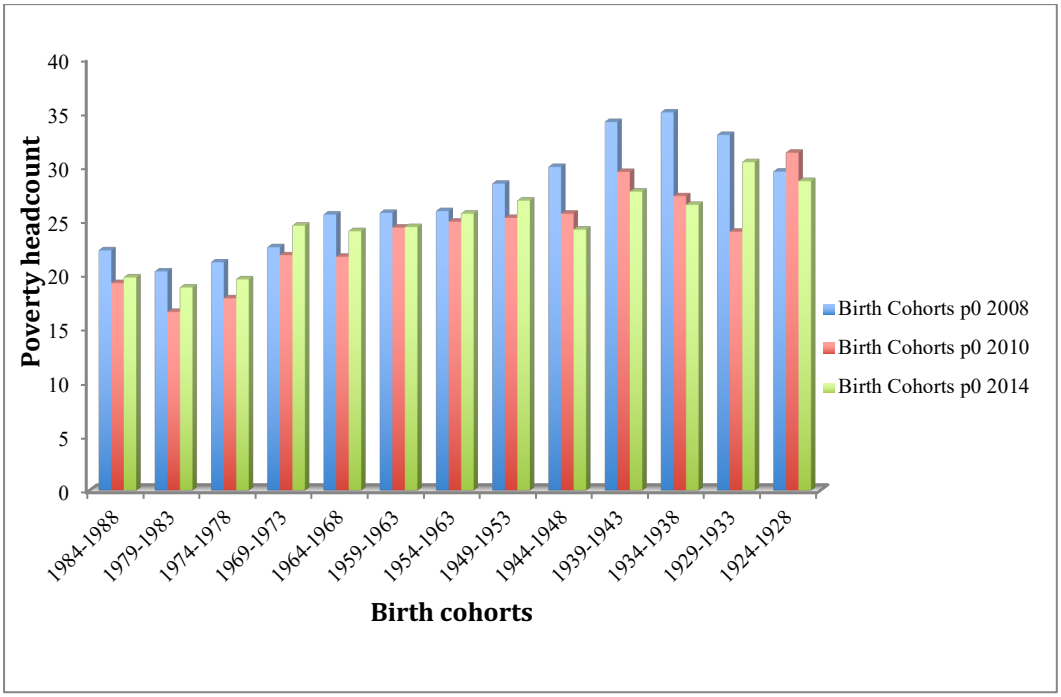
The first section of this chapter analysed the expenditure patterns of the birth cohorts and in the previous segment the analysis attempted to explain the variance on the log of expenditure over a life cycle. In this last section, the analysis is extended to examine the poverty status of the birth cohorts in the pseudo panel between 2008 and 2014. This study uses the Lower Bound Poverty Line and all prices are expressed in 2014 prices. The poverty estimates in this section use the Foster Greer Thornbecke (FGT) poverty indices[15].

This chapter also tested for first order stochastic dominance using the (Stochastic Dominance, Theory, and Practice (DASP), a software package developed by Araar and Duclos (2012). The test was conducted for pairwise groups and a full dominance is assumed when cumulative distribution curves do not cross. (Heshmati et al. 2019). The poverty dominance test done seeks to provide a non-ambiguous welfare ranking. Dominance tests are required in poverty and inequality ranking since these can be offset by the use of different poverty lines, or the use of different aggregation procedures (Gondwe 2015).

Using the poverty headcount, Figure 6.10 below shows that between 2008 and 2010, poverty reduced for all birth cohorts except for the oldest birth cohorts where it increased in the same period. In 2014, poverty increased for most birth cohorts except the 1944-1948, 1939-1943, 1934-1938 and 1924-1928 birth cohorts where it reduced. What is also evident in Figure 7.10 below is that older birth cohorts experience higher poverty rates than younger birth cohorts. Stats SA (2017c:92) in an analysis of age cohorts presented in Chapter Three using the LBPL found that in 2008 and 2009, the 65+ age group had high poverty rates, but this reduced in 2011 and slightly increased in 2015.

[15] **Appendix 2** has tables of the poverty estimates using the Poverty Headcount, Poverty Gap and Poverty Gap Squared for all the Figures in this section.

Figure 6.10: Poverty headcount for all birth cohorts (Poverty headcount).



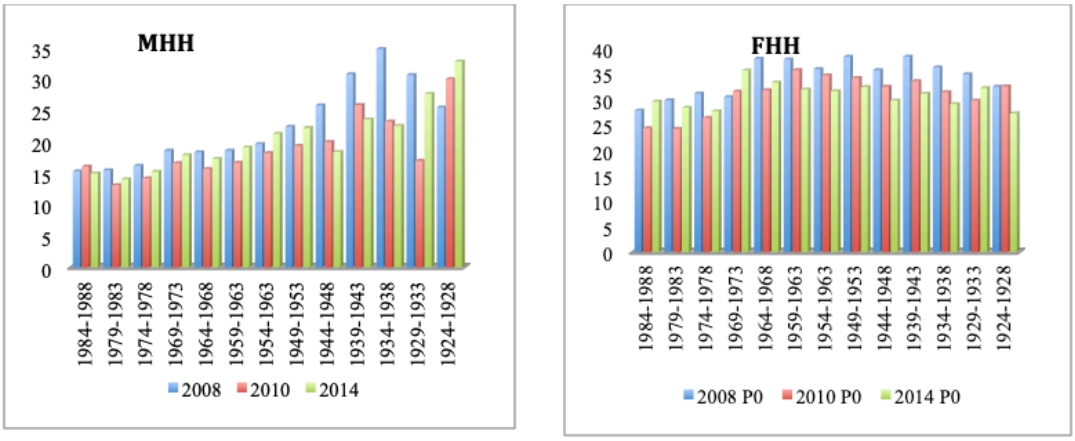
Source: Author's own calculations

A first order stochastic dominance test done for the 2008 and 2010 time periods and there is a clear dominance for all the birth cohorts. But between 2010 and 2014 there is no clear dominance for some of the birth cohorts. The poverty estimates in the graph above shows that poverty fell for most of the birth cohorts between 2008 and 2010 and then increased again in 2014. To check for the robustness of these estimates, poverty dominance was checked for pairwise sub-groups. The first pairwise analysis assessed the dominance of poverty rates for birth cohorts between 2008 and 2010 and the results are dominant using the first order of stochastic dominance for all birth cohorts in this time period. The second pairwise subgroup tested for dominance for poverty rates between 2010 and 2014 and during this period full dominance was not satisfied.

6.4.1 A gendered analysis of poverty estimates

Further poverty analysis was pursued by gender of the household heads. Table 6.11 shows the differences in poverty for MHH and FHH in 2008, 2010, and 2014. FHH have higher poverty rates than MHH and this is observed for all birth cohorts. The results also show that for older birth cohorts, the differences in poverty estimates are less pronounced than the differences in poverty incidence observed for younger MHH and FHH.

Figure 6.11: Poverty incidences by head of household (Poverty headcount).



Source: Author's own calculations

Figure 6.11 also illustrates how poverty changed between 2008 and 2014 for all MHH birth cohorts in the pseudo panel. Younger MHH birth cohorts experience lower poverty rates than older MHH birth cohorts. A majority of MHH birth cohorts experienced a decline in poverty between 2008 and 2010 but in 2014, poverty increased for most MHH birth cohorts.

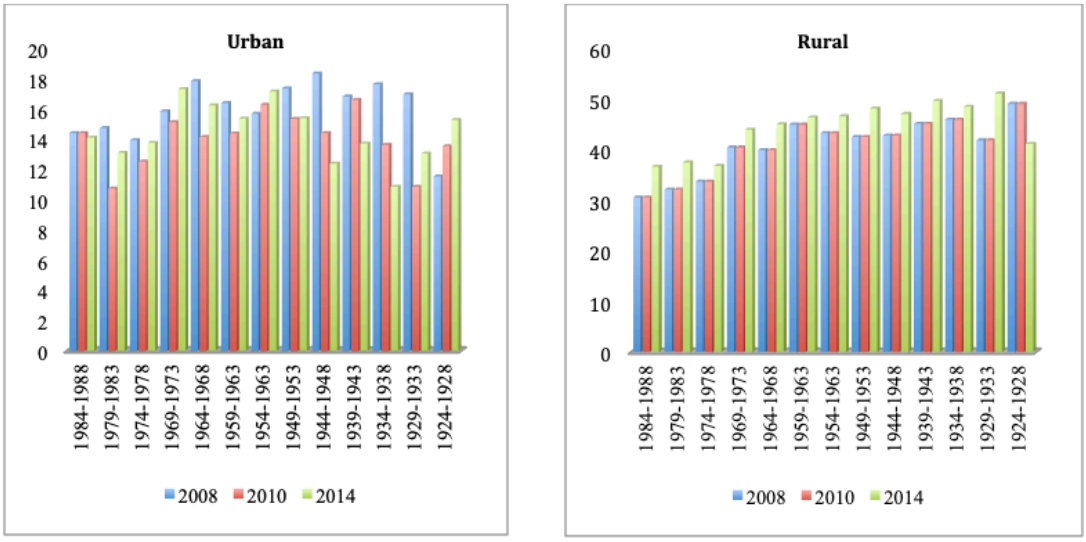
Extending the analysis to FHH the trend observed for MHH is not necessarily repeated for FHH in the pseudo panel between 2008 and 2010. As much as Figure 6.11 also suggests that younger FHH birth cohorts have lower poverty rates when compared to older birth cohorts, the differences in poverty estimates for these two broad groups are less pronounced when compared to Figure 6.12 above.

Poverty dominance was also tested on gender pairwise subgroups. In regard to this, poverty dominance was tested for the poverty rates in 2008 between FHH birth cohorts and 2010 MHH birth cohorts, the second pair consisted of poverty rates for FHH birth cohorts in 2010 and poverty rates for MHH in 2014. Full dominance was recorded for these pairs at a gendered level, suggesting that the differences in the poverty experiences between these two groups are robust. This suggests that FHH are more vulnerable to poverty irrespective of the poverty line used

6.4.2 A geographic analysis of poverty incidence

This section extends the poverty analysis for birth cohorts by geographical variables. Figure 6.14 below shows the differences in poverty rates between these two areas.

Figure 6.12: A comparison of poverty incidences for birth cohorts in urban and rural households (Poverty headcount).



Source: Author's own calculations

What stands out in Figure 6.12 above are the extreme differences in poverty estimates, which are more than double for birth cohorts in rural areas when compared to urban areas. Poverty estimates for all birth cohorts in urban areas are all below 20%, yet in rural areas poverty estimates reach over 50%. To get a better understanding of the differences in poverty estimates within rural and urban areas for all birth cohorts,

In urban areas, poverty incidences for all birth cohorts are crowded between 10.79% (the lowest poverty rate for birth cohorts in urban areas) and 18.25% (the highest poverty rate for birth cohorts residing in urban areas). A majority of birth cohorts in urban areas also experienced higher poverty rates in 2008. However, this declined for a majority of birth cohorts in urban areas only to increase again in 2014 as was the case with the 1969-1973 birth cohort or remained almost the same (1949-1953) or declined (1934-1938).

Extending the analysis to birth cohorts in rural areas, Figure 6.16 below illustrates how poverty incidences differ in rural areas for all birth cohorts. Whereas the differences in poverty estimates for birth cohorts in urban areas were crowded within relatively lower poverty thresholds, birth cohorts in rural areas have poverty rates that are not only higher but are a bit spread out between 32.46% and 56.71%.

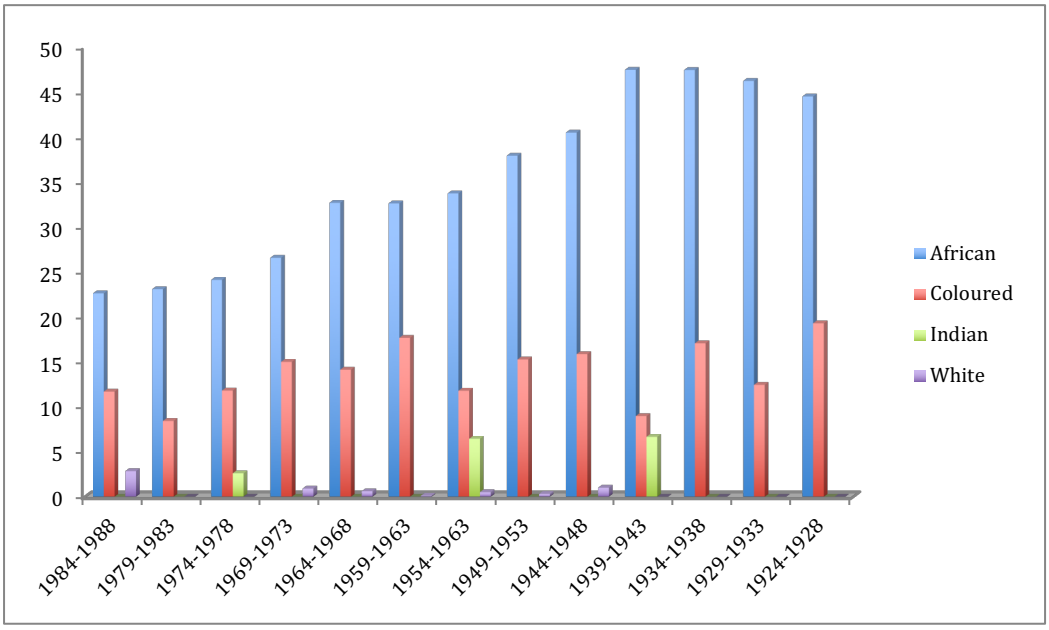
Results from a test of stochastic first order dominance shows that the full dominance condition is satisfied between the 2008 poverty estimates for birth cohorts residing in

rural areas and the 2010 poverty rates for urban birth cohorts. The same pairwise test is conducted for 2010 and 2014 between these two groups of birth cohorts. The poverty rates between these pairs showed full dominance using the first order of stochastic dominance.

6.4.3 A racial analysis of poverty incidence

Poverty analysis in South Africa would be incomplete without a racial component. Figure 6.15 below illustrates the differences in poverty incidences of the birth cohorts by race. The poverty incidences for the birth cohorts when disaggregated by race are significant. In 2008, African birth cohorts were the most affected with poverty, with older African birth cohorts being the most vulnerable to poverty. For the Coloured birth cohorts, poverty headcount rates range between 8.42% to 19.28%. Relatively lower poverty estimates are recorded for the White and Indian birth cohorts.

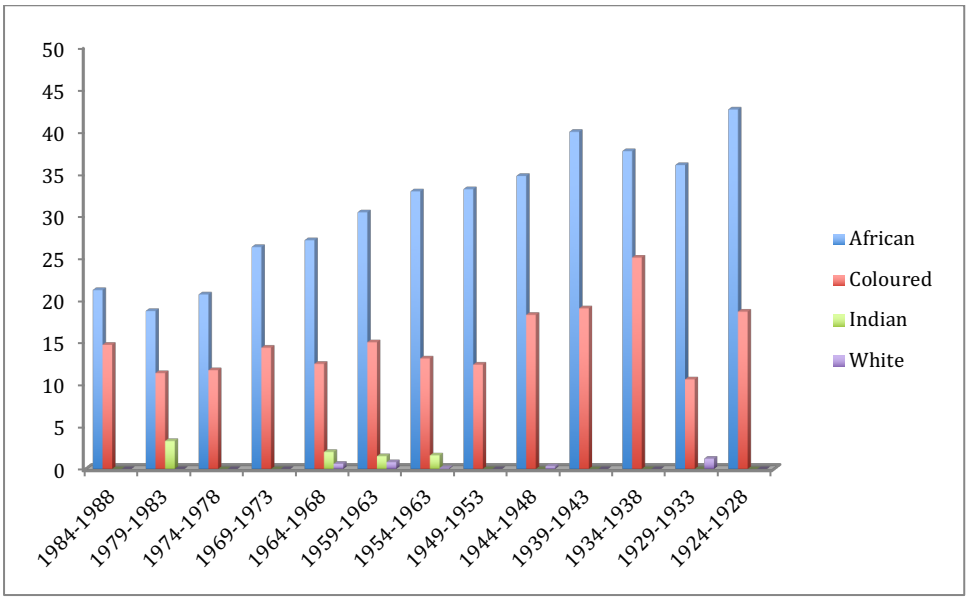
Figure 6.13: Poverty headcount for all birth cohorts by race (2008).



Source: Author's own calculations

In 2010, African birth cohorts continued to have the highest poverty incidences. Despite this, there was a decline in poverty incidence for a majority of the birth cohorts in 2010. Specifically, there was a reduction in poverty for all the African birth cohorts whilst the other racial groups experienced increases and decreases of poverty incidences.

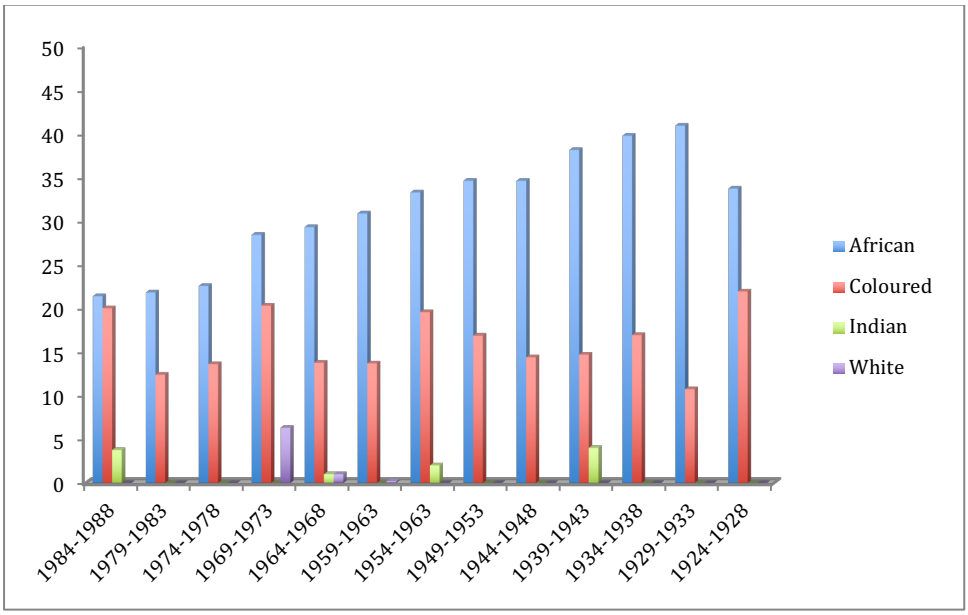
Figure 6.14: Poverty headcount for all birth cohorts by race (2010).



Source: Author's own calculation

In 2014, the majority of African birth cohorts experienced an increase in poverty incidence when compared to 2010. The poverty incidence of the oldest African birth cohort as well as the 1934-1938 African birth cohort reduced consistently between 2008 and 2014. However, all other African birth cohorts experienced an increase in poverty between 2010 and 2014. This is also observed for a majority of the Coloured Indian and White birth cohorts.

Figure 6.15: Poverty headcount for all birth cohorts by race (2014).



Source: Author's own calculation

First order dominance is established when pairwise comparisons are made between African birth cohorts against the other racial groups. The results are robust and they suggest that the African birth cohorts are the most severely affected by poverty.

6.4.4 Discussion: Poverty analysis for the birth cohorts

An analysis of poverty rates when analysed for all birth cohorts shows that poverty increases as birth cohorts get older and declines for the oldest birth cohort. Given the results presented earlier that showed that mean expenditure per capita was increasing as birth cohorts got older, this increase in poverty as birth cohorts got older is surprising. The expectation therefore is that poverty would be higher for the younger birth cohorts.

Extending the analysis of poverty headcount to MHH and FHH, the results indicate that FHH are more susceptible to experiencing poverty than MHH. In fact, FHH experience higher poverty rates than MHH. The findings in the first section in Table 7.1 showed that the FHH are more reliant on social grants as their main source of income when compared to income form salaries/wages. This is compounded by the fact that men tend to have a higher percentage contribution of their income from employment sources. Therefore, the differences in poverty rates for FHH and MHH partly emanates from these observations. The reduction in the differences in poverty rates between MHH and FHH for older birth cohorts indicates the discontinuation of labour market income. Thus, MHH also start to record higher poverty rates after leaving the labour market.

The differences in poverty rates between birth cohorts in urban and rural areas are also hard to miss. Similar to the arguments on the contribution of sources of income to poverty rates for MHH and FHH, the observed differences between these two geographic locations also hint on the same conclusions. Birth cohorts in urban areas receive most of their income from the labour market whilst in rural areas, the major source of income is social grants.

The racial differences in poverty for the birth cohorts show that African birth cohorts are the most vulnerable to poverty when compared to other racial groups. The least poverty incidences are experienced with the Indian and White birth cohorts. It is worth mentioning here that the poverty estimates found in this study for these two groups also reflect the small sample sizes of these two racial groups in the pseudo panel.

6.5 Conclusion

This chapter has analysed expenditure patterns as well as poverty rates of 13 birth cohorts that were 'tracked' in three time periods using a pseudo panel. An analysis of the expenditure patterns conducted in the first section of this chapter concluded that there exist different expenditure patterns within and across birth cohorts and the differences are sustained if the analysis is restricted to MHH and FHH. Differences between MHH and FHH shows that the FHH have lower mean expenditure per capita than MHH in the three time periods under analysis. Disaggregating the different patterns in expenditure by geography also confirms that birth cohorts in urban areas have higher expenditures than birth cohorts in rural areas.

The second section of this thesis conducted a multivariate regression estimation to estimate the trends and patterns of the conditional mean and cross-sectional variance. The use of the conditional mean and variance sought to control for observable characteristics across cohorts such as cohort, time and age effects in order to try and analyse the observed expenditures in the previous section are hump shaped over the lifecycle. The mean was also conditional on household level characteristics including settlement type, gender, and race of household and province. The finding supported that expenditures are hump shaped whilst an analysis of the variance showed increasing inequalities as the birth cohorts get older which starts declining after the cohorts are in their late 50s, an indication of the discontinuation of labour market earnings.

Poverty seems to affect the old aged birth cohorts more than it affects the younger birth cohorts. Part of this can be attributed to increasing household sizes as the cohorts get older. A gendered approach to poverty of the birth cohorts showed that FHH experience poverty rates that are close to double the percentages observed for MHH in the three time periods of the pseudo panel. Further disaggregation of poverty rates by geography proves the existence of extensive expenditure inequalities between households in urban and rural areas. These differences were likely to be driven by unequal contribution of salaries/wages to main income of birth cohorts within the two genders and geographical spaces and an over reliance on social grants by FHH and households in rural areas.

Chapter Seven

Poverty measurement using alternative data: Lessons and future analysis

7.1 Introduction

Data are central in today's development agenda. Existing development targets have a global focus (SDGs), and co-exist with regional development frameworks such as the Africa Union's Agenda 2063. At country level, national governments have their own mandates, with the National Development Plan (NDP) being the national key blueprint to South Africa's development trajectory until 2030. Monitoring progress on these goals translates to a higher demand for data than experienced during the previous two decades. As a result, data availability is expected to increase to meet this increasing need for household survey data.

It is therefore imperative to consider the pressure exerted on National Statistical Offices (NSOs) in developing countries especially in Africa where some of the biggest data gaps have been observed. Sub Saharan Africa has the greatest need for 'better data' (Sandefur and Glassman, 2015:1). This is compounded by the fact that the region continues to lag behind poverty reduction efforts globally. The fact that 27 out of the 28 poorest countries globally are located in sub-Saharan Africa (The World Bank 2018) cannot be overlooked. This leaves the region with a dual burden - having the highest percentage of poor countries globally and at the same time home to a majority of countries classified as data inadequate.

This *status quo* is less promising when SDGs data requirements are added into the picture. SDGs have 17 goals and 232 global targets, marking a sharp increase of more than 100% from the previous targets allocated under the MDG framework. Isolating poverty indicators alone, poverty reduction under the MDGs had three targets (mdgmonitor.org 2017) and in the SDG framework, this has increased to seven targets with 14 indicators (sdgtracker.org 2019).

Critical thinkers such as Jerven (2017) warn that this increasing need for data may cause its deterioration if the demand outstrips supply, hinting back to the re-introduction of a 'statistical tragedy' presented earlier by Devarajan (2013). Therefore, the fact that, 'data has a cost' (Jerven 2017:32) should not be taken lightly.

Currently there is no comprehensive costing and funding strategy in place that has the exact costs associated with providing adequate data for the SDGs that is inclusive of all countries. The consensus is that more funding is required, yet it remains undocumented and unclear as to the exact shortfalls per country and where the remainder of the much-needed funds will be coming from.

Notwithstanding this bleak picture, there have been global, regional and national efforts to improve statistical capacity in Sub Saharan Africa (SSA). Agendas such as 'A world that counts' by the United Nations opened talks on the need for a Data Revolution and are part of a broader set of initiatives that show commitments in improving data availability in developing countries. Kiregyera (2013) contrary to existing critics of African statistics argues that the continent should rather be understood as experiencing a 'statistical renaissance', presenting an optimistic view on the current trajectory of improvements in statistical capacity in African NSOs. The 2015-2030 period therefore presents a chance for the improved conduct of statistics in developing countries and in SSA more specifically. One of the arguments made in Chapter Four is that countries in SSA have a recently started conducting their own statistics after the end of colonialism. In South Africa, Stats SA was established in 1995 after the end of Apartheid.

Yet, the 1990-2015 period gave SSA countries key lessons in survey conduct and the data audits that were conducted with the end of the MDGs gave valuable lessons. Surely, the 2015-2030 period presents new possibilities to the conduct of better quality and comparable statistics, assuming of course that the much-needed funds for the conduct of such statistics will be available. Efforts by NSO's in African countries show improved collaboration efforts to achieve a common goal, and this is encouraging.

Increasing efforts in the improvement of statistical capacities are also demonstrated by the creation of international partnerships such as the Partnerships in Statistics for Development in the 21st Century (PARIS21). Other initiatives in place include the 2004 Marrakech Action Plan for Statistics and the Busan Action Plan for Statistics of 2011. Notable progress has also been observed in the commitments by NSOs and national governments as many[16] have adopted the UN Fundamental Principles of Official

[16] In 2017, 71 countries had signed the United Nations Fundamental Principles of Official Statistics, this number increased to 111 in 2018 (UN 2018).

Statistics in their national statistical legislation (UN 2018) displaying a wide-reaching consensus on the need for quality statistics that meet international standards.

These efforts have also created an improved policy environment needed in survey conduct largely absent in the pre-1990 period. Statistical conduct in the pre-1990 period was jeopardised by political instability, wars of liberation, (Kiregyera 2013) and a largely unfavourable environment that did not prioritize accurate and comprehensive data.

Statistical capacity of NSOs in developing countries in the post-1990 period has improved, and this has been accompanied by an expansion of data relevant for poverty measurement. Even though there is no agreed standard for measuring statistical capacity[17], the World Bank's Statistical Capacity Index supports the existence of significant improvements in the statistical capacity of most developing countries (Chen et al. 2013).

In light of these efforts and achievements, the re-emergence of poverty reduction efforts in the post-2015 development agenda has brought with it the realization that poverty data are less available when compared to data that measures other economic indicators (Alkire 2014; Serajuddin et al. 2015). Serajuddin et al. (2015:7) partly pins this to the assumption that poverty is less 'volatile' when compared to other economic phenomenon such as unemployment. The result is that there is more[18] frequent data collected for the stock exchange (Alkire 2014) and employment (Alkire 2014; Serajuddin et al. 2015) whilst poverty related data are collected every three to ten years (Alkire 2014:2). The assumption however that poverty is less 'volatile' is challenged by findings from Paraguay, India and Argentina where poverty shows significant volatility in short spaces of time (Serajuddin et al. 2015). Therefore, there is need for more frequent poverty data. Annual updates on poverty trends in a country and more frequent panel data would be ideal in informing poverty reduction policy better.

[17] Different organizations have developed their own Statistical Capacity Indices and such organizations include the Food and Agricultural Organization (FAO), The United Nations Economic Commission for Africa (UNECA), PARIS21 and the United States of America Census Bureau (Cameron et al. 2019).

[18] Data on the Gross National Income (GNI) are collected annually, whilst labour force surveys are quarterly and the daily updates on the stock exchange (Alkire 2014:1).

Regardless of the fact that poverty data are less available, poverty reduction continues to be a priority focus. To highlight the urgency given to alleviate this problem, the World Bank's post-2015 agenda seeks 'to end extreme poverty' (Serajuddin et al. 2015:2), whilst the Africa Union Commission similarly prioritises poverty reduction in Africa as one of the main objectives of Agenda 2063 (Africa Union Commission (AUC) 2015). Poverty remains Goal number one under the SDG framework.

This thesis prioritised the growing demand for more accurate, timely and comparable data for poverty measurement in developing countries and used South Africa as a case study. It is the opinion in this thesis that poverty reduction commences with accurate poverty measurement. In fact, the post-2015 development trajectory needs more comparable, frequent, disaggregated and accurate data more than in recent periods. This is partly because data collection methods in developing countries have been improving, in addition to the existence and significant improvement of international, regional and national frameworks that guide data collection and questionnaire design. The environment for generating better statistics looks more promising than it did in the pre-1990 period.

Despite these improvements, it is worth mentioning that household survey conduct is an expensive engagement and it still remains less clear how this recent increase in demand for data will be funded. Jerven (2014c:2) estimated that measuring the MDGs cost an estimated USD$27 billion or USD$1,5 per target. Given the fact that this amount only included the collection of data on 8 MDGs and their accompanying 18 targets, the SDGs with 16 goals and 169 targets is estimated to cost $254 billion.

Due to such high costs of data in the post-2015 period, data manipulation techniques become an alternative way of measuring poverty in a cost-effective way. This thesis approached the measurement of poverty dynamics in South Africa from this angle. There is need for more information to supplement existing knowledge on poverty dynamics in the country, but data availability restrictions limit our ability to use various data sources to pursue this objective.

7.2 Summary of findings

The objectives of this thesis were fourfold; the first sought to review how poverty has been measured in South Africa. Under this objective, this thesis analysed the different

household survey data that have been used to measure poverty trends and dynamics as well as the poverty trajectory they portray. The second objective of this thesis examined factors and issues surrounding the data sources in South African surveys related to this study and incorporated issues pertaining to frequency, volatility, comparability and reliability of poverty data sources. Thirdly, this thesis investigated the robustness of poverty trends and characteristics in South Africa using pseudo panel data. Lastly, the thesis examined the linkages that exists between poverty, race, geography and gender in South Africa using pseudo panel data

This thesis used a mixed method research design, the qualitative enquiry pursued in Chapter Four used key informants from Stats SA and explored how the organization has measured poverty in the post-1995 period. The objective was to give a more nuanced account on the developments engaged to date on poverty measurement in the country to understand the past, present and future engagements in poverty measurement in South Africa. Purposive sampling was used, key informants from Stats SA were interviewed using a semi-structured interview guide.

On the other hand, the quantitative part of this thesis analysed expenditure patterns and poverty estimates using a cohort-based pseudo panel. This analysis was further extended to include the differences in the risk to poverty for birth cohorts in rural and urban areas as well as between the racial groups in South Africa. Prior to this, the study reviewed data gaps in poverty analysis and SSI and pseudo panel approaches were discussed as possible solutions to the data inadequacy problem at the centre of this thesis. The following sections summarises the key discussions and findings made in this study and concludes with a policy and recommendations section.

7.2.1 Expansion of survey data for poverty measurement

The consensus in the post-2000 period in developing countries is that there has been an expansion in the availability of household surveys (Alkire 2014; Beegle et al. 2016, Chan and Ravallion 2010; Demombynes and Sandefur 2014; Gibson 2016; Jolliffe et al. 2014; Serajuddin et al. 2015). Poverty related data has also become available in more countries in the developing world, (Serajuddin et al. 2015) and improvements have been observed not only in the frequency of poverty data but also in terms of credibility of the data sources (Jolliffe et al. 2014).

Panel surveys have also increased in the same period (Hsiao 2007) but with a lesser frequency when compared to cross sectional surveys. Of note as well is the fact that nationally representative panel surveys are fewer with a majority of panel data collected at sub-national level. Existing nationally representative surveys also cover shorter time periods and are available for few countries (Dang and Dabalen 2017). Only seven countries in Sub Saharan Africa had panel data in 2012 (Dang and Dabalen 2017), indicating the gravity of the paucity of panel data in this region. Despite this shortage, household survey conduct in the post 1990 period improved in developing countries.

In South Africa, similar gains in survey conduct have also been observed. This has also been accompanied by an increase in studies that have analysed poverty in South Africa. In the post 1994 period household surveys such as the IEC, LCS, GHS, CS, NIDS and KIDS have contributed immensely to what we know about poverty in the country. Nevertheless, the seemingly 'abundant' source of household data has been associated with incomparability issues, as a result of the use of different methodologies and poverty lines.

The continued expansion of household level data for the analysis of poverty in the post 2020 period is not guaranteed. Findings from the qualitative research conducted in this thesis highlighted the lack of adequate funding as one of the biggest challenges faced by Stats SA. This apart from affecting the ability for new and current household surveys from being conducted also places a risk on Stats Sa's ability to attract and retain the required skilled staff to administer these surveys. In fact, in a recent report from Stats SA (2020), the agency had its budget stripped of R160 millions in 2015 and since then no new staff members have been hired, neither has there been any promotions. The future conduct of statistics in the country therefore hangs in the balance.

The existing vacancy rate of 20% at Stats SA is not sustainable and is inconsistent with the current data needs for the analysis of poverty in line with the 2030 poverty reduction agendas. There is need for more funding to be allocated to Stats SA to avoid data gaps in the 2015-2030 poverty measurement efforts. The reduction of sample sizes as a result of budget cuts apart from resulting in wider error ranges ultimately affects the quality of statistics at Stats SA and removes the credibility of the agency.

7.2.2 Paucity of data for poverty measurement

Despite the efforts observed in survey conduct in developing countries, data gaps remain in poverty measurement and these gaps are more pronounced for panel data. This has been partly attributed to the high costs associated with conducting panel data as well as the fact that they are more complicated to conduct when compared to cross sectional surveys. Baulch (2011) illustrated the availability of panel survey data in developing countries and this study attempted to update as comprehensively as possible, this list of panel surveys. Although it is evident that panel survey availability has improved in developing countries, the production of such data has not matched the existing demand in poverty analysis.

The post-1990 period in developing countries has seen a significant advancement in data availability, but this remains largely inadequate. For instance, Serajuddin et al. (2015) found that only five countries from the 48 countries in sub-Saharan Africa were considered to have satisfactory data whilst 70% of the countries in this region were classified as having inadequate data.

In South Africa, this study has argued that the country on the surface appears to be 'data rich' as far as the production of household survey data are concerned. The country has a number of household surveys that have been used to measure poverty trends over the years. Close scrutiny of the existing data however shows a general lack of comparability within similar surveys (IESs), inadequate consumption/income modules (CSs) or weak income/consumption modules for the analysis of poverty trends (GHSs). The narrative worsens when it is extended to panel surveys, with only two panels available to date; one conducted only in KwaZulu Natal province, the KwaZulu Natal Income Dynamic Study (KIDS) and the National Income Dynamic Study (NIDS), which is nationally representative but has small sample problems. The data gaps identified in this study suggest that our knowledge for poverty dynamics are limited. Fewer national panel surveys have been conducted and these are not available for extensive periods of time. Therefore, in South Africa, orthodox panel types of data are inadequate and alternative data types such as pseudo panels become attractive alternative in the analysis of poverty dynamics.

7.2.3 Role of NSOs in data provision for poverty measurement

The burden of providing data for the measurement of the progress on various development targets largely rests with their respective NSOs. In South Africa Stats SA plays this role. Whilst data gaps in the analysis of poverty dynamics are acknowledged in poverty research, less qualitative research exists on Stats SA's mandate in the post-1994 South Africa. Using key informants, this study explored the history of Stats SA and its efforts in the measurement of poverty after its inception in 1995.

Stats SA is a government organization introduced in 1995 after the previous Apartheid regime-biased Central Statistical Office was abolished with the end of apartheid. Prior to its introduction, statistical analyses in South Africa apart from having a strong racial bias were largely incomprehensive and focused on a White minority. These discriminatory tendencies tended to dominate survey conduct in the country in the pre-1994 period and legislative instruments such as the Registration Act of 1950 further worsened an already biased system of conducting statistics.

Therefore, at the dawn of democracy, the foundation for statistical conduct had to be repositioned to meet the new order of a newly elected government. Between 1995 and 2019, Stats SA has attempted not only to conduct comprehensive statistics, but also to improve data quality, frequency, comparability, availability and access. To achieve these objectives, Stats SA through the SASQAF has attempted to adopt a quality-oriented stance in the production and dissemination of official statistics. The organization has also attempted to incorporate international standards in survey conduct to improve comparisons with other countries.

Despite these notable changes and sound principles on paper, implementation is not always straightforward. Stats SA has managed to increase the conduct of surveys in the country, at the same time, these surveys have been found to be problematic in some cases when measuring poverty. The consensus from the interviews conducted in this study concluded that there is need for more analysis on poverty with existing survey data in South Africa to explore new pathways out of poverty. In line with this suggestion, this study by the use of a pseudo panel is a development in this direction.

7.2.4 Alternative data for poverty measurement

This study has discussed the desirability of alternative data for the measurement of poverty. On a cross sectional level, poverty analysis is often hampered by the

unavailability of comprehensive income and consumption/ expenditure surveys. Data imputation techniques that use asset indices, poverty mapping and SSI regression methods discussed in this thesis have proved to be cost effective and useful in the generation of unorthodox data sources for the analysis of poverty trends.

On a longitudinal dimension, pseudo panel methods have been applied with encouraging results. Deaton (1985); Deaton and Paxson (1994); Banks et al. (2001); and Pancavel (2007) have conducted cohort based pseudo panel, whilst Bourguignon et al. (2004); Guell and Hu (2006); Dang et al. (2014) are some of the studies that have generated another type of pseudo panels that enabled the analysis of poverty dynamics within households or at individual level. The use of pseudo panels has enabled poverty dynamics to be analysed more extensively in the absence of genuine panel data.

The application of pseudo panel methods in poverty also comes with limitations and strengths and following appraisals of pseudo panel methods in Urzanqui (2017) and Herault and Jenkins (2019), they need to be used with reservations. Despite this, pseudo panel methods have allowed the analysis of poverty dynamics where existing 'traditional forms' could not be used for such analysis.

7.2.5 Birth cohort pseudo panel methodology

This study used a birth-year cohort-based pseudo panel under a set of given assumptions. The pseudo panel is derived from the methods of Deaton (1985) generated birth cohorts for household heads born between the years 1924 and 1988. In total 13 birth cohorts were created with a five-year band reference and these are 'tracked' in three time periods. Part of the assumptions applied in this methodology is that household heads within the same cohort are identical in terms of their mean expenditure per capita and cohort means are generalized for the whole cohort group. This study raised a concern that the intra-cohort homogeneity assumption might be problematic given South Africa's HIV/AIDS prevalence rates and patterns which affects different genders and age groups differently.

The second assumption applied cohort means as cohort observations and once again this assumption assumes there is intra group homogeneity. Thirdly, the mean of one birth cohort is assumed to only apply to members within that cohort and is different from the next birth cohorts. And lastly, each household head is assumed to participate in one birth cohort. It is possible however that the same household was interviewed

more than once between 2008 and 2014.

7.2.6 Pseudo panel findings

The analysis conducted for the birth cohorts was on three levels. The first section analysed the differences in mean per capita expenditure for the birth cohorts between 2008 and 2014. The analysis discussed the differences in expenditure patterns of birth cohorts as they got older and within FHH, MHH, rural, urban and within households headed by household heads with different races. Secondly in an attempt to separate age, time and cohort effects, a multiple regression was conducted. In the last section poverty analysis is conducted for the birth cohorts.

This study found that residing in rural areas, as well as in FHH translates to lower mean per capita expenditure and higher susceptibility to poverty. The same is true for birth cohorts headed by African and Coloured household heads. Younger birth cohorts also tend to have lower mean expenditure per capita when compared to older birth cohorts. But older birth cohorts tend to be more vulnerable to poverty than younger birth cohorts.

To partly explain these differentials in mean per capita expenditure and poverty incidence, the study found that vulnerability to poverty is worsened by overreliance on social grants as the main source of income by household heads. For example, MHH and birth cohorts residing in urban areas are more likely to receive their income from salaries/wages whilst birth cohorts with female headship and in rural areas cite social grants as their main source of income. This study acknowledges that expenditure can also be driven by other variables beyond income to other sources such as borrowing or savings. However, incomprehensiveness of savings variables in the LCS and their total absence from the IES of 2010 restricted this analysis.

Moreover, the analysis could not incorporate the use of employment related variables due to the presence of incomparable employment variables in the LCSs and the IES of 2010. Such data would have enabled the analysis of poverty and mean per capita expenditure to be analysed in terms of the types of employment for birth cohorts. Despite this, the presence of salaries/wages suggests that some birth cohorts are deriving their income from employment related sources. Earlier studies that have analysed poverty dynamics in South Africa also found that employment has a strong association with households exiting poverty Woolard and Klasen (2005) came to a

similar conclusion. Similarly, this pseudo panel showed that birth cohorts receiving income/wages are less likely to be vulnerable to poverty.

The findings from a regression conducted in Chapter six confirmed that expenditures are hump shaped over the lifecycle in a multiple regression conducted derived from a similar analysis in Agiuar and Hurst (2013). After factoring in arguments from Reintz et al. (1983), this study tried to isolate age, time and period effects which are confounded when using longitudinal designs.

7.3 Significance and contribution of research

The use of pseudo panels in poverty research is becoming more popular in developing countries due to the lack of adequate panel surveys to analyse dynamic poverty. Pseudo panels have been used in the analysis of poverty dynamics in Cruces et al. (2013); Cuesta et al. (2007); Dang et al. (2011); Dang and Lanchovichina (2016); Dang and Lanjouw (2013); Dang and Lanjouw (2016); Dang et al. (2014); Dang et al. (2017). This growing body of research has included some African countries but has not been comprehensive. The most comprehensive application of pseudo panels with African countries was applied in Dang and Lanjouw (2017) and 21 African countries were included. South Africa was not included. This research using a cohort-based pseudo panel is a contribution of the analysis of poverty dynamics in South Africa.

This study becomes the first to use the LCSs (2008 and 2014) and IES of 2010 to generate a cohort-based pseudo panel for the analysis of poverty in South Africa. This study by using a pseudo panel also generated a new data source for the analysis of poverty dynamics in South Africa. Existing knowledge on poverty dynamics in the country has largely relied on the use of NIDS and KIDS data. This study provides an alternative way of understanding poverty dynamics in South Africa. Existing knowledge of poverty dynamics has relied on the use of analysis at household level, this study allowed the use of birth cohorts.

Discussions conducted with officials from Stats SA in Chapter Four also showed that the agency has a funding problem. The LCS for example in 2008 and 2014 did not have its own budget allocation and instead, it relied on funds allocated to other household surveys such as the IES. Compounding this problem is the fact that the IES can only be conducted every five years, and not sooner.

Recently, evidence has also emerged that in 2015, Stats SA had R160 million stripped from its budget and since then no new posts or promotions have been done within the organization- leaving the organization with a 20% vacancy rate (Stats SA 2020).

Translating these budget constraints on the possibility of new panel data being conducted in South Africa is ambitious and it is within this line of thought that pseudo panels using existing data become significant. Alternative data sources proposed in this thesis are cost effective and therefore valuable in poverty measurement.

7.4 Relevance of study to policy

The study by using a new source of 'panel' data tells a nuanced narrative on the causes of poverty traps for birth cohorts between 2008 and 2014. 'Tracking' birth cohorts in three time periods allowed the identification of birth cohorts who are the most vulnerable to poverty as well as an understanding of the description of such birth cohorts such as household headship, geographic area, age and race. There are birth cohorts that were poor in one time period, while for some birth cohorts their experiences of poverty were more pronounced, lasting for two time periods and three in the most extreme cases. In the absence of panel data, policy effectiveness may be reduced due to a lack of understanding of poverty dynamics (Scott and Marrioti 2014).

The use of pseudo panels in poverty analysis is therefore relevant for poverty reduction policy. The qualitative findings from this study established that Stats SA currently has no plan to conduct a panel survey for poverty analysis. In fact, their efforts are geared towards supporting the conduct of NIDS at the University of Cape Town[19]. A generation of pseudo panels therefore supports existing efforts in monitoring poverty dynamics in the country.

Poverty reduction policy in South Africa could also benefit from the estimation of additional pseudo panels. To fully understand the poverty dynamics of South Africa since 1994 we would need a long running panel with good quality data and to date such data does not exist. Pseudo panels therefore provide possibilities of additional knowledge on the country's poverty dynamics where there are gaps. In light with the promotion of the use of pseudo panels in poverty panels, there is need for statistical training on modelling techniques in the country to promote the generation of more

[19] This information was provided during interviews in April 2017 at Stats SA in Pretoria.

pseudo panels to explore new avenues of poverty in the country.

7.4.1 Relevance to poverty reduction and social protection in South Africa

South African poverty reduction policies in the post-1994 period have invested strongly in the provision a wide-reaching social protection system. Booysen (2004) ranked the country's social protection system as an equivalent to those found in developed countries. The expansion of social grants recipients from 2.4 million to 12.4 million between 1994 and 2008 (Armstrong and Burger 2009) to close to 17 million in 2017 (World Bank and Stats SA 2018), shows that social protection has expanded widely. This high spending on social protection has also paid off and has been associated with the reduction of poverty in the country especially in the post-2000 period (Booysen 2004; Schiel 2014; World Bank and Stats SA2018).

However, this study found that households whose main source of income is social grants are more likely to experience poverty. This could be tied to the fact that the majority of households receiving social grants are recipients of the Child Support Grant and since this grant has the least amount, households that rely entirely on it are at risk of experiencing poverty. Booysen (2004) also argued that the Child Support Grant has little impact on poverty reduction because it is a smaller amount.

7.4.2 Poverty and structural forms of employment

This study found that the lack of labour sources of income for birth cohorts makes them more vulnerable to poverty. Even though the World Bank and Stats SA (2018) found that having a job is not always associated with escaping poverty, employment can create pathways out of poverty. This shows that one of the ways in which chronic and structural forms of poverty can be eliminated in the country is by creating long term employment opportunities. The World Bank and Stats SA (2018) concluded that South Africa experiences high levels of unemployment when compared to other countries in a similar income bracket. Youth unemployment is higher (38.6%) than the country's unemployment rates (27.7%) (World Bank and Stats SA 2018:iv). Addressing the structural nature of unemployment in South Africa will arguably go a long way in alleviating chronic and structural poverty in South Africa.

7.5 Final remarks and recommendations

There is need for more analysis on poverty trends and dynamics using existing data in South Africa. Current knowledge on poverty dynamics in South Africa has been limited to the use of the KIDS and NIDS panel data. Insightful research has emanated from these two panels. However, there is need for more analysis on poverty dynamics in the country and the benefits associated with the use of pseudo panels cannot go unacknowledged. For instance, pseudo panels since they require cross sectional household surveys in at least two time periods can be used to generate 'panel' data in periods where genuine panel data is absent, or for extended periods of time that are not covered by genuine panel data sources. In South Africa, this could enable pseudo panels to be generated in the immediate period after the end of apartheid to date since NIDS only started being available in 2008. The IES has been produced in 1995-2011, the GHS 2002 to date, and the OHS 1993-1999. This study prioritised the LCSs since they were specifically designed to measure poverty, but more research needs to be done on how the above-mentioned surveys can be utilized to understand poverty dynamics in the country.

Therefore, more pseudo panels for the analysis of poverty are required in South Africa to complement our knowledge of poverty dynamics in the country and establish new pathways of ending chronic and structural forms of poverty. Alternative data sources can fill in part of the data gaps in poverty analysis. Moreover, in light of existing data gaps identified during the MDG era, data consolidation on existing household surveys may be more effective than introducing new surveys. There is need for consolidation of econometric skills used in the creation of pseudo panel data methods with national NSOs in developing countries. NSOs should embrace new technologies due to their cost effectiveness and their ability to minimise human error.

7.6 Conclusion

A majority of African NSOs are quite recent, with their national governments having received their independence and democracy a few decades ago (24 years in the case of South Africa). An analysis of South Africa's history of statistical conduct showed the inconsistencies that existed in the conduct of population census, mostly influenced by Apartheid laws that prioritized a few. As a result, at the dawn of democracy the statistical order within the statistical office was in disarray and the post-1994 mandate

of Stats SA has been to produce ‘better’ data. Therefore, NSOs in some African countries, including South Africa, are still in their infancy and it will take a longer time frame for effective implementation of statistical frameworks that guide the collection, analysis and dissemination of official statistics.

Stats SA in this regard has started using the CAPI to collect data, a move that is expected to minimise human error in the collection and inputting of data. Survey conduct for poverty analysis in the post-1994 period in South Africa notably improved after the 2000 period and to date, a number of surveys have been used in the country to analyse poverty. However, the issues associated with such data remain transparent and are an indication that improvements are expected in the future especially on the comparability of poverty related survey data.

References

Adato, M. Carter, M. R. and May, J. (2006) Exploring poverty traps and social exclusion in South Africa using qualitative and quantitative data'. Journal of Development Studies 42(2): 226-247.

Adato, M. Lund, F. and Mhlongo, P. (2007) Methodological innovations in research on the dynamics of poverty: A longitudinal study in KwaZulu Natal, South Africa. World Development, 35(2):247-263

Addison, T. Hulme, D. and Kanbur, R. (2009) Poverty Dynamics: Interdisciplinary Perspectives. 10.1093/acprof/9780199557547.001.0001

Aguiar, M, and Hurst, E. (2013) Deconstructing life cycle expenditure, Journal of Political Economy, University of Chicago Press 121(3):437-492

Agüero, J. Carter, M. and May, J. (2007) Poverty and Inequality in the First Decade of Democracy: Evidence from KwaZulu-Natal. Journal of African Economies, 16(5):782–812.

Alderman, H. Babita, M. Demombynes, G. Makhatha, N.and Ozler, B. (2001) How low can you go? Combining census and survey data for mapping in South Africa

Aliber, M (2001) Study of the incidence and nature of chronic poverty and development policy in South Africa: An overview. Chronic Poverty and Development Working Paper No 1.

Aliber, M (2003) Chronic Poverty in South Africa: Incidence, Causes and Policies, 31:473-490

Alkire, S. (2014) Towards frequent and accurate poverty data. Oxford Poverty and Human Development Initiative (OPHI)

Alkire, S. & Foster, J. (2007, revised in 2008). Counting and multidimensional poverty measurement. OPHI Working Paper 7, University of Oxford.

Allen, T. (1992) Economic development and the feminisation of poverty. In Issues in contemporary economics: Vol. 4. Women's work in the world economy London: Macmillan/IEA.

Angrist, J. and Krueger, A. (1992) The effect of age of school entry on educational attainment: An application of instrumental variables with moments from two samples. Journal of American Statistical Association 87(418): 328-336

Antman, F. and McKenzie D.J. (2007) Earnings mobility and measurement error: A pseudo panel approach

Araar, A. and Duclos, J.Y. (2013) DASP: Distributive Analysis Stata Package. Laval University, World Bank, PEP and CIRPÉE.

Ardington, E. M. (1995) Return to Nkandla: The third survey in a longitudinal study of a rural community in KwaZulu-Natal. Centre for Social and Development Studies (CSDS) Working Paper No. 7, Durban: University of Natal

Arellano, M. and Meghir, C. (1992) Female labor supply and on the job search: An empirical model estimated using complementary data sets. Review of Economic Studies 59(3): 537-559.

Argent, J. Finn, A. Leibbrandt, M. and Woolard, I. (2009) Poverty: Analysis of the NIDS Wave 1 Dataset. NIDS Discussion Document No. 13, Southern Africa Labour and Development Research Unit, University of Cape Town.

Appleton, S. (1996) Problems of measuring changes in poverty over time: the case of Uganda 1989-1992. Institute of Development Studies Bulletin, 27(1): 43-55

Atamanov, A. Jolliffe, D. Lakner C.and Espen B.P. (2018) Purchasing Power Parities used in Global Poverty Measurement. World Bank Group Global Poverty Monitoring Technical Note, no. 5. September 2018

Attanasio, O.P. and Weber, G. (1993) Consumption Growth, the Interest Rate and Aggregation. Rev. Econ. Studies 60: 631-49.

Baigrie, N. and Eyal K. (2013) An evaluation of the determinants and implications of panel attrition in the National Income Dynamics Survey (2008-2010)

Bamberger, M. (2009) 'Review on Statistical issues in measuring poverty from non-survey sources', in Kamanou, G. Morduch, J. David, I.P. and Gibson, J. (2005) Handbook on poverty statistics: Concepts, methods and policy use. United Nations Statistics Division pp.

Bane, M.J. and Ellwood, D.T. (1985) Slipping in and out of poverty: The dynamics of spells National Bureau for Economic Research (NBER) Working Paper Series Working Paper number 1109

Banks, J. Blundell, R. and Brugiavini. A. (2001) Risk Pooling, Precautionary Saving and Consumption Growth. Review of Economic Studies, 68(4): 757-779.

Barrett, C. (2006) Rural poverty dynamics: development policy implications. In Reshaping Agriculture's Contributions to Society, ed. Colman, D. and Vink, N. Oxford: Blackwell.

Baulch, B. (1996) Neglected trade-offs in poverty measurement. Institute of Development Studies

Baulch (2011) (ed.) Why Poverty Persists: Poverty Dynamics in Asia and Africa, Edward Elgar Publishing, Cheltenham, England, pp. 187-218

Baulch, B. (2013) Understanding poverty dynamics and economic mobility. In Shepherd, A. and Brunt, J. (eds) Chronic poverty. Rethinking International Development Series. Palgrave Macmillan, London

Baulch, B., Chant, L. and Robinson, S. (2008) What happens between the waves? Estimating inter-wave dynamics from limited survey data with application to poverty transitions in South Africa and Vietnam. GTAP Conference Paper, Purdue University. https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=2754

Baulch, B and Davis, P. (2008) Poverty dynamics and livelihood trajectories in rural Bangladesh

Baulch B. and Hoddinott, J. (2000) Economic mobility and poverty dynamics in developing countries, The Journal of Development Studies, 36(6):1-24, DOI:10.1080/00220380008422652

Baulch, B. and Masset, E. (2003) Do Monetary and Nonmonetary Indicators Tell the Same Story about about chronic poverty? A study of Vietnam in the 1990s." World Development 31(3): 441-453.

Baulch. B. and McCulloch N. (2000) Being poor and becoming poor: Poverty status and poverty transitions in rural Pakistan. Institute for Development Studies. Working Paper 79

Baulch, B. and Quisimbing, A. (2009) Assets and poverty traps in rural Bangladesh. Centre for Chronic Poverty Research Centre (CPRC) Working Paper Number 143

Baulch, B. and Vu, D. (2011) Poverty dynamics in Vietnam, 2002 to 2006 in Bob Baulch (ed.) Why Poverty Persists: Poverty Dynamics in Asia and Africa, Edward Elgar, Cheltenham, UK, pp. 219-254.

Baulch, B. (n.d.) Moving out of poverty: Growth and freedom from the bottom up. Paper presented for the World Bank.

Baxter, P. and Jack, S. (2008) Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. The Qualitative Report, 13(4):544-559.

Beegle, K. Christiansen, L. Dabalen, A. and Gaddis, I. (2016) Poverty in a rising Africa. Washington, D.C.: World Bank

Benson, T. Epprecht, and Minot, N. (2007) Mapping where the poor live. 2020 Focus Brief on the World's Poor and Hungry People. Washington, D.C: IFPRI

Behrman, J. (2006) Using Micro Data to Understand Better the Intergenerational Transmission of Poverty in Low Income Developing Countries: Methodological Note.

Betti, G. and Cheli B (1998) Panel and pseudo panel techniques for living conditions analysis following a fuzzy approach. Statistica Applicata 10(4): 557-570

Bevan, P. (2003) Extending understanding of chronic poverty dynamics: Towards a post-disciplinary Approach. WeD Working Paper Series. University of Bath

Bhorat, H. and Kanbur, R. (2005) Poverty and Well-being in Post-Apartheid South Africa: An overview of data, outcomes and policy. Development Policy Research Unit, Working Paper.

Bhorat, H. van der Westhuizen,C. and Cassim, A. (2009) Access to household services and assets: Analysis using the NIDS Wave 1 dataset. NIDS Discussion Paper no. 4, no. 4 (2009)

Bhorat, H. Oosthuizen, M. and van der Westhuizen, C. (2011) Estimating a poverty line: An application to free basic municipal services in South Africa. Development Policy Research Unit (DPRU). University of Cape Town

Bhorat, H. and Westhuizen C. (2012) Poverty, inequality and the nature of economic growth in South Africa. DPRU Working Paper 12/151

Bigsten, A. and Shimles A. (2004) Dynamics of poverty in Ethiopia WIDER Research Paper No 2004/39

Bird, K. Higgins, K. and Harris, D. (2010a) Spatial Poverty Traps: An Overview. ODI/CPRC Working Paper Series (ODI WP321, CPRC WP161). London: ODI and Manchester: CPRC, University of Manchester.

Bird, K. McKay, A. and Shinyekwa, I. (2010b) Isolation and poverty: the relationship between spatially differentiated access to goods and services and poverty. ODI/CPRC Working Paper Series (ODI WP322, CPRC WP162). London: ODI and Manchester: CPRC, University of Manchester.

Bird, K. and Shepherd, A. (2003) Livelihoods and Chronic Poverty in Semi-Arid Zimbabwe. World Development. 31:591-610.

Blundell, R. M. Browning and C. Meghir (1994) Consumer Demand and the Life-Cycle Allocation of Household Expenditures, Review of Economic Studies, 61: 57–80.

Bogdan, R. C. and Biklen, S. K. (2007) Qualitative research for education. An introduction to theories and methods, (5th e.d.) Boston. MA: P Education Inc.

Bourguignon, F. Goh, C. and Kim D. (2004) Estimating individual vulnerability to poverty using pseudo-panel data. World Bank Policy Research Working Paper 3375

Bourguignon, F. and Moreno H. (2015) On the construction of synthetic panels. Paper presented at the North East Universities Development Conference, Brown University, Providence RI.

Branson, W. (1997) Macroeconomic Theory and Policy. Pearson, Third Edition

Browning, M. Deaton, A. and Irish, M. (1985) A profitable approach to labour supply and commodity demands over a lifecycle. Econometrica 53:503-545

Budlender, D. (1985) A critique of poverty datum lines. SALDRU Working Paper 63, University of Cape Town.

Budlender, D. Leibbrandt, M. and Woolard, I. (2015) South African Poverty lines: A review and two new money-metric thresholds. SALDRU Working Paper Number 151 UCT

Budlender, D. and Lund, F. (2011) South Africa: A Legacy of Family Disruption. Development and change. 42. 925-46. 10.1111/j.1467-7660.2011.01715.

Bureau for Research and Economic Analysis of Development (BREAD) (2019) http://ibread.org/bread/data

Buvinic, M. (1990) The vulnerability of women headed households: policy questions and options for Latin America and the Caribbean.Paper presented at the Economic Commission for Latin America and the Caribbean Meeting on 'Vulnerable Women', Vienna, 26-30 November

Buvinic, M. and Gupta, G. R. (1993) Responding to insecurity in the 1990s: targeting woman-headed households and woman-maintained families in developing countries. Paper presented at the International Workshop 'Insecurity in the 1990s: Gender and Social Policy in an International Perspective', London School of Economics and European Association of Development Institutes, London, 5-6 April

Buvinic, M. and Gupta, G.R. (1997) Female headed households and female maintained families: Are they worth targeting to reduce poverty in developing countries. Economic Development and Cultural Change 45(2):259-280

Buvinic, M. Nadia H Youssef, Barbara Von Elm (1978) Women-headed households: The ignored factor in development planning

Calvo, C. and Dercon, S. (2009) Chronic Poverty and All That: The Measurement of Poverty over Time. In T. Addison, D. Hulme and R. Kanbur (Eds.), Poverty Dynamics: Interdisciplinary, Oxford University Press.

Carter M.R. and Barrett, C. (2006) The economics of poverty traps an asset based approach. Journal of Development Studies 42 (2): 178-199

Carter, M.R. and May, J. (1999) Poverty, Livelihood and Class in Rural South Africa. World Development.27(1): l-20

Carter, M.R. and May, J. (2001) One kind of freedom: Poverty dynamics in post-Apartheid South Africa. World Development 20(10): 1987-2006

Chant, Sylvia (1997a) Women-headed Households: Diversity and Dynamics in the Developing World Houndmills, Basingstoke: Macmillan.

Chant, Sylvia (1997b) Women-headed Households: Poorest of the Poor? Perspectives from Mexico, Costa Rica and the Philippines, IDS Bulletin, 28:(3): 26-48

Chant, Sylvia (1999) Women-headed Households: Global Orthodoxies and Grassroots Realities, in Haleh Afshar and Stephanie Barrientos (eds) Women, Globalisation and Fragmentation in the Developing World Houndmills, Basingstoke: Macmillan 91-130.

Chant, Sylvia (2003a) Female Household Headship and the Feminisation of Poverty: Facts, Fictions and Forward Strategies (London: London School of Economics, Gender Institute, New Working Paper Series, Issue 9).

Chant, Sylvia (2003b) New Contributions to the Analysis of Poverty: Methodological and Conceptual Challenges to Understanding Poverty from a Gender Perspective.(Santiago de Chile: Comisión Económica para América Latina (CEPAL), Unidad Mujer y Desarrollo, Serie 47 (www.cepal.org).

Chant, S. (2006) Re-visiting the 'feminisation of poverty' and the United Nationas Development Programme (UNDP) gender indices: what case for a gendered poverty index? The London School of Economics and Political Science.

Chaudhuri, S. and Ravallion M. (1994) How well do static indicators identify the chronically poor? Journal of Public Economics Volume 53, Issue 3: 367-394

Chen, S. and Ravallion, M. (2010) The developing world is poorer than we thought, but no less successful in the fight against poverty. Quarterly Journal of Economics

Christopher, A.J. (2011) The Union of South Africa censuses 1911–1960: An incomplete record. Historia 56(20):1–18

Chronic Centre Poverty Research Centre (CPRC) (2005) The Chronic Poverty Report 2004–05. Institute for Development Policy and Management email: info@chronicpoverty.org University of Manchester. United Kingdom

Chronic Centre Poverty Research Centre (CPRC) (2009) Escaping poverty traps. The Chronic Poverty Report 2008-9

Chronic Centre Poverty Research Centre (CPRC) (2011a) Tackling chronic poverty: The policy implications of research on chronic poverty and poverty dynamics.

Chronic Centre Poverty Research Centre (CPRC) (2011b) The value of panel data in chronic poverty analysis. Policy Brief March 2011, No 23

Cichello, P. and Newhouse, D. and Fields, G. and Freije, S. and Menendez, M. (2003) For Richer or for Poorer? Evidence from Indonesia, South Africa, Spain, and Venezuela. Journal of Economic Inequality. 1. 67-99.

Clarke, D. and Hulme D. (2005) Towards A Unified Framework for Understanding the Depth, Breadth and Duration of Poverty. Global Poverty Research Group and Institute for Development Policy and Management University of Manchester, UK.

Clark, A. (2006) Anonymizing research data. Economic and Social Research Centre ESRC National Centre for Research Methods NCRM Working Paper Series 7/06

Coe, R.D. (1978) Dependency and Poverty in the Short and Long Run." in Greg J. Duncan and James N. Morgan, editors, 5000 American Families: Patterns of Economic Progress, Vol. Vi. Ann Arbor, Michigan: Institute for Social Research

Coe, R.D. Duncan, G.J. and Hill, M.S. (1982) Dynamic aspects of poverty and Welfare Use in the United States." Paper presented at Conference on Problems of Poverty, Clark University

Cohen, D. (2006) Qualitative Research Guidelines Project. Semi-structured interviews. New Jersey: Robert Wood Johnson Foundation

Cohen, L. Manion, L. and Morrison, K. (2011) Research Methods in Education (7 ed.). New York: Routledge

Collado, D.M. (1997) Estimating dynamic models from time series of independent cross sections. Journal of Econometrics 82: 37-62

Corcoran, M. (1995) Rags to Rags: Poverty and Mobility in the United States. Annual Review of Sociology, 21:237-267.

Corsi, D.J. Neuman, M. Finlay, J.E. Subramanian, SV. (2012) Demographic and health surveys: a profile, International Journal of Epidemiology, Volume 41, Issue 6, December 2012, Pages 1602–1613, https://doi.org/10.1093/ije/dys184

Coudouel, A. Hentschel, J and Wodon, Q. (2002) Poverty measurement and analysis. World Bank

Creswell, J.W. (2009a) Mapping the Field of Mixed Methods Research Journal of Mixed Methods Research 3(2): 95-108

Creswell, J.W. (2009b) Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage, Los Angeles

Creswell, J. W. (2013) Qualitative Inquiry and Research Design: Choosing among Five Approaches (3rd ed.). Thousand Oaks, CA: SAGE.

Creswell, J.W. (2014) Research design: Qualitative, quantitative and mixed methods approaches. Sage, Los Angeles

Creswell, J.W. and Plano Clark V.L (2011) Designing and conducting mixed method research. 2ND Sage: Thousand Oaks.

Cronin V. and Gunthrie, P. (2011) Alternative approaches to slum upgrading in Kibera, Nairobi. Proceedings of the Institution of Civil Engineers Urban Design and Planning 164(2):129–139

Crow, G. and Pope, C. (2008) The future of the research relationship, Introduction to special issue of Sociology 42 (5);813-19.

Cruces, G. Fields, G. and Viollaz, M. (2013) Can the limitations of panel datasets be overcome by using pseudo-panels to estimate income mobility? Unpublished paper. Ithaca: Cornell University.

Cuesta, J. (2017) Comparing Cross-Survey Micro Imputation and Macro Projection Techniques: Poverty in Post Revolution Tunisia. Journal of Income Distribution

Cumpa, M. and Webb, R. (1999) Mobility and poverty dynamics in the 1990'. Paper presented at the IDS/IFPRI Workshop on Poverty Dynamics, April 1999, Institute of Development Studies, Brighton, United Kingdom.

Dabalen, A. Graham, E. Himelein, K. and Mungai, R. (2014) Estimating poverty in the absence of consumption data: The case of Liberia. Policy Research Working Paper 7024.

Daimon, T. (2001) The spatial dimension of welfare and poverty: Lessons from a regional targeting programme in Indonesia'. Asian Economic Journal 15 (4): 345-367

Dang, H.A. Lanjouw, P. Luoto, J. and McKenzie, D. (2011) Using repeated cross sections to explore movements in and out of poverty. World Bank Policy Research Working Paper 5550

Dang, H., and Lanjouw, P. (2013) Measuring poverty dynamics with synthetic panels based on cross-sections. *Policy Research Working Paper* 6504, World Bank

Dang, H., Lanjouw, P., Luoto, L., McKenzie, D. (2014) Using repeated cross-sections to explore movements into and out of poverty. *Journal of Developmen Economics* 107, 112–128 (2014)

Dang, H. and Dabalen, A (2017) Is poverty chronic or transient: Evidence from synthetic panel data. World Bank Policy Research Working Paper 8033

Dang, H. and Dabalen, A. (2018) Is Poverty in Africa Mostly Chronic or Transient? Evidence from Synthetic Panel Data. The Journal of Development Studies. 1-21. 10.1080/00220388.2017.1417585.

Datafirst (2019) Viewed 21 June 2019< https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/central/about

Davis, B. (2003) Choosing a Method for Poverty Mapping. Agriculture and Economic Analysis Division

Davies, R.B. (1994) From Cross-Sectional to Longitudinal Analysis, in Dale, A. and Davies, R.B. (Eds.) Analysing Social and Political Change. A Casebook of Methods, London: Sage Publications.

Davis, P. and B. Baulch. (2011) Parallel realities: Exploring poverty dynamics using mixed methods in rural Bangladesh. The Journal of Development Studies, 47(1): 118-142.

Desgroppes, A. and Taupin. S. (2011) Kibera: The Biggest Slum in Africa?. Les Cahiers de l'Afrique de l'Est, 2011, 44:.23-34

Demombynes, G. and Sandefur, J. (2014) Costing a data revolution. Centre for Global Development. Working Paper 383

De Vaus, D. (2001) Research design in social research. Sage Publications: London

Deaton, A. (1985) Panel Data from Times Series of Cross-sections, Journal of Econometrics, 30:1985.

Deaton, A. and Paxson, C. (1994) Intertemporal choice and inequality. Journal of Political Economy, 102, 437-67.

Deaton, A. (1997). The Analysis of Household Surveys. A Micro Econometric Approach to Development Policy. Baltimore, Maryland, and London: Johns Hopkins University Press.

Dercon, S. and Porter, C. (2011) A poor life? Chronic poverty and downward mobility in rural Ethiopia, 1994 to 2004. Why Poverty Persists: Poverty Dynamics in Asia and Africa. 65-95.

Devarajan, S. (2013) Africa's statistical tragedy. Review of Income and Wealth Series 59, Special Issue, October 2013 DOI: 10.1111/roiw.12013

Dollar, D. and Kray, A. (2002) Growth is good for the poor. *Journal of Economic Growth* 7(3): 195-225

Dreze, J. Lanjouw, P. and Stern, N. (1992) Economic Mobility and Agricultural Labour in Rural India: A Case Study. Indian Economic Review, Department of Economics, Delhi School of Economics 27:25-54.

Douidich, M. Ezzrari, A. Van der Weide, R. and Verme, P. (2013) Estimating Quarterly Poverty Rates Using Labour Force Surveys. Policy Research Working Paper 6466. Washington DC : World Bank.

Duclos, Jean-Yves, Araar, Abdelkrim and Giles, John, (2010) Chronic and transient poverty: Measurement and estimation, with evidence from China, Journal of Development Economics, 91(2): 266-277

Du Toit, A. (2005a) Chronic and structural poverty in South Africa: challenges for action and research. PLAAS Chronic Poverty and Development Policy Series No. 6 CCSR Working Paper 121

Du Toit, A. and Neves, D. (2008) Chronic and structural poverty in South Africa- An overview

Edwards, R. and Holland, J. (2006) What is qualitative interviewing. Bloomsbury: London

Elbers, C. J. O. Lanjouw, P. Lanjouw, and P. G. Leite (2001) Poverty and Inequality in Brazil: New Estimates from Combined PPV-PNAD Data," Unpublished Manuscript, The World Bank.

Elbers, C. Lanjouw, J.O. Lanjouw, P. (2002) Micro-level estimation of welfare. Policy Research Working Paper 2911. World Bank, Washington DC (2002)

Elbers, C. Lanjouw, J.O. Lanjouw, P. (2003) Micro-level estimation of poverty and inequality. Econometrica, 71(1), 355–364

Elbers, C. Lanjouw, J.O. and Lanjouw, P. (2005) Imputed Welfare Estimates in Regression Analysis, Journal of Economic geography, 5,101-118

Ellis F. and Freeman, H. A.(2004) Rural Livelihoods and Poverty Reduction Strategies in Four African Countries, The Journal of Development Studies, 40(4),1-30, DOI: 10.1080/00220380410001673175

Emmel, N. (2013) Sampling and choosing cases in qualitative research: A realist approach. London: Sage.

Filmer, D. and Pritchett, L. (2001). Estimating Wealth Effects Without Expenditure Data-or Tears: An Application to Educational Enrollments in States of India. Demography, 38(1):115–132.

Finn, A. and Leibbrandt, M. (2013) The Dynamics of Poverty in the First Three Waves of NIDS. SALDRU Working Paper 119/NIDS Discussion Paper 2013/1. Cape Town, South Africa: SALDRU, University of Cape Town

Finn, A. and Leibbrandt, M. (2017) The dynamics of poverty in the first four waves of NIDS. SALDRU Working Paper Number 174/ NIDS Discussion Paper 2016/1. Cape Town, South Africa: Southern Africa Labour and Development Research Unit (SALDRU), University of Cape Town.

Finn, A. Leibbrandt, M. Woolard, I. (2013) What happened to multidimensional poverty in South Africa between 1993 and 2010?. A Southern Africa Labour and Development Research Unit Working Paper Number 99. Cape Town: SALDRU, University of Cape Town

Finn, A. and Ranchold, V. (2013) Genuine Fakes: The prevalence and implications of fieldworker fraud in a large South African survey

Finn, A. Leibbrandt, M. and Oosthuizen, M. (2014a) Poverty, inequality and prices in post-Apartheid South Africa, WIDER Working Papers 2014/127,World Institute for Development Economics Research.

Finn, A. Leibbrandt, M. and Ranchhod, V. (2014b) Post-apartheid poverty and inequality trends, in R. Kanbur, ed. 'The Oxford Companion to the Economics of South Africa', Oxford University Press

Flik, R.J. and van Praag, B.M.S. (1991) De Economist 139:311. https://doi.org/10.1007/BF01423569

Food and Agriculture Organization (FAO), International Fund for Agricultural Development (IFAD), World Food Program (WFP) (2002) Reducing Poverty and Hunger: The Critical Role of Financing for Food, Agriculture and Rural Development

Foster, J.E. (1998) Relative versus absolute poverty. The American Economic Review, 88 (2): 335-341 Papers and proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association.

Foster, J. (2009). A Class of Chronic Poverty Measures. In T. Addison, D. Hulme & R. Kanbur (Eds.), Poverty Dynamics: Interdisciplinary Perspectives Oxford, Oxford University Press

Foster, J. E. and Santos, M. E. (2012) Measuring Chronic Poverty. Poverty and Social Exclusion: New Methods of Analysis. Routledge. Available at SSRN: https://ssrn.com/abstract=2376465

Fourie, J. (2016) The data revolution in African economic history. Journal of Interdisciplinary History, 47(2):193–212.

Friedman, M. (1957). A Theory of Consumption. Princeton University Press.

Gabriel, D. (2013) Inductive and deductive approaches to research | Dr Deborah Gabriel, viewed 20 March <http://deborahgabriel.com/2013/03/17/inductive-and-deductive-approaches-to-research

Gaiha R. and A.B. Deolalikar (1993) Persistent expected and innate poverty: Estimates for semi- arid rural South India, 1975–1984, Cambridge Journal of Economics 17(4); 409–421

Gaiha, R. (1988) Income Mobility in Rural India. Economic Development and Cultural Change, 36(2): 279-302.

Gardner, L (2018) Numerical Data and Statistical Sources

Gender Links (2019) Beyond Numbers: Gender audit of the May 2019 South African elections. Africa Women's Development Fund

Gibson, J. (2016) Poverty measurement: We know less than policy makers realize. Asia and the Pacific Policy Studies. 3 (3):430-442

Gimenez, M. E. (1987) The Feminization of Poverty: Myth or Reality? Insurgent Sociologist, 14(3): 5–30. doi: 10.1177/089692058701400302

Glewwe, P. and Hall, G. (1996) Who Is Most Vulnerable to Macroeconomic Shocks? Hypotheses Tests Using Panel Data from Peru. Living Standards Measurement Study Working Paper 117

Goedhart, T. V. Halberstadt, A. Kapteyn, and B.M.S. van Praag, (1977) The Poverty Line: Concept and Measurement, The Journal of Human Resources, 12:503-520

Gondwe, A.S. (2015) Asset index, child nutritional status and pro-poor growth analysis in Malawi. Stellenbosch Economic Working Ppaers 20/15. University of Stellenbosch

Gottschalk, P. (1980) Earnings Mobility: Permanent Change or Transitory Fluctuations? Review of Economics and Statistics 64:450-56.

Gouws, N.B. (1987)The demography of whites in South Africa before 1820. Department of National Health and Population Development, Pretoria.

Gradín, C. (2012) Race, Poverty, and Deprivation in South Africa. Journal of African Economies, 22(2):187–238

Graham, H. (1987) Being poor: perceptions and coping strategies of lone mothers, in J. Brannen and G. Wilson (eds) Give and Take in Families: Studies in Resource Distribution, London: Allen and Unwin: 56-74

Grant, U. (2010) Spatial inequality and urban poverty traps. Overseas Development Institute London United Kingdom

Greely, M. (1994) Measurement of poverty and poverty of measurement. IDS Bulletin 25.2, 1994

Green, P. (2017) NIDS Survey: A glimpse over high walls, in a country where many don't answer their doors. The Daily Maverick, viewed 19 May 2018,

https://www.dailymaverick.co.za/article/2017-11-21-nids-survey-a-glimpse-over-high-walls-in-a-country-where-many-dont-answer-their-doors/

Green, M. and Hulme, D. (2005) From Correlates and Characteristics to Causes: Thinking About Poverty from a Chronic Poverty Perspective. World Development, 33(6):867-879.

Grosse, M. Klasen, S. and Spatz, J. (2009) Matching Household with DHS Data to Create national Representative Time Series of Poverty: An Application to Bolivia. Courant Research Centre: Poverty, Equity and Growth – Discussion Paper 21. Courant Research Centre PEG.

Grootaert C. and Kanbur, R. (1993) The lucky few amidst economic decline : distributional change in Cote d'Ivoire as seen through panel data sets, 1985-88 (English). Policy, Research working papers; no. WPS 1113. Country operations. Washington, DC: World Bank. -change-in-Cote-dIvoire-as-seen-through-panel-data-sets-1985-88

Guell, M. and Hu, L. (2006) Estimating the Probability of Leaving Unemployment Using Uncompleted Spells from Repeated Cross-Section Data. Journal of Econometrics, 133: 307–341.

Guetterman T.C. (2015) Descriptions of Sampling Practices Within Five Approaches to Qualitative Research in Education and the Health Sciences

Guillerm, M. (2017) Pseudo panel methods and an example of application of household wealth data. Economics and Statistics

Gunning, J. and Hoddinott, J. and Kinsey, B. and Owens, T. (2000) Revisiting forever gained: Income dynamics in the resettlement areas of Zimbabwe, 1983-1997. The Centre for the Study of African Economies Working Paper Series. 36. 10.1080/00220380008422657.

Günther, I. and Klasen, S and (2007) Measuring Chronic Non-Income Poverty. Proceedings of the German Development Economics Conference, Göttingen 2007 10, Verein für Socialpolitik, Research Committee Development Economic

Hagenaars, J.M. and van Praag, B .(1985) A Synthesis of Poverty Line Definitions. Review of Income and Wealth. 31. 139-54. 10.1111/j.1475-4991.1985.tb00504.x

Haughton, J. and Khandker, S. (2009) Handbook on poverty and inequality. The World Bank, Washington D.C.

Handley G. Higgins, K. and Sharma B. Bird, K. and Cammack D. (2009) Poverty and poverty reduction in sub-Saharan Africa: An overview of the issues

Hanival, S and Maia, J, (2008) An overview of the performance of the South African economy since 1994. Paper commissioned by the Presidency as input to the 15 year South African review process. South Africa.

Harttgen, K. (2017) Data for poverty analysis and poverty evaluation in developing countries: Mathematical and data sciences for development. Zurich

Heine, K. and Oltmanns, E. (2016) Towards a political economy of statistics. Statistical Journal of the IAOS. 32. 201-209. 10.3233/SJI-160955.

Henninger, N. (2008) Mapping and geographic analysis of human welfare and poverty: Review and assessment. World Resources Institute, Washington D.C. USA.

Hentschel, J. Lanjouw, J. Lanjouw, P. and Poggi J. (2000) Combining census and survey data to trace the spatial dimensions of poverty: A case study of Ecuador. The World Bank Economic Review, 14(1):147-165.

Herault, N. and Jenkins, S. 2019. How valid are synthetic panel estimates of poverty dynamics? Journal of Economic Inequality 17: 51-76.

Heshmati, A., Maasoumi, E., and Wan, G. (2019) An analysis of the determinants of household consumption expenditure and poverty in India. Economies 7(96)2-27

Heywood, L. and Thornton, J. (1988) African Fiscal Systems as Sources for Demographic History: The Case of Central Angola, 1799–1920. Journal of African History 29 (1988): 213–228.

Higgins, K. Bird, K. and Harris, D (2010) Policy responses to the spatial dimensions of poverty. CPRC

Hill, H. (2002) Spatial Disparities in Developing East Asia: A Survey. Asian-Pacific Economic Literature 16(1): 10-36.

Hill, Martha S. (1981) "Some Dynamic Aspects of Poverty." in Martha S. Hill, Daniel B. Hill and James N. Morgan, editors. 5000 American Families: Patterns of Economic Progress, Vol. IX. Ann Arbor, Michigan: Institute for Social Research, 1981

Horrell, S. and Krishnan, P. (2015) Poverty and Productivity in Female-Headed Households in Zimbabwe. Journal of Development Studies 1-31

Hsiao, C (2005) Why Panel Data. Singapore Economic Review, 50(2): 1-12.

Hsiao, C. (2003) Analysis of Panel Data, Cambridge University Press, Second edition.

Hsiao, C (2007) Panel Data Analysis — Advantages and Challenges. Sociedad de Estad´ısticae Investigaci´on Operativa Test

Hoogeven, J.G. and Özler, B. (2006) Poverty and inequality in post-apartheid South Africa: 1995-2000.Results of ODI research presented in preliminary form for discussion and critical comment

Hulme, D. and McKay A. (2005) Identifying and Understanding Chronic Poverty: Beyond Monetary Measures. In: Kakwani N., Silber J. (eds) The Many Dimensions of Poverty. Palgrave Macmillan, London

Hulme, D. and Moore, K. and Shepherd, A. (2001) Chronic Poverty: Meanings and Analytical Frameworks. SSRN Electronic Journal. 10.2139/ssrn.1754546.

Hulme, D. And Shepherd, A. (2003) Conceptualizing chronic poverty. World Development 31(3), 403-423

Human Sciences and Research Council (HSRC) (2010) gender and poverty reduction: voice, dialogue and targeting policy. Policy Analysis and Capacity Enhancement (PACE)

Jackson, R. Drummond, D. and Camara, S. (2007) What Is Qualitative Research?. Qualitative Research Reports in Communication. 8: 21-28. 10.1080/17459430701617879.

Jalan, J. and Ravallion M. (1998) Transient Poverty in Postreform Rural China1. Journal of Comparative Economics 26:338–357

Jemna, L.M. (2016) Qualitative and mixed research methods in economics: the added value when using qualitative research methods. Journal of Public Administration, Finance and Law. Issue

Jan W. G. Hoddinott, J.F. Bill, K. and Trudy, O. (2000) Journal of Development Studies. 36 (6):134-154

Jerven, M. (2009) The Relativity of Poverty and Income: How Reliable are African Economic Statistics? African Affairs, 109(434): 77-96

Jerven, M. (2012) Briefing: For richer, for poorer: GDP revisions and Africa's statistical tragedy. African Affairs,112:138-147. 10.1093/afraf/ads063

Jerven, M. (2013a) Poor Numbers: How we are Misled by African Development Statistics and What to do About it, Ithaca, NY: Cornell University Press.

Jerven, M. (2013b) Comparability of GDP estimates in Sub-Saharan Africa: The effect of Revisions in Sources and Methods Since Structural Adjustment, Review of Income and Wealth.

Jerven, M. (2014a) Poor numbers and what to do about them. The Lancet, 383(9910): 594–595.

Jerven, M.(2014b) The political economy of agricultural statistics and input subsidies: Evidence from India, Nigeria and Malawi. Journal of Agrarian Change, Vol. 14(1)129–145.

Jerven, M. (2014c) Benefits and costs of the data for development targets for the Post-2015 Development Agenda. Data for Development Assessment Paper Working Paper, September, Copenhagen: Copenhagen Consensus Center.

Jerven, M. (2016) Development by numbers: A Primer, Development Research Institute. Working Paper. Forthcoming, viewed 7 September 2017 https://wp.nyu.edu/dri/publications/working-papers/.

Jerven, M. (2017). How much will a data revolution in development cost?. In Forum for Development Studies 44(1(: 31-50. Routledge

Jerven, M. and Johnston, D. (2015) Statistical tragedy in Africa?: Evaluating the Database for African economic development. Routledge: Abingdon.

Johannes, G. Hoogeveen, G. and Ozler, B. (2005) Nor separate, not equal: Poverty and inequality in post-apartheid South Africa. William Davidson Institute Working Paper 73

Johnston, D. and Abreu, A. (2014) Asset indices as a proxy for poverty measurement in African countries: A reassessment. Paper presented at the African Economic Development: Measuring Success and Failure, Simon Fraser University, Vancouver, Canada.

Johnson, R. B. and Onwuegbuzie, A. J. (2004) Methods Research: A Research Paradigm Whose Time Has Come. Educational Researcher, 33(7):14–26. doi: 10.3102/0013189X033007014.

Jolliffe, D. Lanjouw, P. Chen, A. Kraay, C. Meyer, M. Negre. E. Prydz, R. Vakis, and K. Wethli. (2014) A Measured Approach to Ending Poverty and Boosting Shared Prosperity: Concepts, Data, and the Twin Goals. Policy Research Report. Washington, DC: World Bank, 2014.

Kabeer, N. (2003) Gender Mainstreaming in Poverty Eradication and the Millennium Development Goals: A Handbook for Policy-makers and Other Stakeholders (London: Commonwealth Secretariat)

Kamanou, G. Morduch, J. David, I.P. and Gibson, J. (2005) Handbook on poverty statistics: Concepts, methods and policy use. United Nations Statistics Division.

Kanbur, R. (2002) Notes on the policy significance of inequality decompositions. Mimeo, Cornell University (later published in *Journal of Economic Inequality* 4(3):367-3740).

Kekeya, J. (2016) Analyzing qualitative data using an iterative process. Contemporary PNG Studies: DWU Research Journal, 24:87-94

Klasen, S. (2005) Economic Growth and Poverty Reduction: Measurement and Policy Issues (September 2005). OECD Working Paper No. 246.

Klasen, S. and Lechtenfeld, T. and Povel, F. (2015) A Feminization of Vulnerability? Female Headship, Poverty, and Vulnerability in Thailand and Vietnam. World Development, Elsevier,71: 36-53.

Krätke, F. Byiers, B. (2014) The political economy of official statistics: Implications for the data revolution in Sub-Saharan Africa. (Discussion Paper 170; In partnership with PARIS21- Partnership in Statistics for Development in the 21st Century - Discussion Paper No. 5). Maastricht: ECDPM.

Krishna, A. (2004) Escaping poverty and becoming poor: who gains, who loses, and why? World Development, 32(1):121–136.

Krishna, A. Kristjanson, P. Radeny, M. and Nindo, W. 2004. Escaping poverty and becoming poor in 20 Kenyan villages. Journal of Human Development, 5(2): 211–226

Krishna, A. Lumonya, D. Markiewicz, M. Mugumya, F. Kafuko, A. and Wegoye, J. (2006a) Escaping poverty and becoming poor in 36 villages of Central and Western Uganda. Journal of Development Studies, 42(2):346–370.

Krishna, A. Kristjanson, P. Kuan, J. Quilca, G. Radeny, M. and Sanchez-Urrelo, A. (2006b) Fixing the hole in the bucket: household poverty dynamics in forty communities of the Peruvian Andes. Development and Change, 37(5): 997–1021.

Kristjanson, P. Mango, N. Krishna, A. Radeny, M. and Johnson,. (2010). Understanding poverty dynamics in Kenya. Journal of International Development. 22:978 - 996.

Kumari, R. (1989) Women-headed Households in Rural India, New Delhi: Radiant Publishers

Lanjouw, J.O. (2001) Demystifying poverty lines. UNDP. Accessed online at: http://siteresources.worldbank.org/PGLP/Resources/LanjouwDemystifyingPovertyLines.pdf

Lanjouw, J. and Lanjouw, P. (1997) Poverty comparisons with non-compatible data: Theory and Illustrations. *World Bank Policy Research Working Paper* 1709

Lee, U. (2003) Panel attrition in survey data: A literature review. Centre for Social Science Research. CSSR Working Paper No. 41. University of Cape Town.

Leibbrandt, M. Woolard, I. Finn, A. and Argent, J. (2010) Trends in South African income distribution and poverty since the fall of apartheid, OECD Social, Employment and Migration Working Papers 101, OECD Publishing.

Lipton, M. (1983) The poor of the poorest: Some interim findings. World Bank Discussion Papers

Lohano, H.R. (2011) Poverty dynamics in rural Sindh, Pakistan, 1987-88 to 2004-05 in B. Baulch (Ed.), Why Poverty Persists: Poverty Dynamics in Asia and Africa, Northhampton: Edward Elgar Publishing.

Lokshin, M. and Ravallion, M. (2004) Household income dynamics in two transition economies. Studies in Nonlinear Dynamics and Econometrics, 8(3).

Lehohla, P. (2005) Statistics South Africa in Transition: Reflections on a decade of Statistical Practice (1994-2004)

Lehohla, P. (2002) Promoting statistical literacy: a south African perspective promoting statistical literacy: A South African perspective.

Levy, F. (1977) How Big is the American Underclass? Working Paper #0090—1, Washington, D.C.: The Urban Institute, 1977.

Lewis, D. (1993) 'Going it alone: female-headed households, rights and resources in rural Bangladesh', European Journal of Development Research 5(2:): 23-42

Lusardi, A. (1996) Permanent income, current income and consumption evidence from two datasets. Journal of Business and Economic Statistics 14(1):181-190

Machado, A. and Ribas, R.P. (2010) Do Changes in the Labour Market Take Families Out of Poverty? Determinants of Exiting Poverty in Brazilian Metropolitan Regions Journal of Development Studies, 46(9): 1503-1522.

Magasela, W. (2005) Constructing and adopting an official poverty line for South Africa: some issues for consideration. NALEDI, 2005.

Malaba J. (2006) Poverty Measurement and Gender's Experience. United Nations Department of Economic and Social Affairs. New York.

Malluccio, J. Haddad, M. and May, J. (1999) Social capital and income generation in South Africa, 1993-1998. Paper presented at the IDS/IFPRI Workshop on Poverty Dynamics, April 1999, Institute of Development Studies, Brighton, United Kingdom.

Martinez, A. Western, M. Haynes, M. and Tomaszewski W. (2013) Measuring Income Mobility Using Pseudo-Panel Data. Philippine Statistician, 62(2): 71-99.

May, J. D. (2001) An elusive consensus: Definitions, measurement and analysis of poverty. United Nations Development Programme, New York

May, J. (2003) Chronic Poverty and Older People in South Africa. Chronic Poverty Research Centre Working Paper no. 25

May, J.D. (2008) The persistence of poverty in post-apartheid South Africa: Assets, livelihoods and differentiation in KwaZulu Natal, 1993-2004. PhD Thesis, University of KwaZulu Natal, Durban

May, J. D. (2010) Poverty eradication: The South African Experience. Department of Economic and Social Affairs, Division for Social Policy and Development and Economic Commission for Africa, Economic Development and NEPAD Division Expert Group Meeting on Poverty Eradication. 15-17 September 2010, United Nations Conference Centre, Addis Ababa, Ethiopia.

May, J.D. (2014) Sustaining effective anti-poverty programmes beyond transformation: Challenges and way forward. School of Development Studies, University of KwaZulu-Natal

May, J. D, Jorge A. Carter, M.R,and Timaeus, I. (2007) The KwaZulu-Natal Income Dynamics Study (KIDS) third wave: Methods, first findings and an agenda for future research. Development Southern Africa 24,(5): 629-648

May, J. Carter, M. Haddad, L. and Malluccio, J. (2000) Kwa-Zulu Natal Income Dynamics Study (KIDS) 1993-1998: a longitudinal household data set for Southern African policy analysis, Development Southern Africa 17(4): 567-581.

May, J. and Hunter, N. (2004) Poverty, Inequality and the Social Wage: Balancing the books in Post-Apartheid South Africa. Paper presented at the conference on Politics of Socio-Economic Rights

May J. D. and Lehohla, P. (2005) Counting the costs of a twenty-first century census: South Africa's census 2001, Development Southern Africa, 22:(2):215-232, DOI:10.1080/03768350500163279

May, J. and Woolard, I. (2007) Poverty Traps and Structural Poverty in South Africa: Reassessing the Evidence from Kwazulu-Natal. Why Poverty Persists: Poverty Dynamics in Asia and Africa. 10.2139/ssrn.1752965

May, J. Woolard, I. and Baulch, B. (2011) Poverty traps and structural poverty in South Africa: reassessing the evidence from KwaZulu-Natal, 1993 to 2004 in Bob Baulch (ed.) Why Poverty Persists: Poverty Dynamics in Asia and Africa, Edward Elgar Publishing, Cheltenham, England, pp. 187-218

May, J, Woolard, I. and Baulch, B. (2011) Poverty traps and structural poverty in South Africa: reassessing the evidence from KwaZulu-Natal, 1993 to 2004 in Bob Baulch (ed.) Why Poverty Persists: Poverty Dynamics in Asia and Africa, Edward Elgar Publishing, Cheltenham, England, pp. 187-218.

Marx, B. Stoker, T. and Suri, T. (2013) The Economics of Slums in the Developing World Journal of Economic Perspectives, 27 (4): 187-210

Mckay, A. (2009). Assets and Chronic Poverty. In Shepherd, A. and Brunt, J. (eds) Chronic poverty. Rethinking International Development Series. Palgrave MacMillan, London

McKay, A. and Lawson, D. (2003) Assessing the Extent and Nature of Chronic Poverty in Low Income Countries: Issues and Evidence. University of Nottingham, United Kingdom.

McKenzie, D. (2004) Consumption Growth in a Booming Economy: Taiwan 1976-96, Yale University Economic Growth Center Discussion Paper no. 823.

Meng, Y. Brennan, A. Prushouse, R. Hill-McMnaus, D. Angu, C. (2014) Estimation of own and cost price elasticities of alcohol demand in the UK-A pseudo panel approach using the Living Costs and Food Survey 2001-2009. Journal of Health Economics 2014(34): 96-103

Merola, G. and Baulch, B. (2019) Using sparse categorical principal components analysis to estimate asset indices: a new method with application to south-east Asia. Review of Development Economics 23(May): *640-662*

Merriam, S.B. (2010) Qualitative Case Studies. International Encyclopedia of Education. 10:456-462.

Meth, C. (2006) Income poverty in 2004: A second engagement with the recent van der berg et al figures. Working Paper 47. Durban: University of Kwazulu-Natal.

Milazzo, A. and van de Walle, D. (2015) Women Left Behind? Poverty and Headship in Africa. Demography. 54. 10.1007/s13524-017-0561-7.

Millar, J. (1992) Lone mothers and poverty, in C. Glendinning and J. Millar (eds) Women and Poverty in Britain in the 1990s, Hemel Hempstead: Harvester Wheatsheaf: 149-61

Mincer, J. (1958). Investment in human capital and personal income distribution. Journal of Political Economy. 66(4): 281–302.

Minot, N. and Baulch, B. (2005) Spatial patterns of poverty in Vietnam and their implications for policy. Food Policy. 30:461-475. 10.1016/j.foodpol.2005.09.002.

Minot, N. Baulch, B. and Epprecht, M. (2003) Poverty and Inequality in Vietnam: Spatial Patterns and Geographic Determinants. Washington, DC and Brighton: IFPRI and IDS

Minot, N. and Epprecht, M. and Thi Tram A. Tran and Quang T. (2006) Income Diversification and Poverty in the Northern Uplands of Vietnam. Research Report of the International Food Policy Research Institute. 10.2499/0896291480

Modigliani, F. (1966) The life cycle hypothesis of saving, the demand for wealth and the supply of capital._Social Research. 33(2): 160–217.

Mo Ibrahim Foundation (MIF) 2016) Strength in numbers: Africa's data revolution.

Moghadam, V. (1996) The feminization of poverty: notes on a concept and trends. Report prepared for United Nations Development Programme, Human Development Report, September

Morgan, J.N. (1974) Change in Global Measures in Five Thousand American Families-Patterns of Economic Progress. Vol. 1, ed. Morgan et al. (Ann Arbor: Institute for Social Research, University of Michigan

Moghadam, V. (2005) The Feminization of Poverty and Women's Human Rights.

Mofitt, R. (1993) Identification and estimation of dynamic models with a time series of repeated cross-sections. Journal of Econometrics, 59:99- 123.

Morse, J.M. and Niehaus, L. (2009) Mixed methods design: Principles and procedures. Left Coast Press, Inc: California

Muller, C. (1997) Transient seasonal and chronic poverty of peasants: evidence from Rwanda. The Centre for the Study of African Economies Working Paper Series.

Nargis, N. and M. Hossain. (2006) Income Dynamics and Pathways Out of Rural Poverty in Bangladesh, 1988-2004. Agricultural Economics, 35: 425-435.

National Income dynamics Study (NIDS) (2019) What is NIDS? Accessed online at http://www.nids.uct.ac.za/about/what-is-nids

Neilson, C. D. Contreras, R. Cooper, and J. Hermann, (2008) The Dynamics of Poverty in Chile. Journal of Latin American Studies, 40(2): 251-273.

Newman, I and Benz, C.R. (1998) Qualitative-Quantitative Research Methodology_ Exploring the Interactive Continuum. Southern Illinois University Press.

Newhouse, D. Shivakumaran, S. Takamatsu, S. and Yoshida N. (2014) How survey-to-survey imputation can fail. World Bank Policy Research Paper 6961

Noponen, H. (1991) The Dynamics of Work and Survival for the Urban Poor: A Gender Analysis of Panel Data from Madras. Development and Change, 22: 233-260. doi:10.1111/j.1467-7660.1991.tb00410.x

Oelmaum, J. (2010) Spatial poverty traps and ethnic conflict traps: Lessons from Northern Ghana's 'Blood Yams'. ODI/CPRC Working Paper Series (ODI WP324, CPRC WP164. London and Manchester, University of Manchester.

Oloo, G. (2011) Implementing community based multidimensional poverty Monitoring System in Limpopo province South Africa. Poverty and Economic Policy Research Network. 9th General Meeting Angkor Era Hotel Siem Reap, Cambodia December 3-9, 2011

Omair, A. (2015) Selecting the appropriate study design for your research: Descriptive study designs. Department of Medical Education, Research Unit, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia

Oxford Poverty and Human Development Initiative (OPHI) (2015) Measuring multidimensional poverty: Insights from around the world. University of Oxford.

Orkin, F.H. (1996) Living in Mpumalanga Selected findings of the 1995 October household survey. Central Statistics, South Africa.

Parker, B. and Kozel, V. (2005) Understanding poverty and vulnerability in Indias Uttar Pradesh and Bihar: A Q-squared approach. Q-Squared Working Paper Series, No. 9, Centre for International Studies, University of Toronto.

Parker, J.A. Souleles, N.S. and Caroll, C.D. (2014) The benefits of panel data in consumer expenditure surveys. Prepared for the NBER CRIW Conference on Consumption Measurement, December 2011.

Parsons, J.A. (2008) Key informants in Lavrakas, P. J. (2008) Encyclopedia of survey research methods Thousand Oaks, CA: Sage Publications, Inc. doi: 10.4135/9781412963947

Patel, L. Hotchfield, T. Moodley, J. and Mutwali, L. (2012) The gender dynamics and the impact of the Child Support Grant in Doornkoop, Soweto. Research Report, University of Johannesburg.

Patton, M.Q. (2002) Qualitative research and evaluation methods. 3rd Sage publications: Thousand oaks, CA: 2002

Pearce, D. (1978) The Feminization of Poverty: Women, Work and Welfare. Urban and Social Change Review 11: 28-36.

Pencavel, John. (2007) A Life Cycle Perspective on Changes in Earnings Inequality among Married Men and Women. Review of Economics and Statistics, 88(2): 232-242.

Phiri, I. P. and Tangamodzi, L.T. (2016) Comparative analysis of cross sectional and panel data collection for agricultural statistics in Malawi: Does an integrated vehicle provide an effective framework. Seventh International Conference on Agricultural Statistics I Rome 24-26 October 2016

Porter C. and N. Quinn (2008) Intertemporal poverty measurement: Tradeoffs and policy options. Centre for the Study of African Economies working paper series 2008-21, Oxford.

Posel, D. and Rogan, M. (2009). Women, income and poverty: gendered access to resources in post-apartheid South Africa. Agenda, 81(25)

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

Posel, D. and Rogan, M. (2013) Measured as poor versus feeling poor: Comparing objective and subjective poverty rates in South Africa. Paper prepared for the ESSA Conference, September

Posel P. and Michael Rogan. (2017) Inequality, social comparisons and minimum income aspirations: Evidence from South Africa". ISER Working Paper No. 2017/3. Grahamstown: Institute of Social and Economic Research. Rhodes University

Project for Statistics on Living Standards and Development PSLSD (1994): South Africans rich and poor: Baseline household statistics, South African Labour and Development Research Unit, University of Cape Town, South Africa.

Punch, K. F. (2013). Introduction to Social Research: Quantitative and Qualitative Approaches. SAGE Publications.

Quisumbing, A. (2011) Poverty transitions, shocks and consumption in rural Bangladesh, 1996-97 to 2006-07", in B. Baulch (Ed.), Why Poverty Persists: Poverty Dynamics in Asia and Africa, Northhampton: Edward Elgar Publishing.

Quisumbing, Agnes, Lawrence Haddad and Christine Pena (2001) Are women over-represented among the poor? An analysis of poverty in 10 developing countries', Journal of Development Economics (66):225-269

Quisumbing,A. and B. Baulch (2009) Assets and poverty traps in rural Bangladesh", Chronic Poverty Research Centre Research Working Paper No. 143, Manchester.

Rainwater, L. (1982) Persistent and Transitory Poverty: A New Look,Working Paper 70. Cambridge, Mass.: Joint Center for Urban Studies.

Rajshri J. and Lanjouw, P. (1999) The Evolution of Poverty and Inequality in Indian Villages. World Bank Research Observer. 14(10):1-30.

Ravallion, M. (1998) Poverty lines in theory and practice. Living Standards Measurement Study Working Paper No. 133. Washington: The World Bank.

Ravallion, M. (2012) Poor, or Just Feeling Poor? On Using Subjective Data in Measuring Poverty. Policy Research Working Paper;No. WPS 5968. World Bank.

Ravallion, M. and Chen, S. (1997) What can new survey data tell us about reent changes in distribution and poverty. World Bank Economic Review, 11(2):357-387.

Reardon, T. and Taylor, J. E. (1996) Agroclimatic shock, income inequality, and poverty: Evidence from Burkina Faso, World Development, Elsevier, 24(5):901-914.

Rentz, J. Reynolds, F. and Stout, R. (1983) Analyzing Changing Consumption Patterns with Cohort Analysis. Journal of Marketing Research, 20(1):12-20. doi:10.2307/3151407

Richter, L. M. Norris, S. A. and De Wet, T. (2004) Transition from Birth to Ten to Birth to Twenty: The South African cohort reaches 13 years of age. Paediatric and perinatal epidemiology, 18(4):290–301. doi:10.1111/j.1365-3016.2004.00572.x

Richter L. Norris S. Pettifor J. Yach D. Cameron N. (2007) Cohort profile: Mandela's children: The 1990 Birth to Twenty study in South Africa. Int J Epidemiol, 36(3):504-511.

Riley, M.W. (1972) Aging and Cohort Succession: Interpretations and Misinterpretations. Public Opinion Quarterly, 37:35-49.

Roberts, B. (2000) Chronic and Transitory Poverty in Post-Apartheid South Africa: Evidence from KwaZulu-Natal. CSDS Working Paper No. 28, School of Development Studies: University of Natal

Rowntree, B.S. (1901) Poverty: A Study of Town Life. London.: MacMillan

Ruspini, E. (2000) Longitudinal Research in the Social Sciences. Social Research Update. Issue 20

Ryder, Norman B. (1965) The Cohort as a Concept in the Study of Social Change. American Sociological Review, 30:843-61.

Sahn, D. and Stifel, D. (2003) Exploring alternative measures of welfare in the absence of expenditure data', *Review of Income and Wealth* 49(3): 463-89

Sandefur J. and Amanda Glassman (2015) The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics, The Journal of Development Studies, 51(2): 116-132

Sanga, D. (2011) The challenges of the narrative of African countries development: Data demand and supply mismatches.

Schaie, K.W. (1965) A general model for the study of developmental problems. Psychological Bulletin, 61:92-107.

Schotte, S. Zizzamia, R. and Leibbrandt, M. (2017a) Social Class, Life Chances and Vulnerability to Poverty in South Africa. Southern Africa Labour and Development Research Unit Working Paper 208. Cape Town: SALDRU, University of Cape Town.

Schotte, S. Zizzamia, R. and Leibbrandt, M. (2017b) Assessing the Extent and Nature of Chronic Poverty, Poverty Dynamics, and Vulnerability to Poverty in South Africa .World Bank Technical Background Paper, South African Poverty and Inequality Assessment Report 2018.

Schotte, S. Zizzamia, R. and Leibbrandt, M. (2018) A poverty dynamics approach to social stratification: The South African case. World Development, 110, 88-103.

Scott, L and Mariotti, C. (2014) A place for panel data in the 'data revolution'. ODI Briefing

Scott, C. D. and Litchfield, J. A. (1994) Inequality, mobility and the determinants of income among the rural poor in Chile, 1968-1986. London (The Development Economics Research Programme Suntory Toyota International Centre for Economics and Related Disciplines, London School of Economics and Political Science DEP, 53).

Serajuddin, U. Uematsu, H. Wieser, C. Yoshida, N. Dabalen, A.L. (2015) Data deprivation : another deprivation to end (English). Policy Research working paper; no. WPS 7252. Washington, D.C. : World Bank Group. http://documents.worldbank.org/curated/en/700611468172787967/Data-deprivation-another-deprivation-to-end

Shah, A. (2010) Patterns of poverty in remote rural areas: a case study of a forest-based region in Southern Orissa in India. ODI/CPRC Working Paper Series (ODI WP325, CPRC WP165). London: ODI and Manchester: CPRC, University of Manchester.

Simkins, C. (2005) What happened to the distribution of Income in South Africa between 1995 and 2001?. Unpublished Paper. University of the Witwatersrand.

Stifel, D. and Christiansen, L. (2007) Tracking Poverty Over Time in the Absence of Comparable Consumption Data. The World Bank Economic Review, 21(2):. 317-341

Son H. and Kakwani, N. (2004) Economic growth and poverty reduction: initial conditions matter. Working Papers 2, International Policy Centre for Inclusive Growth.

Spencer, B. May, J.D. Kenyon, S. and Seeskin, Z. (2017) Cost-Benefit Analysis for a Quinquennial Census: The 2016 Population Census of South Africa. Journal of Official Statistics. 33. 10.1515/jos-2017-0013.

Srivastava, P. and Hopwood, N. (2009) A Practical Iterative Framework for Qualitative Data Analysis. International Journal of Qualitative Methods. 8. 76-84. 10.1177/160940690900800107.

Starr, M. (2016) Qualitative and mixed-methods research in economics: Surprising growth, promising future. Journal of Economic Surveys. 28. 10.1111/joes.12004

Statistics South Africa (Stats SA) (1997) Earnings and spending in South Africa: Selected findings of the 1995 Income and Expenditure Survey. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (1999) Rural Survey 1997. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2000) Measuring poverty in South Africa: Pretoria: Statistics South Africa

Statistics South Africa (Stats SA) (2001) South Africa in transition: Selected findings from the October Household Survey of 1999 and changes that have occurred between 1995 and 1999.Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2002) Poverty Report: Measuring poverty in South Africa. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2003) Investigating the appropriate definitions of urban and rural areas for South Africa: Discussing Document Report No 03-02-20 2001. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2006) Deprivation: The provincial indices of multiple deprivation for South Africa 2001. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2007) A discussion note: Constructing comparable household survey data for the analysis of poverty analysis in South Africa (1995-2000). Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2008a) Methodological report on the development of the poverty lines for statistical reporting. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2008b) Income and expenditure of households 2005/2006: Analysis of results. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2010) Social profile of South Africa (2002-2009). Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2011) Social profile of vulnerable groups in South Africa (2002-2010). Pretoria: Statistics South Africa

Statistics South Africa (Stats SA) (2012a) Subjective poverty in South Africa: Findings from the Living Conditions Survey 2008/9. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2012b) Poverty profile of South Africa: Application of poverty lines on the Living Conditions Survey 2008/9. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2012c) Income and Expenditure of Households 2010/2011 Statistical release P0100. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2013a) Social profile of vulnerable groups 2002-2012. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2013b) Men, women and children: Findings from the Living Conditions Survey 2008/9. Pretoria: Statistics South Africa.

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

Statistics South Africa (Stats SA) (2014b) The South Africa Multidimensional Poverty Index: Creating a multidimensional poverty index using Census data. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2015a) Income dynamics in poverty status of households in South Africa. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2015b) Methodological report on rebasing of national poverty lines and development of pilot provincial poverty lines: Technical Report. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2015c) Millennium Development Goals: Country Report 2015. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2016) Vulnerable Groups Series 1: The social profile of youth, 2009-2014.Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2017a) Vulnerable Group Series II: The social profile of older persons, 2011-2015.Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2017b) Poverty trends in South Africa: An examination of absolute poverty between 2006 and 2015. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2017c) Living conditions of households in South Africa: An analysis of the household expenditure and income data using the Living Conditions Survey of 2014/15: Statistical Release P0310. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2017d) Annual Report 2016/17. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2018a) Applying small area estimation techniques on IES (2010/11) and Census (2011). Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2018b) Annual Report 2017/18. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2018c) Statistical release: National poverty lines. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2018d) National poverty lines: Statistical release P0310.1. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2018e) Subjective poverty in South Africa: Findings from the Living Conditions Survey of 2008/9 and 2014/15. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2018f) Men, women, and Children: Findings of the LCS of 2014/15. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2019a) Vulnerable Groups Report: The social profile of children 7-17 years old, 2002-2016. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2019b) Statistical Report P0310.1: National poverty lines. Pretoria: Statistics South Africa.

Statistics South Africa (Stats SA) (2019c) Mid-year population estimates: Statistical Report P0302. Pretoria: Statistics South Africa

Statistics South Africa (Stats SA) (2020) Statement from the South African Statistical Council on Statistics South Africa funding. Statement by Professor Everatt, D. Online: http://www.statssa.gov.za/?p=12992

Tarozzi, A. (2007) Calculating comparable statistics from incomparable surveys, with an application to poverty in India. Journal of Business and Economic Statistics: 25(3): 314-336

Thornbecke, E. (2008) Conceptual and Measurement Issues in Poverty Analysis. UNU World Institute for Development Economics Research (UNU-WIDER) Helsinki, Finland

Thomas, D. R. (2006). A General Inductive Approach for Analyzing Qualitative Evaluation Data. American Journal of Evaluation, 27(2):237–246. https://doi.org/10.1177/1098214005283748

Triegaart, J.D. (2009) Poverty and inequality in South Africa: Policy considerations in an emerging democracy. Paper presented at the annual Association of South African Social Work Education Institutions (ASASWEI) conference organised by University of Venda Department of Social Work. 18-20 September 2006.

Townsend, P. (1962) 'The meaning of poverty', British Journal of Sociology, 13(3):210-19.

United Nations (UN) (2010) Rethinking poverty: Rethinking on the World social situation. Economic and Social Affairs. New York

United Nations (UN) (2015) Millennium Development Goals Report 2015, New York

United Nations Development Program (UNDP) (2004) South Africa Human Development Report 2003: The Challenge of Sustainable Development. UNDP South Africa: Oxford University Press.

United Nations Economic Commission for Africa (UNECA) (2016) The Demographic Profile of African Countries. Addis Ababa, Ethiopia

United Nations Human Settlements Programme (UNHABITAT) (2016) Slums Almanac 2015-2016: Tracking Improvement in the Lives of Slum Dwellers

United Nations (UN) WOMEN (2017) UN Women and the World Bank unveil new data analysis on women and

povertyhttps://www.unwomen.org/en/news/stories/2017/11/news-un-women-and-the-world-bank-unveil-new-data-analysis-on-women-and-poverty

Urzainqui, D.G. (2017) Poverty Transitions without Panel Data? An Appraisal of Synthetic Panel Methods. Paper presented at the ECINEQ Conference, New York City

United States Agency for International Aid (USAID) (1996) Center for Development Information and Evaluation. Conducting Key Informant Interviews Performance Monitoring and Evaluation TIPS. Washington DC, USAID.

Van der Berg, S. Louw, M. and Yu, D (2006) Post-Transition poverty trends based on an alternative data source. Stellenbosch Economic Working Papers 08/07

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

Van der Berg S. and Ronelle Burger and Rulof Burger and Megan Louw and Derek Yu, (2006) Trends in Poverty and Inequality since the Political Transition. Working Papers 06104, University of Cape Town, Development Policy Research Unit

Van der Berg, S. Louw, M. and Yu, D. (2008) Post-Transition Poverty Trends Based On An Alternative Data Source. South African Journal of Economics, Economic Society of South Africa, 76(1);58-76.

Van der Berg, S. Burger, C. Burger, R. De Vos, M. Du Rand, G. Gustafsson, M. Moses, E. Shepherd, D. Spaull, N. Taylor, S. Van Broekhuizen, H. and Von Fintel, D. (2011) Low quality education as a poverty trap. University of Stellenbosch: Stellenbosch.

Van der Berg, S. Burger, R. Burger, R. Louw, M. and Yu, D (2005) Trends in poverty and inequality since the political transition. Stellenbosch Economic Working Papers:1/2005. University of Stellenbosch

Van Wyk B. (2000) Research design and methods. Lecture Notes Part 2. Postgraduate enrolment and throughput. University of the Western Cape

Van Wyk, B. (2012) Research design and methods Part 1," University of Western Cape. Viewed 09 August, 2016, https://www.uwc.ac.za/Students/Postgraduate/Documents/Research_and_Design_I.pdf

Vandemoortele, J. (2009) The MDG Conundrum: Meeting the Targets without Missing the Point', Development Policy Review, Vol 27, No. 4, pp. 355–71

Verbeek, M. (1993) Pseudo panel data. Reprint series/ Centre for Economic Research

Verbeek, M. (2008) Pseudo-panels and repeated cross sections.

Vermaak, C. (2005) Trends in income distribution, inequality and poverty in South Africa, 1995 to 2003. Paper presented at the Biennial Economics Society of South Africa Conference held at Durban, 7-9 September.

Visser, P. S. Krosnick, J. A. and Lavrakas, P. (2000) Survey research methods. In H. T. Reis and C. M. Judd (Eds.), Handbook of research methods in social psychology (pp. 223–252). New York: Cambridge University Press.

Von Fintel. D. and Zoch, A. (2015) The dynamics of child poverty in South Africa between 2008 and 2012: An analysis using the National Income Dynamics Study. A Working Paper of the Department of Economics and the Bureau for Economic Research at the University of Stellenbosch

Woolard, I. and Leibbrandt, M. (2006) Towards a poverty line for South Africa: Background note. Cape Town: Southern Africa Labour and Development Research Unit.

World Bank (2011) Perspectives on poverty in India: Stylized facts from survey data. Washington D.C. USA.

Wijesekere W. (2009) The role of household panel surveys in evidence-based policy. a revised version of the paper presented to the Australasian Evaluation Society International Evaluation

Williams, R. (2018) Panel data 4: Fixed effects and random effects models. University of Notre Dame.

World Bank (2009) Reshaping economic geography: World Development Report 2009. Washington D.C, United States of America.

World Bank (2016) Poverty and shared prosperity: Taking on inequality. Washington D.C. USA

World Bank (2017) Monitoring global poverty: Report of the commission on global poverty. Washington D.C, United States of America.

World Bank (2018) Piecing together the poverty puzzle: Poverty and shared prosperity 2018

World Bank and Statistics South Africa (2018) Overcoming poverty and inequality in South Africa: An assessment of drivers, constraints and opportunities

World Bank Central Data Catalogue (2019) viewed 29 August 2019, https://microdata.worldbank.org/index.php/catalog

Woolard, I. and Klasen, S. (2005) Determinants of Income Mobility and Household Poverty Dynamics in South Africa, Journal of Development Studies, 41(5):865-897,

Wooldridge, J.M. (2009) Econometric analysis of cross section and panel data. The MIT Press Cambridge.

Yin, R.K. (1981a) The case study crisis: Some Answers. Administrative Science Quarterly, 26(1):58-65.

Yin, R.K. (1981b) The case study as a serious research strategy. Knowledge 3(1);97-114.

Yin, R. K. (1989) Case study research: Design and methods. London: Sage Publications

Young, Alwyn, (2013) The African Growth Miracle, No 18490, NBER Working Papers, National Bureau of Economic Research.

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

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

Yaqub, S. (2000) Poverty dynamics in developing countries. Institute of Development Studies, University of Sussex

Zizzammia, R, Schotte, S. and Leibbrandt, M. (2018) Snakes and Ladders and Loaded Dice: Poverty dynamics and inequality in South Africa between 2008-2017. SALDRU

Zuberi, T. and Khalfani, A.K. (1999) Racial classification and the census in South Africa.

Zuberi, T. Sibanda, A. and Udjo E. (Eds.) (2005) The Demography of South Africa. M.E. Sharpe Inc. Armonk, New York, London.

Appendix

APPENDIX A

Information sheet

UNIVERSITY of the WESTERN CAPE
INSTITUTE FOR SOCIAL DEVELOPMENT
UNIVERSITY of the WESTERN CAPE
University of the Western Cape

Private Bag X17, Bellville 7535, Cape Town, South Africa

Telephone :(021) 959 3858/6 Fax: (021) 959 3865

E-mail: ***pkippie@uwc.ac.za*** *or* spenderis@uwc.ac.za

INFORMATION SHEET FOR STUDY PARTICIPANTS

DATE: 31/01/2017

Study title: A gendered approach to poverty and inequality analysis in South Africa: An application of synthetic panel data

Purpose of the study:

This research seeks to promote the use of alternative data sources in poverty analysis in South Africa. This study uses a pseudo panel data that will be created using data from the Living Conditions Surveys of 2008/9 and 2014/15 as well as the IES of 2010 to analyze poverty. The objective of this qualitative study is to understand what Statistics South Africa (Stats SA) has done and is doing as far as the measurement of poverty is concerned in the post-Apartheid period. This qualitative research seeks to capture the experiences of experts involved in the collection and processing of official poverty statistics. Issues relating to how the organization deals with data problems relating data unavailability, incomparability, and infrequency of poverty related data will form part of the discussion. The central argument in this thesis is centered on the importance of accurate and timely poverty surveys for effective poverty monitoring and subsequently its reduction.

Contact details for researcher and supervisor:

Researcher:

Rejoice Mabhena

APPENDIX B

Letter of consent

University of the Western Cape
Private Bag X17, Bellville 7535, Cape Town, South Africa
Telephone :(021) 959 3858/6 Fax: (021) 959 3865
E-mail: ***pkippie@uwc.ac.za*** *or* spenderis@uwc.ac.za

Letter of consent: For Structured Questionnaires

I……., have had the opportunity to ask any questions related to this study, and received satisfactory answers to my questions, and any additional details I wanted.

I agree to take part in this research.

I understand that my participation in this study is voluntary. I am free not to participate and have the right to withdraw from the study at any time, without having to explain myself.

I am aware that this interview might result in research which may be published, but my name may be/ not be used **(circle appropriate)**.

I understand that if I don't want my name to be used that this will be ensured by the researcher.

I may also refuse to answer any questions that I don't want to answer.

Date:……

Participant Name:……

Participant Signature:……

Interviewer name: Rejoice Mabhena

Interviewer Signature:……..

If you have any questions concerning this research, feel free to call (Rejoice Mabhena 0027849977684/ rejoicemabhena01@gmail.com or my supervisor, (Professor Julian May +27 (021) 959 3858 /56/ julian.may120@gmail.com).

APPENDIX C

Interview guide

Private Bag X17, Bellville 7535, Cape Town, South Africa
Telephone :(021) 959 3858/6 Fax: (021) 959 3865
E-mail: pkippie@uwc.ac.za or spenderis@uwc.ac.za

Interview guide

Making sense of poverty statistics in South Africa and beyond: South Africa's statistical readiness in the post-2015 poverty reduction agenda

Section A: Demographic information

Gender:

Age:

Occupation:

Years in current position:

Section B: Poverty measurement in South Africa

i. How has poverty measurement and analysis in South Africa improved in the post-1994 period?

ii. StatsSA uses three poverty lines, why, and how effective have they been in capturing the level and trends of poverty in the country? Why do you say this?

iii. How has StatsSA embraced the multidimensionality of poverty in its surveys? Please give examples of surveys and reports?

1

iv. What further improvements could be made?

v. StatsSA has a number of surveys that have showed the poverty trends in the country in the post-1994 period, how have you embraced the issue of poverty ***dynamics*** in South Africa?

vi. What current activities and future activities are being considered to improve the analysis of poverty dynamics in the country?

vii. Are there plans to improve the measurement of poverty and poverty dynamics in the near future?

viii. How effective has your production of poverty statistics helped in informing policy for poverty reduction in South Africa. Please give examples? How could this be improved?

ix. In the different categories you conduct your surveys, would you consider poverty related surveys to be sufficient to the current demand of such information from users in South Africa? What else is needed?

Section C: Survey conduct in South Africa

i. What is the level of involvement of non-government and international donors in the collection of statistics in your organization?

ii. Does StatsSA require more donor support/involvement in the conduct of poverty related surveys?

iii. Does StatsSA require more technical support/involvement in the conduct of poverty related surveys?

iv. Does StatsSA require more user-group support/involvement in the conduct of poverty related surveys?

v. Does StatsSA have the statistical capacity needed to analyze poverty dynamics? What plans are in place?

Section D: Data quality in the post-2015 development Agenda

i. The Data Revolution[1] is guided by a number of principles including the following;

- Data quality and integrity
- Data disaggregation
- Timeliness
- Transparency and openness
- Usability
- Data governance and independence
- Data resources and capacity

To what extent is StatsSA's strategy adopted for the 2015/16-2019/20 timeframe as discussed in the StatsSA 2015/16 Annual Report, in line with some of these principles?

Section E: Staff statistical capacity, training and development

i. Can you comment on the training and development initiatives that StatsSA has engaged in over the years?

- Probe: The Centre for Regional and Urban Innovation and Statistical Exploration, (CRUISE) Stellenbosch University,
- Postgraduate Diploma in Survey Data Analysis for Development, UCT
- Cote de Ivoire, Statistics degree

ii. What new training and development initiatives are you planning between 2017 and 2030?

iii. Are there any challenges you face in trying to achieve your goals in building staff statistical capacity and development?

Section F: The political economy of African surveys

i. What is your take on the comparability of South Africa's surveys with other countries in the region?

ii. Comment on the issue of regional cooperation and collaboration of statistical offices in Sub-Saharan Africa and South Africa's position?

[1] The data revolution is a recent concept that summarizes the increasing demand of quality data and consequently the need to meet such a demand.

iii. Some analysts suggest that there are poor official statistics in most developing countries? What is your view? Why do you think that this is the case?

iv. In your view what are some of the causes of this 'statistical tragedy'? Do you think poor economies translate to poor data?

v. There is a view[2] that suggests that household surveys in African countries 'systematically exaggerate development progress'. What is your view?

vi. How may other statistical offices in the region improve their statistical capacities?

Section G: From Millennium Development Goals to Sustainable Development Goals

i. How effective were your surveys in capturing the poverty related MDG indicators?

ii. Are there any challenges you faced in providing information on the poverty related indicators for the MDG's.

iii. The SDG's have more indicators for poverty measurement, what is your level of preparedness in capturing these indicators? What actions are you taking?

iv. How are you planning to capture all the required indicators for poverty measurement during the SDG era, (2016-2030)?

v. To what extent did MDG 1, ' Eradicate extreme poverty and hunger' reflect South Africa's poverty reduction objectives during 1990 and 2015?

vi. What is the relevance of SDG 1 to the poverty reduction agenda of South Africa in the next 15 years?

vii. What steps are you following to localize the SDG indicators?

Section H: Challenges faced by African Statistical agencies

[2] Sandefur and Glassman 2014 in "The Political Economy of Bad Data: Evidence from African Survey & Administrative Statistics".

i. What are the some of the challenges faced by Statistical agencies in the region?

Probe: Devarajan (2013)[3] cites the following;

- Weak capacity to collect, manage and disseminate data
- Funding
- Fragmentation in the conduct of surveys

ii. What challenges would you say StatsSA is facing?

iii. Does funding pose a challenge in the conduct of any surveys within your organization?

iv. Are there any surveys you would like to be conducted more frequently, or methodological improvements that you wish to implement but are restricted by lack of funds?

v. In light of the challenges you mentioned what mechanisms are in place to address them both in the country and regionally?

[3] Devarajan, S. 2013 in "Africa's statistical tragedy", Review of income and wealth

APPENDIX D

Birth cohort descriptive statistics

Table Ap1: Distribution of birth cohorts in the 2011 Population Census 10% sample

Number of birth cohorts	Observations
1984-1988	105849
1979-1983	138533
1974-1978	136948
1969-1973	137143
1964-1968	126191
1959-1963	120146
1954-1958	101671
1949-1953	85477
1944-1948	62536
1939-1943	50800
1934-1938	33862
1929-1933	23660
1924-1928	11529

Source: Author's own calculations using t Population Census 10% sample e 2011

Table Ap2: Distribution of birth cohorts by gender

Birth cohorts	2008 MHH	2008 FHH	2010 MHH	2010 FHH	2014 MHH	2014 FHH
1984-1988	591	835	1539	468	649	883
1979-1983	1337	1639	1831	872	904	1093
1974-1978	1882	2035	1617	1129	1064	1026
1969-1973	1851	1889	1568	1222	1249	1116
1964-1968	1513	1593	1364	1117	1156	1184
1959-1963	1499	1454	1151	1137	1213	972
1954-1958	1339	1324	970	1077	1118	871
1949-1953	1130	1089	722	924	885	737
1944-1948	871	807	476	812	836	600
1939-1943	644	619	328	716	722	493
1934-1938	418	347	172	536	438	327
1929-1933	181	201	93	321	305	202
1924-1928	110	89	33	204	146	97

Source: Author's own calculations

Tab AP4: Table AP3: Distribution of birth cohorts by geography

Birth cohorts	Urban 2008	Rural 2008	Urban 2010	Rural 2010	Urban 2010	Rural 2010
1984-1988	709	988	1581	430	420	707
1979-1983	1603	1765	1936	564	620	777
1974-1978	2124	2248	1718	766	737	765
1969-1973	2128	2212	1692	800	786	842
1964-1968	1726	1857	1579	785	769	874
1959-1963	1678	1703	1290	837	868	738
1954-1958	1513	1574	1119	842	769	665
1949-1953	1222	1188	805	797	726	664
1944-1948	959	931	600	673	658	465
1939-1943	686	684	420	689	636	411
1934-1938	450	422	241	507	353	261
1929-1933	235	269	132	268	225	176
1924-1928	157	107	52	147	118	88

Source: Author's own calculations

Table Ap4: Household size for birth cohorts

Year of Birth Cohorts	2008	2010	2014
1984-1988	**2,7**	2,4	2,5
1979-1983	2,7	2,8	2,9
1974-1978	3,1	3,2	3,4
1969-1973	3,7	3,7	3,8
1964-1968	4,1	4,1	3,9
1959-1963	4,3	4,5	4
1954-1958	4,5	4,5	4,2
1949-1953	4,5	4,5	4,3
1944-1948	4,6	4,6	3,9
1939-1943	4,5	4,6	4,1
1934-1938	4,5	4,4	3,7
1929-1933	4,6	4	4,2
1924-1928	4,2	4,3	4,4

Source: Author's own calculations

APPENDIX E

Expenditure patterns of birth cohorts

Table Ap5: Median Expenditure per Capita (Overal)

Birth cohorts	Expenditure 2008	Expenditure 2010	Expenditure 2014
1984-1988	1244	1536	1510
1979-1983	1510	1608	1632
1974-1978	1442	1627	1487
1969-1973	1335	1430	1390
1964-1968	1382	1436	1388
1959-1963	1242	1376	1388
1954-1963	1203	1311	1411
1949-1953	1198	1258	1221
1944-1948	1044	1212	1411
1939-1943	961	1095	1162
1934-1938	920	1136	1324
1929-1933	926	1372	1007
1924-1928	1109	965	1108

Source: Author's own calculations

Table Ap6 Expenditure per capita by race of head

Birth Cohorts	**MHH 2008**	**MHH 2010**	**MHH 2014**	**FHH 2008**	**FHH 2010**	**FHH 2014**
1984-1988	1478	1637	1643	1105	1425	1140
1979-1983	1694	1834	1879	1144	1235	1126
1974-1978	1645	1863	1724	977	1148	1085
1969-1973	1629	1795	1778	966	999	977
1964-1968	1847	1847	1853	920	953	960
1959-1963	1635	1893	1660	852	893	1129
1954-1963	1630	1754	1823	836	887	1029
1949-1953	1559	1858	1632	853	894	996
1944-1948	1305	1786	2106	830	915	1086
1939-1943	1167	1296	1474	806	897	1031
1934-1938	1045	1387	1876	839	962	1197
1929-1933	1146	2138	1292	850	1064	953
1924-1928	1378	1275	1295	1049	870	1030

Source: Author's own calculations

Table Ap7: Expenditure per capita by geography

Birth Cohorts	Urban 2008	Urban 2010	Urban 2014	Rural 2008	Rural 2010	Rural 2014
1984-1988	1679	1969	1866	787	894	835
1979-1983	1791	1980	1975	779	842	823
1974-1978	1745	1956	1862	665	913	801
1969-1973	1724	1894	1822	698	734	705
1964-1968	1895	1933	1871	681	705	716
1959-1963	1752	2031	1949	637	685	671
1954-1963	1795	1896	**2013**	613	699	**656**
1949-1953	1917	2032	1890	612	695	671
1944-1948	1783	2075	2579	593	692	686
1939-1943	1844	2136	2046	567	663	650
1934-1938	1916	2495	2883	536	674	644
1929-1933	2689	3933	1915	596	674	628
1924-1928	3278	3067	2225	573	662	715

Source: Author's own calculations

APPENDIX E

Table Ap8: Regression Results

Source	SS	df	MS		
				Number of obs =	42522
				F(108, 42413) =	208.28
Model	14899.4956	108	137.958293	Prob > F =	0.0000
Residual	28093.0639	42413	.662369177	R-squared =	0.3466
				Adj R-squared =	0.3449
Total	42992.5595	42521	1.01109004	Root MSE =	.81386

LogExpenditure	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
age_cat2	.1902169	.0622201	3.06	0.002	.0682642	.3121696
age_cat3	-.068406	.0620863	-1.10	0.271	-.1900964	.0532844
age_cat4	.2309802	.0829402	2.78	0.005	.0684157	.3935446
age_cat5	-.0369958	.0870486	-0.43	0.671	-.2076129	.1336213
age_cat6	.2989604	.096224	3.11	0.002	.1103593	.4875614
age_cat7	-.06367	.106018	-0.60	0.548	-.2714674	.1441274
age_cat8	.3018263	.1059832	2.85	0.004	.0940971	.5095555
age_cat9	-.1348751	.1238085	-1.09	0.276	-.3775422	.107792
age_cat10	.3896436	.1140486	3.42	0.001	.166106	.6131812
age_cat11	-.1311653	.1386118	-0.95	0.344	-.4028471	.1405165
age_cat12	.3916848	.1202492	3.26	0.001	.1559938	.6273757
age_cat13	-.1186748	.1524079	-0.78	0.436	-.4173974	.1800477
age_cat14	.4150711	.124904	3.32	0.001	.1702567	.6598854
age_cat15	-.1154938	.1658139	-0.70	0.486	-.4404923	.2095047
age_cat16	.4445399	.1290256	3.45	0.001	.1916471	.6974326
age_cat17	-.1191923	.17891	-0.67	0.505	-.4698595	.2314748
age_cat18	.4592315	.1320784	3.48	0.001	.2003553	.7181077
age_cat19	-.1782575	.1915212	-0.93	0.352	-.5536428	.1971278
age_cat20	.4458754	.13505	3.30	0.001	.1811747	.7105761
age_cat21	-.1653324	.2035731	-0.81	0.417	-.5643398	.233675
age_cat22	.5022352	.137226	3.66	0.000	.2332694	.771201
age_cat23	-.1092808	.2157768	-0.51	0.613	-.5322076	.313646
age_cat24	.4370355	.1388283	3.15	0.002	.1649292	.7091418
age_cat25	-.0908774	.2279724	-0.40	0.690	-.5377078	.3559529
age_cat26	.4081822	.1400077	2.92	0.004	.1337644	.6826
age_cat27	-.0254698	.2403178	-0.11	0.916	-.4964975	.445558
age_cat28	.4323652	.1406757	3.07	0.002	.1566381	.7080924
age_cat29	.0045467	.2528281	0.02	0.986	-.4910014	.5000947
age_cat30	.3872514	.1406158	2.75	0.006	.1116417	.6628611
age_cat31	-.0341512	.265494	-0.13	0.898	-.5545247	.4862223
age_cat32	.4445186	.1396213	3.18	0.001	.170858	.7181792
age_cat33	-.1183702	.27861	-0.42	0.671	-.6644513	.4277109
age_cat34	.4374777	.1377202	3.18	0.001	.1675435	.707412
age_cat35	-.1945512	.2918639	-0.67	0.505	-.7666103	.377508
age_cat36	.39265	.1345159	2.92	0.004	.1289961	.6563039

Source	SS	df	MS		
				Number of obs =	42522
				F(108, 42413) =	208.28
Model	14899.4956	108	137.958293	Prob > F =	0.0000
Residual	28093.0639	42413	.662369177	R-squared =	0.3466
				Adj R-squared =	0.3449
Total	42992.5595	42521	1.01109004	Root MSE =	.81386

LogExpenditure	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
age_cat2	.1902169	.0622201	3.06	0.002	.0682642	.3121696
age_cat3	-.068406	.0620863	-1.10	0.271	-.1900964	.0532844
age_cat4	.2309802	.0829402	2.78	0.005	.0684157	.3935446
age_cat5	-.0369958	.0870486	-0.43	0.671	-.2076129	.1336213
age_cat6	.2989604	.096224	3.11	0.002	.1103593	.4875614
age_cat7	-.06367	.106018	-0.60	0.548	-.2714674	.1441274
age_cat8	.3018263	.1059832	2.85	0.004	.0940971	.5095555
age_cat9	-.1348751	.1238085	-1.09	0.276	-.3775422	.107792
age_cat10	.3896436	.1140486	3.42	0.001	.166106	.6131812
age_cat11	-.1311653	.1386118	-0.95	0.344	-.4028471	.1405165
age_cat12	.3916848	.1202492	3.26	0.001	.1559938	.6273757
age_cat13	-.1186748	.1524079	-0.78	0.436	-.4173974	.1800477
age_cat14	.4150711	.124904	3.32	0.001	.1702567	.6598854
age_cat15	-.1154938	.1658139	-0.70	0.486	-.4404923	.2095047
age_cat16	.4445399	.1290256	3.45	0.001	.1916471	.6974326
age_cat17	-.1191923	.17891	-0.67	0.505	-.4698595	.2314748
age_cat18	.4592315	.1320784	3.48	0.001	.2003553	.7181077
age_cat19	-.1782575	.1915212	-0.93	0.352	-.5536428	.1971278
age_cat20	.4458754	.13505	3.30	0.001	.1811747	.7105761
age_cat21	-.1653324	.2035731	-0.81	0.417	-.5643398	.233675
age_cat22	.5022352	.137226	3.66	0.000	.2332694	.771201
age_cat23	-.1092808	.2157768	-0.51	0.613	-.5322076	.313646
age_cat24	.4370355	.1388283	3.15	0.002	.1649292	.7091418
age_cat25	-.0908774	.2279724	-0.40	0.690	-.5377078	.3559529
age_cat26	.4081822	.1400077	2.92	0.004	.1337644	.6826
age_cat27	-.0254698	.2403178	-0.11	0.916	-.4964975	.445558
age_cat28	.4323652	.1406757	3.07	0.002	.1566381	.7080924
age_cat29	.0045467	.2528281	0.02	0.986	-.4910014	.5000947
age_cat30	.3872514	.1406158	2.75	0.006	.1116417	.6628611
age_cat31	-.0341512	.265494	-0.13	0.898	-.5545247	.4862223
age_cat32	.4445186	.1396213	3.18	0.001	.170858	.7181792
age_cat33	-.1183702	.27861	-0.42	0.671	-.6644513	.4277109
age_cat34	.4374777	.1377202	3.18	0.001	.1675435	.707412
age_cat35	-.1945512	.2918639	-0.67	0.505	-.7666103	.377508
age_cat36	.39265	.1345159	2.92	0.004	.1289961	.6563039

age_cat37	-.174004	.3051032	-0.57	0.568	-.7720124	.4240045
age_cat38	.4369934	.1298707	3.36	0.001	.1824442	.6915427
age_cat39	-.1550593	.3187786	-0.49	0.627	-.7798716	.4697531
age_cat40	.4334341	.1219243	3.55	0.000	.1944599	.6724082
age_cat41	-.3072404	.332809	-0.92	0.356	-.9595528	.3450719
age_cat42	.3817577	.1087097	3.51	0.000	.1686846	.5948308
age_cat43	-.2671707	.3467975	-0.77	0.441	-.9469007	.4125593
age_cat44	.3069465	.0902077	3.40	0.001	.1301376	.4837555
age_cat45	-.3397272	.3622776	-0.94	0.348	-1.049799	.3703441
age_cat46	.2657854	.0666126	3.99	0.000	.1352234	.3963475
age_cat47	-.338334	.3768797	-0.90	0.369	-1.077026	.4003577
age_cat48	.3591665	.0733436	4.90	0.000	.2154117	.5029214
age_cat49	0 (omitted)					
age_cat50	0 (omitted)					
age_cat51	0 (omitted)					
year_cat2	.1266089	.0114702	11.04	0.000	.1041271	.1490907
year_cat3	0 (omitted)					
cohort_cat2	.6249745	.3761651	1.66	0.097	-.1123165	1.362266
cohort_cat3	.0251153	.0649071	0.39	0.699	-.1021039	.1523346
cohort_cat4	.6268998	.3612089	1.74	0.083	-.081077	1.334876
cohort_cat5	.0286715	.0903976	0.32	0.751	-.1485096	.2058526
cohort_cat6	.6713975	.3467763	1.94	0.053	-.008291	1.351086
cohort_cat7	.0172569	.1072257	0.16	0.872	-.1929077	.2274215
cohort_cat8	.6130113	.3328994	1.84	0.066	-.0394782	1.265501
cohort_cat9	-.0362061	.1185806	-0.31	0.760	-.2686264	.1962142
cohort_cat10	.5523269	.31938	1.73	0.084	-.0736644	1.178318
cohort_cat11	-.0198011	.1252576	-0.16	0.874	-.2653084	.2257062
cohort_cat12	.6261273	.3056869	2.05	0.041	.0269749	1.22528
cohort_cat13	-.0201005	.1294108	-0.16	0.877	-.2737483	.2335472
cohort_cat14	.5520075	.2927598	1.89	0.059	-.0218075	1.125823
cohort_cat15	.0065336	.1320852	0.05	0.961	-.2523561	.2654233
cohort_cat16	.533425	.2794206	1.91	0.056	-.014245	1.081095
cohort_cat17	.0118848	.1335858	0.09	0.929	-.249946	.2737157
cohort_cat18	.4556523	.2668141	1.71	0.088	-.0673087	.9786133
cohort_cat19	.0252019	.134055	0.19	0.851	-.2375485	.2879523
cohort_cat20	.4496135	.2545136	1.77	0.077	-.0492382	.9484653
cohort_cat21	.0062625	.1336029	0.05	0.963	-.255602	.2681269
cohort_cat22	.5486554	.2418038	2.27	0.023	.0747151	1.022596
cohort_cat23	-.0246652	.1325128	-0.19	0.852	-.2843928	.2350625
cohort_cat24	.5584469	.2301196	2.43	0.015	.1074078	1.009486
cohort_cat25	-.005962	.1309165	-0.05	0.964	-.262561	.250637
cohort_cat26	.6656396	.2178273	3.06	0.002	.2386936	1.092585
cohort_cat27	-.0076098	.1287893	-0.06	0.953	-.2600395	.2448198
cohort_cat28	.663477	.2059742	3.22	0.001	.2597634	1.067191
cohort_cat29	.0504767	.1259052	0.40	0.688	-.1963001	.2972534

cohort_cat30		.5798361	.1939262	2.99	0.003	.1997368	.9599353
cohort_cat31		-.0047457	.1226328	-0.04	0.969	-.2451085	.2356171
cohort_cat32		.5362747	.1811324	2.96	0.003	.1812517	.8912977
cohort_cat33		.0169074	.1186615	0.14	0.887	-.2156714	.2494862
cohort_cat34		.502728	.168358	2.99	0.003	.1727429	.8327131
cohort_cat35		-.040933	.1140103	-0.36	0.720	-.2643955	.1825295
cohort_cat36		.4896528	.1548838	3.16	0.002	.1860774	.7932282
cohort_cat37		-.02778	.1079832	-0.26	0.797	-.2394293	.1838693
cohort_cat38		.4545535	.1416937	3.21	0.001	.176831	.732276
cohort_cat39		-.0342493	.0998619	-0.34	0.732	-.2299807	.1614821
cohort_cat40		.357513	.1261363	2.83	0.005	.1102833	.6047428
cohort_cat41		-.0030345	.0902329	-0.03	0.973	-.1798928	.1738237
cohort_cat42		.2948436	.1087417	2.71	0.007	.0817077	.5079794
cohort_cat43		-.0401793	.0774155	-0.52	0.604	-.1919153	.1115566
cohort_cat44		.2655794	.0900021	2.95	0.003	.0891735	.4419853
cohort_cat45		-.0104861	.0585669	-0.18	0.858	-.1252784	.1043063
cohort_cat46		.2477194	.0656634	3.77	0.000	.1190178	.3764209
cohort_cat47		0 (omitted)					
cohort_cat48		0 (omitted)					
cohort_cat49		0 (omitted)					
cohort_cat50		0 (omitted)					
Settlementtype							
	2	-.6133541	.0154949	-39.58	0.000	-.6437244	-.5829839
	4	-.4186137	.0112288	-37.28	0.000	-.4406224	-.396605
	5	-.6159718	.0204826	-30.07	0.000	-.656118	-.5758256
Province							
	2	-.2192742	.0182503	-12.01	0.000	-.2550452	-.1835032
	3	-.3350019	.0208599	-16.06	0.000	-.3758877	-.294116
	4	-.2949971	.0196183	-15.04	0.000	-.3334494	-.2565449
	5	-.0947363	.0184652	-5.13	0.000	-.1309284	-.0585441
	6	-.1564672	.019863	-7.88	0.000	-.1953991	-.1175353
	7	.0242298	.0172165	1.41	0.159	-.0095149	.0579746
	8	-.0793419	.0199824	-3.97	0.000	-.1185077	-.0401761
	9	-.1446663	.020109	-7.19	0.000	-.1840803	-.1052523
2.Genderofhead		-.1578712	.0082416	-19.16	0.000	-.1740249	-.1417174
Raceofhead							
	2	.3061576	.0163571	18.72	0.000	.2740973	.3382179
	3	.9198849	.0307539	29.91	0.000	.8596066	.9801632
	4	1.513922	.016124	93.89	0.000	1.482319	1.545526
_cons		8.147297	.0494618	164.72	0.000	8.050351	8.244243

APPENDIX F

Poverty estimates for birth cohorts

Table Ap 9: Poverty incidences for all birth cohorts

Birth Cohorts	2008			2010			2014		
	P0	**P1**	**P2**	**P0**	**P1**	**P2**	**P0**	**P1**	**P2**
1984-1988	**22,24**	7,95	3,89	**19,21**	6,25	2,8	**19,73**	6,91	3,4
1979-1983	**20,29**	6,15	2,66	**16,54**	5,52	2,68	**18,8**	6,41	3,03
1974-1978	**21,14**	7,14	3,32	**17,8**	5,89	2,7	**19,55**	7,05	3,48
1969-1973	**22,53**	7,81	3,61	**21,8**	7,51	3,55	**24,53**	9,14	4,66
1964-1968	**25,56**	9,07	4,38	**21,65**	7,81	3,85	**24,02**	8,92	4,58
1959-1963	**25,72**	9,57	4,83	**24,35**	8,75	4,35	**24,4**	8,93	4,46
1954-1963	**25,88**	9,61	4,78	**24,9**	9,26	4,57	**25,64**	9,37	4,8
1949-1953	**28,42**	10,72	5,36	**25,25**	9,25	4,53	**26,85**	10,44	5,39
1944-1948	**29,97**	11,25	5,5	**25,64**	9,07	4,43	**24,17**	8,61	4,25
1939-1943	**34,14**	12,13	5,92	**29,51**	10,05	4,65	**27,69**	10,21	5,09
1934-1938	**35,03**	12,62	6,02	**27,27**	9,68	4,58	**26,46**	10,5	5,51
1929-1933	**32,94**	11,19	5,3	**23,97**	8,17	3,77	**30,42**	11,98	6,03
1924-1928	**29,54**	10,02	4,59	**31,29**	11,17	5,61	**28,66**	9,15	4,16

Source: Author's own calculations

Table Ap10: Poverty incidences for MHH

Birth Cohorts	P0	P1	P2	P0	P1	P2	P0	P1	P2
1984-1988	**15,22**	5,02	2,29	**15,99**	5,21	2,48	**14,92**	4,99	2,32
1979-1983	**15,42**	4,52	1,97	**13,06**	4,4	2,16	**13,98**	4,36	1,93
1974-1978	**16,09**	5,08	2,28	**14,13**	4,58	2,12	**15,18**	5	2,36
1969-1973	**18,5**	5,98	2,62	**16,53**	5,35	2,46	**17,81**	6,11	2,96
1964-1968	**18,25**	5,95	2,75	**15,63**	5,8	3,01	**17,23**	6,18	3,1
1959-1963	**18,5**	6,76	3,38	**16,6**	5,83	2,92	**19,04**	6,83	3,41
1954-1963	**19,56**	7,11	3,55	**18,13**	6,69	3,32	**21,19**	7,39	3,7
1949-1953	**22,33**	8,02	3,81	**19,32**	7,05	3,47	**22,12**	8,4	4,3
1944-1948	**25,71**	9,56	4,74	**19,92**	7,17	3,52	**18,29**	6,47	3,19
1939-1943	**30,67**	10,94	5,4	**25,78**	8,31	3,78	**23,47**	9,25	4,98
1934-1938	**34,65**	12,84	6,38	**23,11**	7,84	3,55	**22,45**	7,94	4,13
1929-1933	**30,53**	9,56	4,2	**16,91**	6,29	3,1	**27,54**	10,15	4,67
1924-1928	**25,4**	8,31	3,7	**29,89**	8,97	4,33	**32,69**	9,27	3,7

Source: Author's own calculations

Table AP11: Poverty incidences for FHH

Birth Cohorts	P0	P1	P2	P0	P1	P2	P0	P1	P2
1984-1988	27,61	9,97	4,93	24,13	7,86	3,74	29,34	10,75	5,55
1979-1983	29,58	9,33	4,03	24,05	7,94	3,81	28,09	10,36	5,14
1974-1978	30,91	11,16	5,38	26,14	8,88	4,03	27,47	10,77	5,5
1969-1973	30,21	11,21	5,43	31,27	11,4	5,5	35,37	14,02	7,41
1964-1968	37,71	14,26	7,1	31,54	11,1	5,23	33,01	12,55	6,55
1959-1963	37,53	14,18	7,21	35,43	12,92	6,4	31,69	11,79	5,9
1954-1963	35,69	13,49	6,69	34,44	12,87	6,32	31,33	11,91	6,21
1949-1953	38,05	14,95	7,77	33,94	12,47	6,09	32,17	12,74	6,61
1944-1948	35,47	13,43	6,49	32,22	11,26	5,49	29,54	10,55	5,23
1939-1943	38,07	13,5	6,53	33,31	11,83	5,55	30,91	10,94	5,17
1934-1938	35,99	12,63	5,8	31,18	11,42	5,55	28,88	12,05	6,34
1929-1933	34,66	12,34	6,06	29,51	9,65	4,3	31,96	12,97	6,76
1924-1928	32,23	11,13	5,16	32,31	12,76	6,54	27,06	9,1	4,34

Source: Author's own calculations

Table Ap12: Poverty incidence for birth cohorts in urban areas

Birth Cohorts	P0	P1	P2	P0	P1	P2	P0	P1	P2
1984-1988	14,33	4,82	2,32	14,32	4,32	1,94	14,03	4,57	2,12
1979-1983	14,66	4,08	1,63	10,67	3,31	1,52	13,02	3,93	1,7
1974-1978	13,86	4,18	1,79	12,45	4	1,73	13,69	4,54	2,09
1969-1973	15,77	5,5	2,4	15,06	4,66	2,04	17,22	6	2,93
1964-1968	17,74	5,96	2,77	14,07	4,67	2,23	16,16	5,36	2,54
1959-1963	16,3	5,61	2,67	14,31	4,86	2,37	15,29	4,72	2,03
1954-1963	15,6	5,12	2,36	16,2	5,72	2,68	17,07	5,34	2,49
1949-1953	17,27	5,87	2,77	15,27	5,24	2,41	15,31	5,78	3,05
1944-1948	18,25	6,83	3,38	14,33	4,39	2,02	12,33	3,65	1,57
1939-1943	16,73	5	2,16	16,52	4,66	1,86	13,65	4,39	1,97
1934-1938	17,55	5,84	2,,70	13,55	4,48	1,99	10,8	3,8	1,75

1929-1933	16,87	5,16	2,31	10,79	3,15	1,32	12,98	4,89	2,4
1924-1928	11,48	3,55	1,42	13,47	3,7	1,91	15,2	3,44	1,13

Source: Author's own calculations

Table AP13: Poverty incidence for birth cohorts in rural areas

Birth Cohorts	P0	P1	P2	P0	P1	P2	P0	P1	P2
1984-1988	37.19	15,06	7,36	30.41	11,72	5,97	36.48	14,54	7,63
1979-1983	32.29	14,32	6,85	31.95	13,23	6,67	37.30	16,05	8,44
1974-1978	42.57	17,92	9,02	33.54	11,63	5,71	36.53	16,69	8,82
1969-1973	42.34	16,87	8,08	40.24	16,01	8,1	43.75	18,42	9,76
1964-1968	45.11	18,09	9,22	39.72	16,85	8,52	44.79	18,42	9,97
1959-1963	47.01	19,86	10,56	44.93	17,5	8,75	46.13	19,98	10,96
1954-1963	48.00	20,52	10,67	43.07	17,46	8,96	46.36	20,05	10,83
1949-1953	48.87	20,7	10,78	42.30	17,42	8,91	47.89	18,93	9,72
1944-1948	49.47	19,65	9,51	42.60	16,87	8,39	46.84	18,26	9,55
1939-1943	56.33	21,64	10,92	44.91	16,87	8,13	49.37	18,66	9,42
1934-1938	54.41	21,43	10,42	45.71	16,63	8,13	48.25	20,27	10,97
1929-1933	5138	18,39	8,81	41.67	15,64	7,41	50.86	20,4	10,31
1924-1928	53.55	19,77	9,43	48.82	17,84	8,91	49.94	14,64	7,06

Source: Author's own calculations

Table AP14: Unpaired ttest results for 2008, 2010 and 2014 expenditure for all birth cohorts

Obs/cohorts	Mean expenditure for all birth cohorts in 2010	Mean expenditure for all birth cohorts in 2008	Difference	t	df	p-value
13	3783.54	3175,23	608,31	3.5491	24	0.0016
Obs/cohorts	**Mean expenditure for all birth cohorts in 2014**	**Mean expenditure for all birth cohorts in 2010**				
13	3501.85	3783.54	-281.69	-130	24	0.2053
Obs/cohorts	**Mean expenditure for FHH birth cohorts in 2010**	**Mean expenditure for MHH birth cohorts in 2008**				
13	2710.31	3683.92	-973.62	-5.22	24	0.0000
Obs/cohorts	**Mean expenditure for FHH birth cohorts in 2014**	**Mean expenditure for FHH birth cohorts in 2010**				
13	2868.31	4446.46	1377.85	9.67	24	0.0000
Obs/cohorts	**Mean expenditure for Rural birth cohorts in 2010**	**Mean expenditure for Urban birth cohorts in 2008**				
13	1490.46	2868.31	1377.85	9.67	24	0.0000
Obs/cohorts	**Mean expenditure for Urban birth cohorts in 2014**	**Mean expenditure for Rural birth cohorts in 2010**				
13	1396.69	5202.85	3806.15	10.84	24	0.000
Obs/cohorts	**Mean expenditure for Coloured birth cohorts in 2010**	**Mean expenditure for African birth cohorts in 2008**				
13	3011.23	1563.39	-1447.85	-6.63	24	0.0000
Obs/cohorts	**Mean expenditure for Coloured birth cohorts in 2014**	**Mean expenditure for African birth cohorts in 2010**				
13	3033.62	1791.39	-815.58	-6.00	24	0.000
Obs/cohorts	**Mean expenditure for**	**Mean expenditure for**				

	Indian birth cohorts in 2010	**Black birth cohorts in 2008**				
13	6915.85	1563.39	-5352.46	-6.77	24	0.0000
Obs/cohorts	**Mean expenditure for Indian birth cohorts in 2014**	**Mean expenditure for African birth cohorts in 2010**				
13	5529.69	1791.39	-3738.31	-7.87	24	0.000
Obs/cohorts	**Mean expenditure for White birth cohorts in 2010**	**Mean expenditure for African birth cohorts in 2008**				
13	12 889	1563.39	-11325	-16.35	24	0.000
Obs/cohorts	**Mean expenditure for White birth cohorts in 2014**	**Mean expenditure for African birth cohorts in 2010**				
13	12 765	1791.39	-10128.65	-26.80	24	0.0000

Author's own calculations