The lived experience of inequality in post-apartheid South Africa: measuring exposure to socio-economic inequality at small area level

David James McLennan

Green Templeton College

University of Oxford

DPhil Social Policy

Table of Contents

List of Tables

List of Figures

Abstract

Thesis title	The lived experience of inequality in post-apartheid South Africa: measuring exposure to socio-economic inequality at small area level.
Name	David James McLennan
College	Green Templeton College
Degree	DPhil Social Policy
Year/term of submission	Hilary Term 2016

South Africa has undergone a remarkable political transformation since the birth of democracy in 1994, yet it remains plagued by extremely high levels of socio-economic inequality, violent crime and social unrest. Although inequality is often regarded as a major driver of many social problems, the evidence base concerning inequality in South Africa is relatively limited, consisting primarily of national level Gini coefficients or General Entropy measures based upon household income, expenditure or consumption data. In this thesis I argue that these broad national level measures say little about people's actual day-to-day lived experiences of inequality and how these individual experiences of inequality may be shaped by the local geographical areas in which people live and go about their daily lives. I construct a series of empirical measures of exposure to socio-economic inequality which reflect the socio-spatial environments in which people live. I argue that these new measures can be used as explanatory factors in the study of other social outcomes, both at an individual level (for example, individuals'

attitudes) and at an area level (for example, rates of violent crime). Exposure to inequality is measured both from the perspective of the ‘poor’ population and the perspective of the ‘non-poor’ population and the measures are constructed and presented at small area level using the Datazone statistical geography. I analyse the spatial distribution of exposure to inequality and find that exposure to inequality is typically highest in urban neighbourhoods, particularly in the major metropolitan areas. I develop a measure of intensity of exposure in order to highlight areas with both high exposure and high levels of deprivation. I also present one example of how my new measures can be used to explore associations with other outcomes, specifically looking at the relationship between people’s lived experience of inequality and their attitudes towards inequality and redress.

Acknowledgements

I would like to thank my transfer of status examiners – Mr George Smith and Dr Martin Evans – and my confirmation of status examiners – Mr George Smith and Professor Erzsébet Bukodi – for their helpful suggestions on each occasion.

Although the research presented in this thesis is entirely my own work, it is based upon data and methodologies emerging from a research project funded by the UK Economic and Social Research Council (ESRC) and the South African National Research Foundation (NRF). As detailed in Appendix A, this ESRC/NRF project was a collaborative undertaking between researchers from the Centre for the Analysis of South African Social Policy (CASASP) at the University of Oxford and the Human Sciences Research Council (HSRC) in South Africa. I was the Principal Investigator on the project and I am very grateful to the other members of the project team for their respective contributions which helped me to shape my own conceptual and methodological approaches which are reflected in this thesis. I am particularly grateful to Ben Roberts at the HSRC for his insights concerning social attitudes research in South Africa and for permitting me to utilise the South African Social Attitudes Survey for this thesis. Thanks also to my colleague Helen Barnes for her advice and support across a range of research projects and on my thesis.

I would like to thank Dr Dan Lunn and Dr George Nicholson from the Department of Statistics at the University of Oxford for their invaluable advice and support on the statistical modelling presented in Chapter 7 of this thesis, including their help with rapidly learning the R statistical software package. I am also grateful to David Avenell for his advice on ArcGIS matters.

I am extremely grateful for all the support I have received from my family, without which I would not have been able to complete this thesis. I would particularly like to thank my wife, Katharine, and my three young children for their endless encouragement and for tolerating my absences during the final write-up phase. I would also like to thank my parents and my brother for their support and I would like to thank my late-grandparents for their generosity which enabled me to take some time away from work each week during the final phase of the thesis.

My greatest thanks are to Professor Michael Noble and Dr Gemma Wright for their unwavering support and encouragement, both as colleagues of 14 years and as doctoral supervisors. As colleagues, first at the University of Oxford and now at Southern African Social Policy Research Insights (SASPRI), we have undertaken an incredible breadth of research projects which has enabled me to develop the skill set upon which this thesis is based. I would like to thank Michael and Gemma for introducing me to South Africa and for giving me an invaluable insight into the multitude of economic and social challenges that the country faces, which has undoubtedly helped me to shape the focus of this thesis. I would also like to thank Michael and Gemma for the clarity of their advice and guidance as doctoral supervisors and for their continual reassurance that this thesis was both achievable and academically valuable. I am also very grateful for their help in juggling the responsibilities of being an employee with the responsibilities of being a doctoral student, especially during the final write-up stage.

Chapter 1: Introduction

1.1 Why examine ‘the lived experience of inequality’?

Despite South Africa's classification as an 'upper middle income' economy by the World Bank, significant proportions of the population continue to be characterised by high levels of poverty, unemployment, poor health, criminal victimisation, and lack of services. The country's colonial, segregationist and apartheid history prior to 1994 produced a highly unequal distribution of resources, skills and opportunities between different racial groups. There was also a strong spatial dimension to discriminatory policies of the pre-1994 period, with citizenship and land ownership rights determined by one's designated population group[1]. Consequently, not only does South Africa rank among the world's most unequal societies, but clear spatial clustering of deprived groups has produced visibly high levels of spatial inequality (Noble et al., 2009a).

Although census and household survey data have demonstrated that incomes have been growing for all race groups in the country since 1994, income inequality has remained stubbornly high. In fact, aggregate inequality estimates have shown that the post-apartheid period has been characterised by a moderate *rise* in inequality with income increasingly concentrated in the richest income decile (Leibbrandt et al., 2012; Leibbrandt et al., 2010).

As elsewhere internationally, persistent inequality, rather than poverty *per se*, is increasingly regarded as a major correlate of various social ills in South Africa

[1] Statistics South Africa defines ‘population group’ as “A group with common characteristics (in terms of descent and history), particularly in relation to how they were (or would have been) classified before the 1994 elections. The following categories are provided in the census: black African, coloured, Indian or Asian, white, other.” (Statistics South Africa, 2004, p.12)

(Kaufmann D et al., 2010; Pickett and Wilkinson, 2015; Wilkinson and Pickett, 2010). This is particularly the case in respect of service delivery protests and crime. For example, the escalating service delivery protests, particularly in urban informal settlements, have been attributed to, among other things, socio-economic inequalities (Alexander, 2010; Botes, 2007). However, this assertion has only limited empirical validation (Noble and Wright, 2013b). Similarly, notwithstanding international evidence that supports the general contention that inequality is associated with crime (e.g. Bourguignon, 1999; UNODC, 2011), there is very little empirical evidence from South Africa, the main exception being a study that demonstrated a relationship between income inequality and burglary/vehicle theft using police recorded crime data from 1996 (Demombynes and Ozler, 2006).

The lack of empirical evidence linking inequality to service delivery protests and crime should not be read as rejecting the relationship, but rather it highlights the lack of adequate data on what one might term the 'lived experience of inequality'. The vast majority of quantitative research concerning inequality in South Africa has utilised the 'classical' measures of (income) inequality such as the Gini coefficient, General Entropy measures and the Atkinson Index, expressed at national level. Whilst these classical measures are powerful tools for assessing change in South African inequality over time or for comparing South Africa with other countries worldwide, they say little about people's day-to-day lived experience of inequality.

The mechanisms through which experiences of societal inequality may shape attitudes and subsequently encourage or discourage particular forms of action, whether this be involvement in violent crime, participation in (peaceful or non-peaceful) public protests, or other forms of demonstration, are indeed complex. It has been argued that people's experience of inequality plays an important role in shaping their beliefs as to the sort of society in which they wish to live, and the means through which they aim to achieve that goal (Wikstrom, 2006b).

Research has shown that people may experience inequality in a variety of ways and settings (Kingdon and Knight, 2007; Posel and Casale, 2011). For instance, inequality may be apparent between different broad sectors of the population, such as between gender groups, racial groups or age groups. It may be apparent at a social network level, such as between different members of the same household, immediate family, extended family or friendship network. It may be apparent spatially, for instance between residents of different cities, town, villages, or neighbourhoods. Indeed, it may even be apparent between different streets within a neighbourhood. In a globalising world characterised by mass media, inequality may also be apparent at a much more macro level, for instance between different provinces, countries or even continents.

Whilst I acknowledge the multitude of ways in which individuals may potentially experience inequality, I argue here that a common feature of all these social dynamics is that they occur (or may potentially occur) within geographical places. In terms of routine experiences of inequality, I contend that place matters, and that, more specifically, the *neighbourhoods* in which people live, work, socialise etc matter. With regards to the process through which people experience inequality, I argue that an individual's first-hand experience of inequality may be shaped both by the degree to which they come into personal contact with those from the opposite ends of the socio-economic spectrum, and also by the stark visual signs of inequality that do not involve direct personal contact. For example, a rich person driving along a highway may observe the highly deprived informal settlements without actually coming into personal contact with the residents of those settlements. Similarly, a poor person may travel through or spend time in affluent neighbourhoods whilst seeking work and thus observe the very obvious visual signs of affluence, but with little or no personal interaction with the affluent residents of those neighbourhoods. People's first-hand experience of inequality is therefore contoured by the geographical settings in which they live, work, socialise and travel.

To illustrate my argument concerning the importance of *place* in shaping people's lived experiences of inequality, Figure 1.1 and Figure 1.2 show two contrasting geographical areas within South Africa. Figure 1.1 shows the view across the highly deprived Alexandra township towards the affluent suburbs of Sandton in the distance. These two parts of Johannesburg sit geographically side-by-side, separated only by the M1 highway. Due to the close spatial proximity between Alexandra and Sandton, it seems likely that residents of these two suburbs are well aware of the high degree of inequality between different sections of the South African population.

Figure 1.1: Alexandra township with Sandton in the distance

In contrast, Figure 1.2 shows a typical view within the former Transkei homeland in the Eastern Cape. As I discuss in detail later in this thesis, areas such as the former Transkei are characterised by extremely high (and almost uniformly high) rates of deprivation across considerable geographical expanses. The disparities in deprivation levels between neighbouring areas are typically much smaller in areas such as the former Transkei than in areas such as Johannesburg. In light of this, one might imagine that the lived experience of *inequality* may be somewhat less pronounced for residents of the former Transkei than for residents of Alexandra township (whilst residents of both areas may experience similar levels of *poverty and deprivation*).

Figure 1.2: The former Transkei homeland, Eastern Cape

Despite the strong images of spatial inequality conveyed visually by Figure 1.1 and Figure 1.2, until now, there has been no neighbourhood level measure reflecting people's experience of inequality developed for South Africa.

In this thesis I seek to address the gap in the existing evidence base by developing a new empirical measure that reflects people's lived experience of inequality as they go about their daily lives. Furthermore, I use a case study to test how the measure might be used as an explanatory variable in social enquiry. As an example, I test for statistical associations between this new measure of inequality and key attitudinal outcomes concerning inequality and options for redress. As such, I contend that this thesis makes an important contribution to the literature on spatial inequality, attitudes to inequality, and the linkage between the two.

The research comprising this thesis draws upon elements of work undertaken as part of a Pathfinder research grant funded jointly by the Economic and Social Research Council (ESRC) in the UK and the National Research Foundation (NRF) in South Africa. The project, titled 'Exploring the relationship between spatial inequality and attitudes to inequality in South Africa', was a collaborative undertaking between the Centre for the Analysis of South African Social Policy (CASASP) at the University of Oxford and the Human Sciences Research Council (HSRC) in South Africa. I was principal investigator on the grant. I detail my input into the project, and the respective inputs of the other four members of the project team, in Appendix A. All the research presented in this thesis is my own work.

In the remainder of this introductory chapter I first discuss the historical developments that have contributed to the contemporary patterns of inequality in South Africa and I then set out my research questions and thesis structure.

1.2 Historical context

The current unequal distribution of poverty and deprivation in South Africa continues to reflect the outcomes of the colonial (from 1652 to 1910), segregationist (from 1910 to 1948) and apartheid (from 1948 to 1994) political regimes. Population group, or race, was the discriminating factor in the policies of segregation, with employment opportunities, access to services, entitlement to land ownership and many other opportunities and rights determined by one's ascribed population group (Lester et al., 2000). The distinctly spatial nature of the racially discriminatory policies prior to democracy in 1994 meant that a unique geographical pattern of deprivation developed which is, in many cases, still evident today. In order to understand the unequal distribution of deprivation between racial groups and between geographical areas in present day South Africa it is first necessary to understand the social and political forces that shaped the country prior to democracy in 1994 and then subsequently from 1994 to the present.

Colonial, segregationist and apartheid periods: 1652-1994

The primary goal of the various political regimes from the beginning of colonial rule in the mid seventeenth century to the end of apartheid in late twentieth century was to ensure white minority rule and dominance. Achievement of this goal was dependent upon two key factors: first, the prevention of racial mixing to ensure the preservation of a distinct white population group, and; second, the establishment of power structures whereby the white population could exploit the other population groups without assuming any responsibility for their welfare.

According to Terreblanche, the colonial powers and white colonists enriched themselves, mostly at the cost of indigenous people, in three main ways:

> "firstly, by creating political and economic power structures that put them in a privileged position vis-à-vis the indigenous population groups; secondly, by depriving indigenous people of land, surface water, and cattle; and, thirdly, by reducing slaves and indigenous people to different forms of unfree and exploitable labour." (Terreblanche, 2002, p.6).

The importance of land, and therefore the spatial distribution of people, was a crucial element in this exploitation. Enactments such as the Natives Land Act (1913), the Native Trust and Land Act (1936) and the Natives (Urban Areas) Act (1923) all served to impose a structural geography of segregation upon the country's population. Disparities between population groups were evident in almost every socioeconomic indicator. For example, in 1946, just prior to the National Party assuming power, the White one-fifth of the population controlled two-thirds of personal income (Christopher, 1994).

With the rise to power of the National Party in 1948, the enactments became intensified through the policies of apartheid. As Wolpe states:

> "The National Party came to power on a policy aimed at suppressing the emergent black opposition which threatened the reproduction of white domination, that is, threatened the conditions which would enable the regime to meet, inter alia, the demands of the white farmers and protect the interests of the white working class" (Wolpe, 1988, p.6).

The Population Registration Act (1950), the Group Areas Act (1950), the Bantu Authorities Act 1951 and the Promotion of Bantu Self-Regulation Act (1959) all sought to stipulate and control the rights of residence and movement of the different ethnic groups in order to maximise the benefits to the minority white population (Beinart, 1994; Terreblanche, 2002). The apartheid policies resulted in distinct spatial divisions in the distribution of different population groups, with little or no residential mixing permitted (Christopher, 1994; Lemon, 1991).

The complex web of legislative Acts passed by the government prior to 1994 can therefore be seen as intended to achieve two main inter-related objectives: (i) to ensure racial segregation of population groups in order to preserve the concept of a distinctly white racial group which underpinned the ideology that the white population was superior to the coloured, Indian and black African populations (Beinart, 1994); and (ii) to restrict employment and land-ownership opportunities for the non-white population thereby generating and sustaining a pool of cheap labour to service the demands of white economic interests (Lemon, 1991).

Under the apartheid regime property ownership rights for non-white population groups were restricted to either homeland areas or townships. The homelands, or Bantustans, were created by the National Party in 1951 through the Bantu Authorities Act 1951. Ten such homelands were established which together accounted for 13% of the total land of South Africa. These areas typically consisted of swathes of poor quality rural land located considerable distances from the main urban centres. The Bantu Authorities Act 1951 simultaneously assigned the black population citizenship rights to the homeland areas whilst removing their citizenship rights (such as they were) to the then Union of South Africa.

The National Party prohibited any investment of white capital in the homeland areas but did attempt to generate some economic development close to the boundaries of the homelands. It was envisaged that European entrepreneurs would be attracted to sites close to the borders of homelands due to the availability of cheap labour and subsidised power, water and transport. Economic development outside of but close to the homelands was seen as a means of encouraging the black African population to remain in these designated areas rather than migrate to the urban areas for employment. However, a variety of factors acted to discourage investment in these areas, including distance from main markets, absence of external economies (as benefited the principal urban cores) and other excessive costs (Feinstein, 2005). While some labour-intensive enterprises were attracted to these decentralised sites they remained disappointingly few as far as the National Party was concerned and the rate of joblessness in the homelands was hardly affected by the policies (Lester et al., 2000). Some homelands, such as the Transkei and Ciskei, were relatively large single-entity geographical areas whereas other homelands, such as KwaZulu, consisted of a series of smaller non-contiguous spatial areas surrounded by white-controlled farm land.

Townships were also defined by the white-controlled government and were typically located either within or on the periphery of urban areas for the purpose of providing a source of cheap labour for the white-owned industrial and commercial businesses operating in the towns and cities (Lemon, 1991). Christopher (1994) details the often piecemeal way in which land was allocated for the development of townships for non-Whites. In many cases, an individual farm on the periphery of an urban area would be purchased from a White land owner and demarcated as a township area for a particular non-White population group. This would lead to rapid in-migration by members of the designated population group and the establishment of a new, often densely populated, racially homogenous residential settlement. By their very nature, these areas would contain concentrations of individuals with restricted employment opportunities, limited economic and social capital, and no or poor quality service provision. These townships, often surrounded by white-owned farming land or sandwiched between white-owned suburban areas and white-owned farming land, therefore represented ‘islands’ or ‘pockets’ of severe poverty in otherwise affluent and privileged areas. The separation of the townships from the main urban cores, often by these coherent buffer strips, was also intended to increase the effectiveness with which they could be policed (Lester et al., 2000).

Many of the townships from the apartheid era have continued to grow in size since 1994. For example, the township of Khayelitsha in Cape Town was established in 1985 by the National Party and grew to a population of approximately 250,000 by 1996 and continued to grow to approximately 330,000 by 2001 and 392,000 by 2011 (City of Cape Town, 2005). It is common to find relatively formal and established communities within many township areas, such as Khayelitsha. However, townships are often accompanied by informal shack settlements that have grown up and continue to develop on the edges of the original township boundary (Parnell and Pirie, 1991). Thus spatial differentiation is apparent even within townships, with pockets of relatively more deprived and relatively less deprived people living in separate enclaves (Cook, 1991). So, for example, as Table 1.1 shows, in the township of Khayelitsha in Cape Town just under half the housing is formal (i.e. made conventional materials) whilst just over half comprise shacks either stand-alone backyards of formal housing.

Table 1.1: Dwelling type in Khayelitsha, 2001 Census

Dwelling type	**Number**	**Percentage**
Formal Housing	52,865	44.5%
Informal (i.e. shack)	64,761	54.5%
Other (including caravan and tent)	1,184	1.0%

In summary, therefore, the centuries of rule by colonial, segregationist and apartheid regimes created a country which, during the process of transition to democratic rule, faced a multitude of social and economic challenges. Large parts of the population were deprived, not only in terms of income and wealth, but also in terms of employment, education, housing, safety and access to services. These deprivations were structured along racial lines and the enforced spatial distribution of different population groups led to a unique spatial distribution of poverty, deprivation and inequality.

Post-apartheid period: 1994-present

When the transitional government, led by the African National Congress (ANC), replaced the apartheid regime in 1994 it gave a commitment to tackle the high levels of poverty, deprivation and discrimination afflicting large sections of the population. The legacy of centuries of colonial, segregationist and apartheid rule prior to 1994 were starkly apparent in the unequal distribution of resources and opportunities between different racial groups and between different parts of the country:

> "We have, at last, achieved our political emancipation. We pledge ourselves to liberate all our people from the continuing bondage of poverty, deprivation, suffering, gender and other discrimination" (Nelson Mandela, Inauguration Address, 10th May 1994, Pretoria).

In 2004 the incoming President, Thabo Mbeki, reiterated the need to remain focused on tackling the challenges that continued to face the nation:

> "Endemic and widespread poverty continues to disfigure the face of our country. It will always be impossible for us to say that we have fully restored the dignity of all our people as long as this situation persists. For this reason the struggle to eradicate poverty has been and will continue to be a central part of the national effort to build the new South Africa" (Thabo Mbeki, Inauguration Address, 27th April 2004, Pretoria).

Five years later, Mbeki's successor, Jacob Zuma repeated the need for a continued focus on tackling poverty and the whole range of South Africa's many interrelated socio-economic problems:

> "For as long as there are South Africans who die from preventable disease; for as long as there are workers who struggle to feed their families; for as long as there are communities without clean water, decent shelter or proper sanitation; for as long as there are rural dwellers unable to make a decent living from the land on which they live; for as long as there are women who are subjected to discrimination, exploitation or abuse; for as long as there are children who do not have the means nor the opportunity to receive a decent education; for as long as there are people who are unable to find work, we shall not rest, and we dare not falter" (Jacob Zuma, Inauguration Address, 9th May 2009, Pretoria).

Political statements such as the three examples provided here have been backed up by commitments enshrined through legislation and international agreements. For example, South Africa had obligations under the United Nations Millennium Development Goals (United Nations, 1995b) to halve the proportion of people whose income is less than $1US per day and halve the proportion of people who suffer from hunger between 1990 and 2015.

South Africa's poverty reduction commitment was articulated well before the international millennium targets were set, and was at the centre of the Reconstruction and Development Programme commitment to "meeting basic needs" that informed the democratic government's policy framework from 1994. The multi-dimensional nature of deprivation was asserted in the Reconstruction and Development Programme (RDP):

> "It is not merely the lack of income which determines poverty. An enormous proportion of very basic needs are presently unmet. In attacking poverty and deprivation, the RDP aims to get South Africa firmly on the road to eliminating hunger, providing land and housing to all our people, providing access to safe water and sanitation for all, ensuring the availability of affordably and sustainable energy sources, eliminating illiteracy, raising the quality of education and training for children and adults, protecting the environment, and improving our health services and making them accessible to all" (African National Congress, 1994).

South Africa is also a signatory to the Copenhagen Declaration 1995 (United Nations, 1995a) in which the country made a commitment to address poverty, promote social integration, create an enabling environment for social development, promote full employment, build the capacity of its people and mobilise resources for social development.

There are a number of other international covenants and treaties that South Africa has signed which commit signatory states to the advancement of the socio-economic rights of their inhabitants which have a direct bearing on people's states of well-being or impoverishment. These include:

- The International Covenant on Economic, Social and Cultural Rights (signed by South Africa on 3rd October 1994)
- The African Charter on Human and Peoples Rights (signed 9th July 1996)
- The African Charter on the Rights and Welfare of the Child (signed 10th October 1997)
- The Convention on the Rights of the Child (signed in 1989)
- The Convention on the Elimination of All Forms of Discrimination against Women (signed 29th January 1993)
- The International Convention on the Elimination of All Forms of Racial Discrimination (signed 3rd October 1994).

The challenges posed by the unique spatial distribution of poverty, deprivation and inequality are also reflected in the Constitution (Republic of South Africa, 1996). Under Section 214 of the Constitution, the Parliament must provide equitable shares and allocations of revenues to provinces and local governments, taking into account ten criteria, including: the 'need to ensure that the provinces and municipalities are able to provide basic services and perform the functions allocated to them'; the 'developmental and other needs of provinces, local government and municipalities'; and the 'economic disparities within and among the provinces'. In other words, financial resources must be distributed equitably among provincial and sub-provincial governments, based partly on levels of poverty, deprivation and inequality. In order for this constitutional requirement to be fulfilled, a robust evidence base is required quantifying the distribution of poverty, deprivation and inequality at both national and sub-provincial level.

Although I recognise that many commentators use the terms 'poverty', 'deprivation' and 'inequality' interchangeably, in this thesis I treat them as separate concepts that have separate measurement approaches. I follow Townsend's approach to distinguishing between poverty and deprivation. Townsend defined people as poor if 'they lack the resources to obtain the types of diet, participate in the activities and have the living conditions and amenities which are customary, or at least widely encouraged or approved in the societies to which they belong' (Townsend, 1979, p.31), whereas he defined people as deprived if 'they lack the types of diet, clothing, housing, household facilities and fuel and environmental, educational, working and social conditions, activities and facilities which are customary' (Townsend, 1979, p.131). Deprivation therefore refers to peoples' unmet needs, whereas poverty refers to the lack of resources required to meet those needs (Noble et al., 2006a). Inequality, on the other hand, refers to the unequal distribution of opportunities or outcomes (including financial resources such as income) amongst members of a given population. By extension, spatial inequality refers to the unequal distribution of opportunities or outcomes (or of people experiencing those opportunities or outcomes) between geographical areas within a given spatial constraint (e.g. between neighbourhoods within a country).

The recognition of challenges posed by the high levels of poverty, deprivation and inequality has led to considerable policy attention since the advent of democracy in 1994. The South African government has introduced an array of redistributive social policies, including expanding social security provision, free primary education in deprived areas, land reform, job creation programmes, and affirmative action in the labour market. Certain area-based interventions have also been introduced, most notably the Presidential Poverty Nodes that focused on a small number of highly deprived geographical areas and supported cross-cutting developmental priorities.

Despite this broad range of redistributive policies, it has been contended that prior to the 2007 ANC conference in Polokwane, the ANC government tended to prioritise tackling poverty and deprivation rather than reducing inequality (Habib, 2012). Following the 2007 ANC conference, however, inequality has risen to the forefront of policy attention and is now regarded, alongside poverty and unemployment, as one of the three core economic challenges facing South Africa's development. This emphasis on striving for a more equitable society is explicitly set out in two key policy documents: the New Growth Path (NGP, 2010) produced by the Ministry of Economic Development; and the National Development Plan 2030: Our Future Make it Work (NDP, 2012) produced by the Presidency.

Both these documents highlight inequality reduction as a core policy objective over the coming decades in South Africa and include explicit targets, including reducing the Gini coefficient from 0.69 to 0.60 by the year 2030 (NDP, 2012). The National Development Plan also explicitly recognises the importance of tackling spatial inequalities, both within and between urban and rural areas. For instance, there is an aim to reshape urban planning by minimising housing development in marginal areas and instead encouraging and facilitating housing near to labour market opportunities, thereby reducing travel times and costs. There is also an aim to bring rural areas into mainstream economic planning and to ensure that densifying rural areas are adequately structured (NDP, 2012).

I argue in this thesis that the increasing policy emphasis on tackling inequality requires a broader and more nuanced evidence base than currently exists in South Africa. This is all the more important given the international evidence base (primarily northern hemisphere-based research) suggesting the links between inequality and other social problems, such as violent crime, which can also be seen to afflict South Africa. I contend that it is imperative to go beyond the classical measures of income inequality at national level and instead look at a much more detailed level and how people's experience of inequality varies according to where they live and go about their daily routines.

1.3 Research Questions and thesis structure

This thesis is motivated by three key factors discussed above: first, a recognition of the expanding international evidence base concerning the role that socio-economic inequality may play in shaping other social outcomes, such as health, as well as people's attitudes and subsequently their actions; second, an acknowledgement that South Africa is widely regarded as being amongst the most unequal countries in the world and as having one of the highest levels of crime and social unrest in the world; and, third, an appreciation of the current dearth of suitable data on people's lived experience of inequality in South Africa which has so far limited the possible analytical approaches to exploring links between inequality, attitudes and actions within the country. This thesis is a response to the combination of these three key factors.

The main overarching research question pursued in this thesis is:

> How does the lived experience of inequality vary spatially across South Africa and how is this associated with people's attitudes towards inequality?

The aim throughout this thesis is therefore to establish whether an empirical measure can be developed which can be said to reflect people's lived experience of inequality and which can be used to explore associations with other socio-economic outcomes. As I will elaborate below, in this thesis I will, as a case study, look at whether the lived experience of inequality may be associated with people's attitudinal outcomes, and particularly their attitudes towards inequality and options for redress. I deliberately do not look at possible associations between the lived experience of inequality and action outcomes such as crime in this thesis (for reasons expanded upon in Chapter 7), although future research (outside the scope of this thesis) may do so.

In order to address the overarching research question I have identified five separate sub-questions which I discuss further below:

> Sub Q1: To what extent is deprivation distributed unequally across neighbourhoods in South Africa?

> Sub Q2: Can an empirical measure be developed that reflects people's lived experience of inequality?

> Sub Q3: Where within South Africa is people's lived experience of inequality highest and lowest?

Sub Q4: Are there any neighbourhoods across South Africa with high rates of deprivation and high lived experience of inequality?

Sub Q5: To what extent are people's attitudes to inequality and options for redress associated with their lived experience of inequality?

This thesis contains eight chapters plus seven appendices. The eight chapters consist of this introduction, a literature review, five empirical chapters, and a conclusion. Each of the five empirical chapters addresses a specific research sub-question from the list above. I now provide a short summary of the purpose of each of the following seven chapters of the thesis plus each of the seven appendices and in doing so elaborate the research sub questions outlined above.

In Chapter 2, I begin by reviewing the international literature linking inequality to a variety of social and economic outcomes, such as crime and poor health. I then consider how levels of inequality in South Africa compare to other countries globally and then particularly within sub-Saharan Africa. The focus then turns to South Africa-specific research concerning inequality, including inequality measured across the country as a whole, between different population sub-groups and between different geographical areas. As part of this I highlight certain key studies on poverty and deprivation within South Africa which adopt a spatial focus. I conclude with a brief overview of the specific literature relating to spatial inequality measurement. However, the main review of this literature is undertaken in the relevant empirical chapters.

Chapter 3 is the first of the five empirical chapters. The aim in this chapter is to establish the foundations for the following empirical chapters by demonstrating the *spatial* nature of inequality in South Africa. The exemplar photographs of urban Johannesburg and rural former Transkei provided here in Chapter 1 illustrate some of the differing socio-spatial environments across South Africa. As discussed above, I contend that it is the socio-spatial environment in which a person lives and in which they carry out their routine daily activities that shapes their lived experience of inequality. As such, in order to understand how people's lived experience of inequality varies spatially across the country, it is first necessary to understand the spatial structuring of levels of *deprivation* across the country and thereby address the first research sub-question:

> Sub Q1: To what extent is deprivation distributed unequally across neighbourhoods in South Africa?

Although the two exemplar photographs of urban Johannesburg and rural former Transkei *suggest* that deprivation *may* be distributed unequally across neighbourhoods in South Africa, the aim within Chapter 3 is to explore this *analytically*. I begin by providing a summary of the different geographical units available in South Africa and their respective strengths and weaknesses for measuring deprivation and inequality at neighbourhood level. I proceed to introduce in more detail the main dataset used in this thesis: the South African Index of Multiple Deprivation (SAIMD) 2001 at Datazone[2] level (Noble et al., 2009a). In order to explicitly address Sub Q1, I utilise the SAIMD and its constituent domains of deprivation to undertake an empirical analysis of the spatial distribution of deprivation across South Africa. I explain how the SAIMD

[2] As I explain in Chapter 3, Datazones are a statistical geography comprising on average 2000 people per Datazone.

contains five constituent domains of deprivation that are each are measured separately and that four of these domains are technically suitable for forming the basis of the spatial inequality measures I review in Chapter 4. I begin by presenting an analysis of deprivation levels at national and provincial levels as context, before drawing upon the datazone level deprivation rates and ranks to examine the unequal distribution of deprivation between neighbourhoods. The datazone analyses include consideration of the 'types' of neighbourhoods that are most deprived and least deprived nationally, and I highlight those municipalities that are characterised by the greatest heterogeneity of deprivation levels amongst their constituent neighbourhoods. In summary, Sub Q1 requires me to look at the spatial configuration of deprivation levels across the country to make an assessment of the differing socio-spatial environments in which people live and carry out routine daily activities. As noted above, it is this socio-spatial environment which I propose forms the basis of people's lived experience of inequality.

In Chapter 4, the focus turns to developing a new statistical measure of the lived experience of inequality. Specifically, the objective is to develop a new empirical measure which reflects how people's lived experience of inequality is shaped by the socio-spatial environment in which they live their lives. In Chapter 4, I therefore address:

> Sub Q2: Can an empirical measure be developed that reflects people's lived experience of inequality?

Sub Q2 requires me to go beyond the spatial analysis of deprivation patterns addressed in Chapter 3 to consider how these complex patterns may manifest as determinants of people's lived experience of inequality. I refer to Massey & Denton's (1988) seminal work on measures of residential segregation and I review each of their five dimensions of segregation against a set of criteria to determine which statistical measures might be suitable for my purpose. I also refer to the distinction between ‘global’, ‘spatial’ and ‘local’ measures of segregation as discussed by Lloyd & Shutleworth (2012). Based upon a combination of conceptual reasoning and empirical testing, I identify a particular measure of residential segregation that I believe satisfies my three criteria for a measure of the lived experience of inequality. My chosen measure is a form of a ‘local distance-weighted Exposure index’. I proceed to describe the construction of this particular measure, including the specification of key methodological parameters. My approach to specifying these parameters represents a divergence from the conventional approach most widely used in northern hemisphere countries and instead, I argue, is designed to better reflect the social context of South Africa. The methodological development work undertaken here generates two complementary variants of the exposure index: one reflects exposure to inequality as experienced by the ‘poor’, and the other reflects exposure to inequality as experienced by the ‘non-poor’. In order to operationalise these two variants of my chosen exposure measure I initially use data from the income and material deprivation domain of the SAIMD 2001.

Having developed an empirical measure of the lived experience of inequality, the objective in Chapter 5 is to analyse the spatial patterning of the results. In this chapter, I explicitly address research sub-question Q3:

> Sub Q3: Where within South Africa is people’s lived experience of inequality highest and lowest?

In the analysis in response to sub Q3, I begin by looking across all neighbourhoods in South Africa to identify where, geographically, people's lived experience of inequality is highest and lowest using the measure developed in Chapter 4. As explained at the beginning of Chapter 5, the focus here is explicitly on the measure of exposure as experienced by the 'poor' population, with equivalent analyses for the 'non-poor' population presented in Appendix C. Within Chapter 5, I initially focus the analyses on the exposure measure underpinned by the income and material deprivation domain of the SAIMD, as used to detail the operationalisation of the measure in Chapter 4. I show in Chapter 5 how levels of exposure to inequality vary at datazone level both between and within larger areas such as provinces and municipalities. A distinction is made between metropolitan and non-metropolitan datazones with regard to the patterns of results. I then proceed to consider levels of exposure to inequality as measured using the other three domains of the SAIMD as input data, this time with a particular focus on metropolitan datazones.

In Chapter 6, I review the spatial patterns of exposure to inequality revealed in Chapter 5 in conjunction with the spatial patterns of deprivation presented in Chapter 3. In light of the literature discussed in Chapter 2 that suggests both deprivation *and* inequality may play important roles in shaping people's attitudes and actions, the aim in Chapter 6 is to overlay these two sets of results and look for commonalities and differences. Specifically, the aim in Chapter 6 is to explicitly address research sub Q4:

> Sub Q4: Are there any neighbourhoods across South Africa with high rates of deprivation and high lived experience of inequality?

As I demonstrate in Chapter 6, some areas of South Africa are indeed characterised by high levels of deprivation *and* exposure. In Chapter 6, I therefore develop a new empirical measure that draws upon the deprivation data and the exposure data to produce a small area level measure of community 'intensity' of exposure. Areas that score highly on this new 'intensity' measure are those with high levels of deprivation and high levels of exposure to inequality. I propose in Chapter 6 that it is those neighbourhoods where intensity is highest that may be most vulnerable to crime and unrest due to the combined effects of deprivation and exposure to inequality on people's attitudes and actions.

The new intensity measure was constructed using the deprivation rate and exposure score relating to the income and material deprivation domain of the SAIMD. Chapter 6 presents national analysis of the intensity measure across all datazones in South Africa and examines spatial patterns both between and within provinces and municipalities. I identify differential in results between metropolitan and non-metropolitan datazones, and so then focus on metropolitan areas in more detail. I additionally generate intensity measures using the other three suitable domains of the SAIMD (following the approach adopted in Chapters 3 and 5) and analyse the spatial distribution of results over the set of metropolitan datazones.

Chapter 7 is the fifth and final empirical chapter. The objective here is to present one example of how my new exposure measures (and the associated intensity measures) can be used to explore associations with other socio-economic outcomes. For the purpose of this thesis I have chosen to look at associations between people's exposure to inequality and their attitudes to inequality and options for redress, using the South African Social Attitudes Survey (SASAS). Specifically, the aim in Chapter 7 is to address research sub-question:

Sub Q5: To what extent are people's attitudes to inequality and options for redress associated with their lived experience of inequality?

Having established the motivation for the chapter, I proceed to introduce the data and methods that are employed. A feature of the analysis here is the use of data linkage to attach datazone level information on deprivation rates and exposure scores to each individual SASAS respondent based upon the sampled Enumeration Area which contains the respondent's home address. This permits individual level attitudinal outcomes to be considered in the context of individual level characteristics and area level characteristics. The primary aim in this chapter is to assess whether there are any statistical associations between my new exposure measures (and/or the intensity measures) and attitudinal outcomes concerning inequality and options for redress. Having discussed the datasets and data preparation steps, I then proceed to discuss the specification of four multilevel statistical regression models. With regards to the presentation and discussion of the modelling results, I focus explicitly within Chapter 7 on associations between exposure and attitudes amongst the 'poor' population. Equivalent analyses for the 'non-poor' population are provided in Appendix E. I conclude that there is indeed some evidence of associations between my exposure measures and attitudinal outcomes concerning inequality and options for redress, although these associations tend to manifest through interactions with other independent variables.

Finally, in Chapter 8, I reflect upon the findings of the five empirical chapters and I highlight how each component part of the thesis addresses a particular research sub-question and how this contributes to addressing the main overarching research question. I end by acknowledging the main limitations of my research and I make a number of recommendations on how the research can be further developed in the future.

There are also eight appendices to this thesis. In Appendix A, I confirm that all the research presented in this thesis is my own work and I acknowledge the input I have received from others that helped to shape my conceptual and analytical approaches. In Appendix B, I provide additional technical details and empirical workings concerning a series of residential segregation measures that I rejected in Chapter 4. In Appendix C, I provide analyses of exposure to inequality from the perspective of the ‘non-poor’ population, thereby complementing the equivalent analyses in Chapter 5 concerning the ‘poor’ population. In Appendix D, I provide further output on the base models developed in Chapter 7. In Appendix E, I present and discuss the results of the statistical models of attitudinal outcomes from the perspective of the ‘non-poor’ population, thereby complementing the equivalent analyses in Chapter 7 concerning the ‘poor’ population. In Appendix F, I provide further output on the base models developed in Appendix E. In Appendix G, I present the geographical variance terms relating to each of the four final statistical models discussed in Chapter 7 and Appendix E. In Appendix H, I provide some additional background technical details concerning the sampling methodology of the SASAS dataset that I utilise in Chapter 7 and Appendices D, E and F. Finally, Appendix I contains a short glossary of selected key terms.

Chapter 2: Literature review

The main objective of this thesis is to develop a measure of spatial inequality that reflects the lived experience of inequality. Although there is an extensive international literature on the links (and potential causal relationships) between inequality and a wide variety of social and economic ills, much less research has been undertaken to quantify the extent to which people's experience of inequality varies spatially and how this might relate to a range of outcomes at the sub-national level.

This chapter has a number of objectives which relate directly to the main focus of the thesis. In consequence, I do not analyse in detail the international literature relating to measurement of inequality (except in so far as it is relevant to the measurement of the spatial patterning of inequality *within* South Africa). Instead I first consider the international literature linking inequality to a variety of social and economic outcomes. I then situate the levels of inequality in South Africa in the broader global context. The focus then turns to South Africa-specific research concerning inequality, including inequality measured across the country as a whole, between different population sub-groups and between different geographical areas. As part of this discussion I also highlight certain key studies on poverty and deprivation within South Africa which are of relevance to the spatial inequality focus of this thesis. I conclude with a brief overview of the specific literature relating to spatial inequality measurement. This literature is then further developed within each of the empirical chapters as necessary.

2.1 Inequality and its links with social and economic outcomes

As I noted briefly in Chapter 1, there has been a notable increase in academic and policy attention directed at socio-economic inequality and its negative social and economic impact over recent years (Atkinson, 2015a). Since the global recession at the end of the last decade, inequality has also increasingly become a matter of key public debate, with particular questions asked of the legitimacy of a large proportion of global resources being in the hands of very few individuals (e.g. Atkinson, 2015a; Dorling, 2014b; Hardoon et al., 2016; Piketty, 2014; Stiglitz, 2012). Reflecting this renewed emphasis on inequality, a number of key international actors and organisations have made explicit reference in public addresses or documents to the challenges that inequality may pose to our future economies and societies.

The United Nations has placed the reduction of inequality squarely on the development agenda (Transforming our world: the 2030 Agenda for Sustainable Development) by the adoption of the sustainable development goals (SDGs) in paragraph 54 of United Nations Resolution A/RES/70/1 which was signed in September 2015. In particular, SDG 10 is "to reduce inequality within and among countries". There are ten targets associated with this goal, the first four of which are of relevance to the issue of within-country inequality:

> "10.1 By 2030, progressively achieve and sustain income growth of the bottom 40 per cent of the population at a rate higher than the national average
>
> 10.2 By 2030, empower and promote the social, economic and political inclusion of all, irrespective of age, sex, disability, race, ethnicity, origin, religion or economic or other status
>
> 10.3 Ensure equal opportunity and reduce inequalities of outcome, including by eliminating discriminatory laws, policies and practices and promoting appropriate legislation, policies and action in this regard

> 10.4 Adopt policies, especially fiscal, wage and social protection policies, and progressively achieve greater equality.[3]"

Indeed this particular SDG was specifically addressed at the 2013 General Assembly Thematic Debate on Inequality, where the UN Secretary General, Ban Ki-moon, stated:

> "…in many countries, rich and poor, we are seeing social and economic inequalities widening… Social and economic inequalities can tear the social fabric, undermine social cohesion and prevent nations from thriving. Inequality can breed crime, disease and environmental degradation and hamper economic growth. If inequalities continue to widen, development may not be sustainable. That is why equity is emerging as a central plank in discussions on the post-2015 development agenda."" Ban Ki-moon 8th July 2013[4].

Ban Ki-moon's statement highlights two important types of possible consequences of inequality: first, social consequences such as crime and poor health; and second, economic consequences, including reduced overall economic growth. Whilst the links between inequality and social problems such as crime (e.g. Becker, 1968; Sampson and Groves, 1989; Shaw and McKay, 1942; Wikstrom, 2004) and poor health (e.g. Pickett and Wilkinson, 2015; Wilkinson, 1996; Wilkinson and Pickett, 2010) have been posited for some time, the potential for inequality to hamper broad economic growth and development has emerged more recently (e.g. Dabla-Norris et al., 2015). Before I turn to discuss the literature on the links between inequality and other social outcomes, I will therefore first highlight a selection of key public statements and/or documents from international actors or organisations that emphasise the need to address inequality in order to maximise economic growth. Although I do not concern myself with the relationships between inequality and economic growth in this thesis, I highlight this issue briefly here as it is important context to the increased policy discussion concerning

[3] See https://sustainabledevelopment.un.org/post2015/transformingourworld
[4] http://www.un.org/sg/statements/index.asp?nid=6955

inequality internationally. It is also highly relevant to the building of a broad consensus about the need to reduce inequalities through pro-poor redistributive policies.

The International Monetary Fund (IMF) has begun to acknowledge that inequality can adversely affect economic development and growth, and that increased incomes of the rich do not necessarily trickle down to benefit the rest of the population. In 2012 at the Annual Meetings of the IMF and World Bank, the head of the IMF, Christine Lagarde, stated:

> "…recent IMF research tells us that less inequality is associated with greater macroeconomic stability and more sustainable growth" [5]

The IMF subsequently released a report focused explicitly on the causes and consequences of inequality across the globe:

> "…if the income share of the top 20 percent (the rich) increases, then GDP growth actually declines over the medium term, suggesting that the benefits do not trickle down. In contrast, an increase in the income share of the bottom 20 percent (the poor) is associated with higher GDP growth. The poor and the middle class matter the most for growth via a number of interrelated economic, social, and political channels." (Dabla-Norris et al., 2015, p.4)

The World Bank has also acknowledged that high levels of inequality can hamper economic growth and poverty reduction as well as causing wider social problems:

> "… high levels of inequality can impose heavy socioeconomic costs on society. Mechanically, higher initial inequality results in less poverty reduction for a given level of growth. Tentative evidence also suggests that inequality leads to lower and less sustainable growth and thus less poverty reduction (Berg et al., 2012)" (Beegle et al., 2016, p.15).

Similar concerns have been raised by OECD Secretary-General Angel Gurria in his speech at the launch of the OECD report *'In it together: why less inequality benefits all'* on 21st May 2015:

[5] https://www.imf.org/external/np/speeches/2012/101212a.htm

> "We have reached a tipping point. Inequality in OECD countries is at its highest since records began. The evidence shows that high inequality is bad for growth. The case for policy action is as much economic as social. By not addressing inequality, governments are cutting into the social fabric of their countries and hurting their long-term economic growth."[6]

Furthermore, on 4th Dec 2013, Barack Obama described rising inequality as “the defining challenge of our time” and that “over the course of the next year and for the rest of my presidency, that's where you should expect my administration to focus all our efforts".[7]

These statements by powerful and influential international actors and organisations have contributed to the increasing academic and policy discourse concerning the potentially damaging effects of inequality on economies across the world.

As noted above, there is also an established and growing body of literature arguing that inequality may fuel specific social problems such as crime and poor health which needs to be considered alongside the material relating to economic challenges posed by inequality. For instance, in their book, *The Spirit Level*, Wilkinson & Pickett (2010) present an array of empirical evidence through which they demonstrate that more unequal societies tend to fare worse on a range of outcome measures when compared to countries where inequality is lower. They constructed a composite index of health and social problems consisting of indicators of life expectancy; maths and literacy; infant mortality; homicides; imprisonment; teenage births; trust; obesity; mental illness (including drug and alcohol addiction); and social mobility. The health-related outcomes within the composite index therefore include measures of both physical ill health and mental ill health, as well as mortality indicators. The positive relationship

---

[6] http://www.oecd.org/social/reducing-gender-gaps-and-poor-job-quality-essential-to-tackle-growing-inequality.htm
[7] https://www.whitehouse.gov/the-press-office/2013/12/04/remarks-president-economic-mobility

between inequality and this composite health and social problem index is shown in Figure 2.1 which is reproduced directly from Wilkinson and Pickett.

Figure 2.1: Relationship between income inequality and composite index of health and social problems in selected developed countries

Source: Wilkinson and Pickett (2010, p.20)

It is clearly evident from Figure 2.1 that the more unequal countries (e.g. USA, Portugal, UK) score highest on the index of health and social problems, whilst the more equal countries (e.g. Japan, Finland, Norway, Sweden) score at the lowest end of the health and social problem index. Wilkinson and Pickett also looked at the relationships between income inequality and various health and social problems separately and found positive relationships at country level between the level of inequality and various physical health outcomes (e.g. obesity and infant mortality rates) and various mental health outcomes (e.g. percent with any mental illness), as well as outcomes relating to drug use and violence. They also found negative relationships between the level of inequality and human capital indicators (e.g. school attainment) and between inequality and social capital indicators (e.g. trust).

The publication of The Spirit Level stimulated increased international debate on the subject of inequality and its links with other socio-economic outcomes. While a small minority of commentators have criticised the statistical analysis and interpretation of Wilkinson and Pickett (e.g. Saunders, 2010; Snowdon, 2010), the majority have broadly supported Wilkinson and Pickett's analyses and interpretations (e.g. Rowlingson, 2011).

More recently, Pickett and Wilkinson (2015) reviewed new empirical evidence emerging since the publication of *The Spirit Level* against an epidemiological framework for causal inference. They conclude that:

> "The body of evidence on income inequality and health points strongly to a causal connection. The major criteria of temporality, biological plausibility, consistency and lack of alternative explanations are well supported. Of the small minority of studies which find no association, most can be explained by income inequality being measured at an inappropriate scale, the inclusion of mediating variables as controls, the use of subjective rather than objective measures of health, or follow up periods which are too short." (Pickett and Wilkinson, 2015, p.323)

They go on to say:

> "We suggest that the most parsimonious explanation for the effects of income inequality is that larger income differences increase social distances, accentuating social class or status differences…Rather than income inequality being a new and independent determinant of health, it is likely to act by strengthening the many causal processes (known and unknown) through which social class imprints itself on people throughout life. This would suggest why, not only health, but a wide range of other outcomes with social gradients are also related to inequality. It also suggests that if class and status are to become a less powerful influence both on individual lives and on whole societies, it will be necessary to reduce the material differences which are so often constitutive of

> the cultural markers of social differentiation." (Pickett and Wilkinson, 2015, p.323-324)

A crucial finding from Wilkinson & Pickett's work is that outcomes tend to be worse *for all levels of society* in countries where inequality is relatively high compared to countries where inequality is relatively low. These analyses suggest that the damaging effects of high levels of inequality are not felt only by the poor and disadvantaged, but rather by the broader population as a whole. In terms of the basis upon which to build a broad consensus for inequality reduction, the arguments of commentators such as Wilkinson & Pickett can be seen to complement the separate arguments regarding economic development and growth noted above.

Within the criminological literature, inequality is frequently cited as an important factor in crime causation. Whilst it is outside the scope of this thesis to consider the criminological debates in detail, it is pertinent to summarise some key propositions concerning inequality's direct and indirect roles in criminal behaviour and the resulting crime patterns.

The basic premise of *economic* models of crime (e.g. Becker, 1968) is that individuals make a rational choice as to whether to offend or not based upon a consideration of (i) anticipated gains from illegal activity compared to anticipated gains from legal activities; (ii) the likelihood of being apprehended (and the subsequent likelihood of actually being convicted); (iii) the severity of likely punishment if apprehended and convicted; and (iv) the perceived likelihood of being able to successfully access opportunities for legal activities. These factors are as applicable to an individual at the top of the income distribution as an individual at the bottom of the income distribution (consider, for instance, the crimes committed by bankers such as the rigging of the Libor rates). Inequality can be seen to directly feed into this rational choice model in two main ways. First, in a highly unequal society an individual may perceive there to be far greater potential gains to be achieved through illegal activities than through legal labour market activities, especially when the income distribution is highly skewed towards the highest earners. Second, less well-off individuals in highly unequal societies may perceive there to be fewer potential legitimate opportunities for legal activities, such as through the labour market.

It has also been contended that inequality may play an indirect role in determining the likelihood of apprehension, and this is addressed in a range of *social* theories of crime. For example, in their theory of social disorganisation, Shaw & McKay (1942) contend that socially organised communities can resist crime more effectively than socially disorganised communities. They argue that this is because socially organised communities are more able to exert effective control over their members and more able to identify perpetrators (thus increasing the risk of offender apprehension). A large body of research now exists on the subject of social disorganisation (see, for example, Bursik, 1988; Kawachi et al., 1999; Sampson and Groves, 1989; Weisburd et al., 2014) and how this relates to the broader concept of social cohesion. Sampson and colleagues have further advanced this field of enquiry and developed a theory of collective efficacy in which the protective role of social organisation is operationalised through the willingness of community members to actively intervene when observing acts of crime, disorder or delinquency (Sampson, 2006; Sampson et al., 1999; Sampson et al., 1997; Sampson et al., 1998). According to the theory of collective efficacy, socially incohesive communities where members are less willing to intervene to stop acts of crime, disorder or delinquency for the collective benefit of the local community are regarded as being more vulnerable to criminal activity.

The routine activity theory (RAT) proposed by Cohen & Felson (1979) contends that crime occurs when a motivated offender meets a suitable victim in the absence of a capable guardian. Although Cohen & Felson do not regard inequality as *direct* determinant of crime occurrence in their RAT, there are a number of ways in which inequality may feed into the RAT. For instance, as discussed above in relation to economic models of crime, inequality may be an important motivation for an offender to commit crimes if, based upon his/her rational choices, the net returns from illegal activity over legal are seen to outweigh the risks of the illegal activity. Furthermore, as discussed in relation to social disorganisation and collective efficacy, inequality may affect community resilience to crime and, as such, it is of relevance to the component of RAT concerned with capable guardianship.

An important recent development in the criminological literature, and one which is particularly relevant for this thesis, is the Situational Action Theory crime causation (SAT) articulated by Wikstrom (Wikstrom, 2004; Wikstrom, 2006a; Wikstrom, 2006b; Wikström, 2005; Wikström, 2014; Wikström and Treiber, 2009). Wikstrom argues that poverty, deprivation and inequality are some of the "causes of the causes" of crime occurrence (Wikstrom, 2006a, p.62), in that poverty, deprivation and inequality shape individuals' attitudes towards "rule breaking" and this in turn shapes the "moral context" of the community. Wikstrom suggests it is the interaction between an individual's propensity to break moral rules and the community's propensity to reject, permit or facilitate moral rule breaking that is the primary cause of criminal activity. In this sense, poverty/deprivation and inequality are seen as some of the causes of certain social attitudes and it is these social attitudes which are the direct causes of crime.

In summary, there is growing international concern about the impact of high levels of inequality on both economic and social outcomes. Whilst there is ongoing debate around the extent to which inequality is a *direct* or *indirect* determinant of social problems such as crime and poor health, there is a growing consensus that high inequality is associated with worse outcomes across a range of measures.

The focus in this thesis is on South Africa which, as noted in Chapter 1, is one of the most unequal countries in the world and fares poorly on many other socio-economic outcomes, including having one of the highest rates of violent crime in the world (UNODC, 2011). This is despite being classed as an upper middle income country by the World Bank[8]. I will now proceed to consider the empirical evidence concerning inequality levels in South Africa.

2.2 Inequality patterns and trends in South Africa

The majority of research on inequality in South Africa is mainly at a national level and is based upon the 'classical' measures of inequality, such as the Gini coefficient or General Entropy (GE) measures[9]. A number of measures of inequality have been produced in recent years by different individuals and organisations. These estimates of inequality differ due to various reasons, including using different underpinning datasets (i.e. Census; household surveys); relating to different points in time; and using different metrics (i.e. income; expenditure), as well as different types of measures.

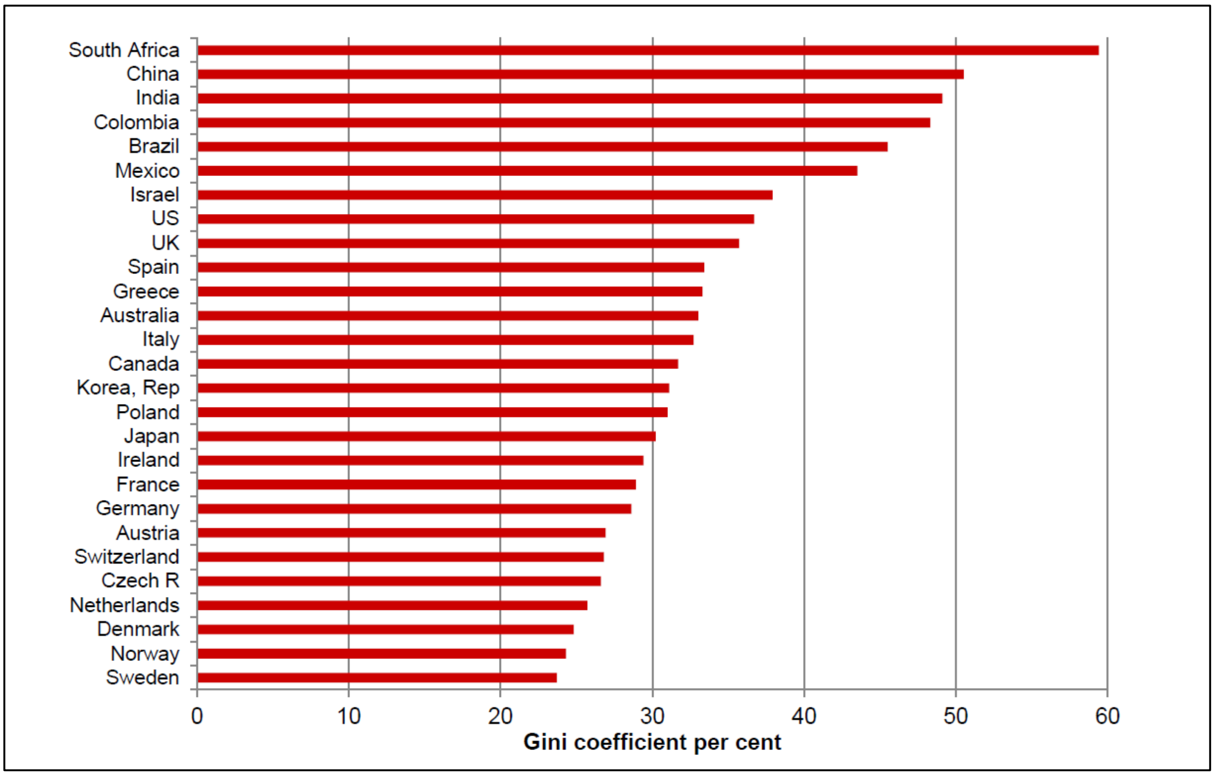
Irrespective of the inequality measure chosen (i.e. Gini; GE; etc), the underpinning dataset, the time point or whether income or expenditure is used, South Africa exhibits one of the highest levels of inequality in the world (Leibbrandt et al., 2010).

[8] http://data.worldbank.org/country/south-africa
[9] See Glossary for brief description of these classical inequality measures

Although I argue throughout this thesis that broad national level measures of inequality, such as the commonly used Gini coefficient, cannot adequately reflect people's 'lived experience' of inequality, it is nevertheless instructive to give a brief overview of the existing literature in order to provide background context for the analyses that follow in subsequent chapters.

Figure 2.2 shows national level Gini coefficients for a wide selection of countries calculated using equivalised household disposable income. The data for most countries depicted in the chart relate to the year 2010, with a small number of countries deviating from this year (as detailed in the note below the chart).

Figure 2.2: Gini coefficient in a selection of countries



Source: Atkinson (2015b, p.7)

"Note: This graph shows the Gini coefficient for equivalised household disposable income in different countries, ranked in decreasing order. The coefficient in Sweden was 23.7 per cent. Sources: LIS Key Figures http://www.lisdatacenter.org/data-access/key-figures/download-key-figures/, downloaded 30 November 2014. The data are for 2010 except for Austria (2004), Brazil (2011), China (2002), Czech Republic (2004), Korea (2006), India (2004), Japan (2008), Sweden (2005), and Switzerland (2004)." (Atkinson, 2015b, p.7)

South Africa can be seen at the very top of the chart, with the highest score on this inequality measure of all countries considered here. Furthermore, South Africa appears to be somewhat of an outlier, with a Gini coefficient far in excess of any other country shown. The Latin American countries that sit below South Africa, and which are widely regarded as being very unequal countries, can be seen to register Gini coefficients around the mid-40s when expressed as a percentage as in this chart. South Africa, on the other hand, can be seen to exhibit a Gini of close to 60 on this same measure.

It could perhaps be argued that situating South Africa in the context of the heterogeneous group of countries above, which contains many highly developed nations such as the Nordic states, presents a somewhat unfair comparison. The new World Bank report *Poverty in rising Africa* (Beegle et al., 2016) contains a section focused solely on the African continent (effectively sub Saharan Africa as it excludes North African countries). Here the Gini coefficient is calculated based upon consumption. The World Bank analyses show that the southern African countries of South Africa, Namibia and Botswana rank the highest on this Gini measure of all African countries for which data are presented. The African countries with the lowest Gini scores are those located across the Sahara itself (e.g. Niger, Mali and Sudan) and in the tropics just south of the Sahara (e.g. Sierra Leone). Of course, levels of poverty and deprivation are very high in these countries, so having a relatively low Gini score does not mean these countries have fewer socio-economic challenges than highly unequal countries such as South Africa.

To unpick this issue further, the UNDP's 2015 Human Development Report (UNDP, 2015), shows South Africa's position on various indicators of human development. Based upon the overall composite Human Development Index (HDI), South Africa is ranked at position 116 out of 188 nations globally (for background context, Norway is ranked 1st and the UK is ranked 14th). Sierra Leone, Guinea, Mali, Niger, Sudan and Ethiopia, which together constitute the band of relatively low inequality countries (relative to the other African nations) identified in the World Bank (Beegle et al., 2016) analysis, are ranked between positions 167 and 188 on the UNDP's HDI in 2015, confirming that poverty and deprivation (as measured by the HDI) are more widespread in these counties than in South Africa.

The UNDP's 2015 Human Development Report also contains information on the levels of inequality within the global list of countries, including South Africa. Estimates of inequality in each of the three constituent components of the HDI (life expectancy, education and income) were produced by the UNDP using the Atkinson inequality index. These inequality measures were then used to adjust the standard scores on each of the three constituent components of the HDI. The three measures of inequality were also averaged to produce a composite 'coefficient of human inequality'. Furthermore, the three inequality-adjusted HDI components were then combined to produce an inequality-adjusted HDI. It is apparent from the UNDP's analysis that the inequality-adjusted HDI score for South Africa (0.428) is notably lower than the standard HDI score (0.666). This is an overall loss of 35.7%, which translates to a rank on the inequality-adjusted HDI of 15 places lower than on the standard HDI (UNDP, 2015, p.217). Based upon this study, it would seem that South Africa's level of human development is considerably restrained by the high levels of inequality within the country.

The discussion of the inequality-adjusted HDI provided here highlights the potential to measure inequality in dimensions other than income, expenditure or consumption (as noted above, the HDI contains components on life expectancy and education as well as a component on income) and the importance of considering actual levels of poverty/deprivation *alongside* levels of inequality. As will become apparent later in this thesis, I have also adopted an analytical approach that reflects both these factors.

These international comparisons provide a valuable means by which to situate South Africa in the context of its neighbouring countries and more widely across the globe. However, in order to unpick the aggregate inequality figures for South Africa it is necessary to now shift the focus squarely onto South Africa and consider the more detailed evidence relating to inequality within the country. Although the focus of this section will be on inequality, I will also briefly touch on the measurement of poverty and deprivation which, as I will demonstrate, also have a bearing on this thesis.

Before discussing the findings from empirical research on inequality, poverty and deprivation within South Africa, the main data sources and measurement metrics utilised are summarised.

In terms of main data sources within South Africa, approaches to measuring poverty, deprivation and inequality typically utilise data from the Census (1996, 2001 and 2011) and/or a range of household surveys. Commonly used household surveys include the Income and Expenditure Surveys (IES), national Labour Force Surveys (LFS), the National Income Dynamics Study from 2008 onwards, the Community Survey 2007, and the Living Conditions Surveys. Unlike in the United Kingdom and many other more developed countries, administrative data are rarely if ever used to measure poverty, deprivation and inequality in South Africa. However, this situation is changing and it is likely that administrative data will become increasingly important over the coming years as the quality and availability of such data improve (Barnes et al., 2007a).

As regards the main measurement metrics used in the study of *poverty* in South Africa, there is no single official poverty line, but rather a range of poverty lines developed by different individuals and organisations at different points in time and for different purposes. Both so called absolute approaches (such as $1 per day or $2 a day) and relative poverty lines (such as proportions of mean or median equivalised household income) have been used in order to calculate headcount measures of poverty and poverty gap measures (e.g. May, 1998; Noble et al., 2006a). Although not classed as *official* poverty lines, the 'upper bound' and 'lower bound' income-based poverty lines of Hoogeveen & Ozler (2006) are widely used in academia and for the calculation of inequality (e.g. Leibbrandt et al., 2010). Consumption-based poverty lines are mainly used by government and consumption data is also used for the calculation of inequality (e.g. Statistics South Africa, 2012).

With regards to the main measurement metrics used in the study of *deprivation* in South Africa, some commentators have constructed measures of particular dimensions of deprivation and examined these independently of one-another (see, for example, Aguero et al., 2007; Bhorat and Kanbur, 2005; Bhorat et al., 2004; Burger et al., 2004; Leibbrandt et al., 2005), while in other cases measures of deprivation are combined together into composite indices at national or provincial level (see, for example, Hirschowitz et al., 2000; Klasen, 2000; Mattes et al., 2003; UNDP, 2003).

With regards to the main measurement metrics used in the study of inequality in South Africa, as has been noted above, the Gini coefficient is the most widely used, but the GE measures and the Atkinson index have also been adopted. Furthermore, some commentators have used asset indices for measuring poverty or inequality (e.g. Bhorat et al., 2006; McKenzie, 2003), although other commentators have cautioned against using asset indices for measuring inequality on methodological grounds (Wittenberg and Leibbrandt, 2015).

The wide variety of approaches taken to explore inequality and its associated bases of poverty and deprivation using national level data has generated an ongoing debate as to the actual levels and trends over time. I will now provide a brief summary of empirical research in South Africa on levels of inequality, poverty and deprivation in turn.

With regards to inequality, Leibbrandt et al (2010) analysed patterns and trends within South Africa using both Census data and household survey data. Based upon their analyses of income inequality using Census data between 1970 and 2001 they concluded that South Africa's national level Gini coefficients were very high across this period by international standards. They found considerable inequality between the four main population groups (Black African, Coloured, Indian/Asian, and white) at each time point and they identified increasing income inequality within each of the four main racial groups over the period 1975 to 2001. These income dynamics resulted in an increase in the overlap of the within-racial group income distributions over the period. They concluded that the increase in aggregate income inequality was driven primarily by increasing inequality within the Black African population group and that this was accentuated by the increasing population share of the Black African group over the time period.

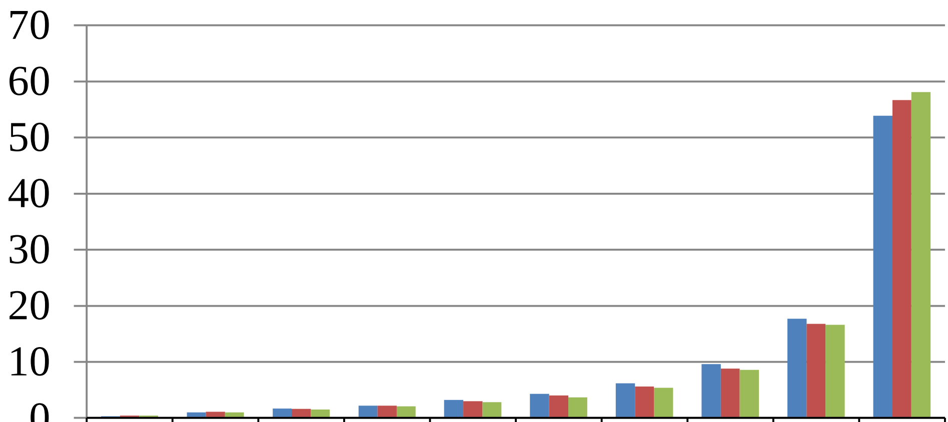
Leibbrandt et al (2010) also undertook analyses of income inequality patterns and trends using household survey data for 1993, 2000 and 2008. The results of their survey analyses generally support their findings from the Census analysis in that levels of income inequality were found to be very high by international standards and were found to increase over the time period under consideration. The overall national level Gini for income inequality was calculated to be 0.66 in 1993, increasing to 0.68 in 2000, and increasing again to 0.70 in 2008. In addition to the disaggregation by race group, Leibbrandt et al (2010) also provide a disaggregation by urban/rural geotype. Here they find that the Gini coefficient for rural residents increased from 0.58 in 1993 to 0.62 in 2000, before falling back to 0.56 in 2008. In contrast, the Gini coefficient for urban residents increased from 0.61 in 1993 to 0.64 in 2000 and again to 0.67 in 2008. Leibbrandt et al (2010) suggest that the respective changes in urban and rural Gini coefficients are largely due to the relatively large scale rural to urban migration (particularly of relatively poor individuals) over the time period, which has changed the population composition of urban areas. Leibbrandt et al (2010) also present some Gini coefficients reflecting inequality in expenditure, rather than income, and again find that South Africa registers very high levels of inequality compared to other countries in the world[10]. The Gini analyses are also complemented by the calculation of General Entropy measures, and both GE(0) and GE(1) measures indicate that inequality increased progressively from 1993 to 2000 to 2008 and that within-race group inequality increased relative to between-race group over the period, irrespective of whether either the standard GE approach or the modified GE approach proposed by Elbers, Lanjouw, Mistiaen, & Özler (2008) was used.

The broad conclusions of Leibbrandt et al (2010) concerning the increasing level of inequality in South Africa prior to and/or since the advent of democracy largely reinforce other commentators who used a mixture of Census and survey-based analyses (see, for example, Fedderke et al., 2003; Hoogeveen and Özler, 2006; Simkins, 2005; van der Berg et al., 2005; Van der Berg et al., 2008), although the exact scores on the inequality measures do vary somewhat between the different studies.

[10] Although they note that expenditure-based measures of inequality tend to produce somewhat lower scores on the Gini coefficient than income-based measures of inequality.

Before moving on to highlight key research concerning poverty patterns and trends, it is first instructive to consider the actual income distribution across income deciles in South Africa. Leibbrandt et al (2010) present these data for 1993, 2000 and 2008 based upon their analyses of household survey data. Figure 2.3 below is derived from Leibbrandt et al (2010, p.25) and shows these three data series. Two main points are conveyed by this chart: first, the income distribution is highly skewed towards the top income decile, with this particular group accounting for over half the total income of the entire population; and, second, the income share of the top income decile increases from 1993 to 2000 and again from 2000 to 2008, while the middle and upper-middle income deciles see relative reductions in their income shares and the lowest income deciles see little change.

Figure 2.3: Income shares by national income decile: 1993, 2000 and 2008



	1	2	3	4	5	6	7	8	9	10
1993	0.3	1	1.7	2.2	3.2	4.3	6.2	9.6	17.7	53.9
2000	0.4	1.1	1.6	2.2	3	4	5.6	8.8	16.8	56.7
2008	0.4	1	1.5	2.1	2.8	3.7	5.4	8.6	16.6	58.1

Source: derived from data presented in (Leibbrandt et al., 2010)

Leibbrandt et al (2010) state that, based upon their detailed investigation of changing incomes across the income spectrum, it appears that the increasing share of income enjoyed by the top decile is in fact actually largely accounted for by the top vigintile. This particular group at the very top of the income distribution saw their share of total income rise from 38.2% to 41.1% to 42.6% across the three years considered. Although data are not presently available on a more detailed disaggregation of income shares *within* the top vigintile, other research internationally has found that the proportional share of income becomes increasingly extreme as one moves up through the very highest levels of the income distribution (see, for instance, Atkinson, 2015a; Dorling, 2014b).

Whereas there is a broad consensus that inequality has increased since the fall of apartheid, there is less agreement concerning trends in poverty over the same period. Leibbrandt et al (2010) begin by using the 1996 and 2001 Census data to examine the proportion of population living on incomes below a lower poverty line of $2 per day (converted to a Rand amount of R91 per person per month) and an upper poverty line of R250 per person per month. Their results from this Census-based analysis indicates “a slight but unambiguous increase in measured poverty between 1996 and 2001” (Leibbrandt et al., 2010, p.14). However, Leibbrandt et al (2010) also calculated the Foster-Greer-Thorbecke (FGT) poverty indices[11] for the upper and lower bound Hoogeveen & Ozler (2006) poverty lines using household survey data for 1993, 2000 and 2008. They found that poverty levels fell between 1993 and 2008 across all these measures, with the largest falls typically seen between 1993 and 2000, and smaller falls (or no change) between 2000 and 2008. The mixed results produced by Leibbrandt et al

[11] P0 is the poverty headcount ratio; P1 the mean poverty gap (as a percentage of the poverty line); and P2 the squared mean poverty gap. These measures put increasing emphasis on the poorest people.

(2010) are also seen across a range of other poverty studies. Statistics South Africa (2002), Hoogeveen & Ozler (2006) and Simkins (2005) all found that poverty rates rose between 1994 and 2000, whereas Adelzadeh (2004), Van der Berg & Louw (2004) and Van der Berg et al. (2005) all found that poverty either stabilised or even declined slightly over this period. With regards to the post 2000 situation, Van der Berg et al. (2005), Van der Berg et al. (2008) and Meth (2006; 2007) used competing methodological approaches but all concluded that poverty rates fell between 2000 and 2004. Both the work of Van der Berg and Meth identified social grants as a key factor behind the reductions in poverty since the year 2000. However, even if income poverty has stabilised or even fallen over the post-apartheid period, poverty is still highly unevenly spatially distributed (Noble et al., 2014) and thus very unequally experienced – a theme to which I will return later in this chapter.

Before turning to consider patterns and trends in indicators of wider dimensions of deprivation, it is first informative to briefly consider some of the factors, such as social grants, that have been identified as potentially driving the changes in poverty and inequality in South Africa. Although I do not explicitly consider the effects of social grants on people's lived experiences of inequality in this thesis, I acknowledge the potential important role that such grants can have in redistributing income across the income range. South Africa has introduced or deracialised a range of means-tested social grants since the fall of apartheid, including (but not limited to) the Child Support Grant, Old Age Grant, Care Dependency Grant, Foster Child Grant and Disability Grant.

Leibbrandt et al (2010) examined the composition of household income across national income deciles and found that wages constitute by far the greatest component of household income across the upper income deciles in 1993, 2000 and 2008. In the top two deciles of the income distribution wages account for around 80% or more of total household income. However, as one moves down through the income deciles in any of the years the contribution of wages to total household income falls so that approximately only 20% to 35% (depending on the year) of total household income amongst the lowest income decile is obtained through wages. In 1993 approximately half of the total household income amongst the lowest income decile was accounted for by remittances, however, this proportional share reduced over the period to 2000 and again over the period to 2008 so that less than 10% of total income was due to remittances in 2008. An important factor in the changing composition of income source amongst the income distribution, and particularly amongst those in the lower deciles of the distribution, was the growth in social grants. By 2008, approximately two-thirds of total household income amongst the lowest income quintile was accounted for by social grant receipt, primarily the Child Support Grant (Leibbrandt et al., 2010). The impact of social grants on the incomes of the poor and therefore on the levels of poverty since 2000 has also been identified by other commentators too (Aguero et al., 2007; Meth and Dias, 2003). Leibbrandt et al (2010) contend that although social grants are likely to have made an important contribution to poverty levels through providing a regular (albeit small) income stream to very poor households, the impact of social grants on inequality measures has been neutral. Rather, their view is that the main determinant of inequality levels is inequality in labour market outcomes, both in terms of difficulty in accessing labour market opportunities (especially amongst individuals in the lower income deciles) and the high degree of wage inequality amongst those who are engaged in the labour market. There is, however, evidence that social grants may have positive effects on wider social outcomes such as school attendance, health status and nutritional outcomes (e.g. Leibbrandt et al., 2010) although a detailed discussion of this is outside the scope of this thesis.

In summary, therefore, there is a general consensus that inequality increased from the mid-1990s to the end of the century and into the next decade, although the magnitude of the increases vary from study to study. The findings with regard to poverty levels since the turn of the twenty-first century are more mixed, although there is again a general consensus that overall aggregate poverty has fallen over this period, most probably due to the expansion of social grant provision.

Bhorat and Kanbur (2005) provide a helpful overview of poverty and wellbeing in the first decade of post-apartheid South Africa. They highlight three key issues: (i) data quality and comparability have made it difficult for any sort of consensus to be reached on how outcomes for individuals and households have changed over the period; (ii) while the outcomes on poverty, inequality and unemployment appear quite poor, there are more encouraging outcomes on social indicators and access to services; and (iii) the prospects for rapid and sustained economic growth (which they argue are needed to address poverty, unemployment and inequality in the long-term) are themselves negatively affected by increasing inequality, poverty and unemployment. They also highlight the importance of geography to poverty, deprivation and inequality in South Africa: “[The] role of space in reinforcing or potentially overcoming South Africa’s welfare challenges is likely to feature increasingly in future policy debates” (Bhorat and Kanbur, 2005, p.5).

Whilst there does seem to have been some progress against the objective of reducing poverty, there is evidence of greater progress in terms of outcomes concerning access to basic services. Leibbrandt et al (2005); Bhorat et al (2006); and Woolard & Woolard (2006) all found measurable improvements in indicators relating to housing, water and electricity since 1994. For example, using data from the 1996 and 2001 Censuses, Leibbrandt et al (2005) showed improvements over this period in dwelling type (i.e. a reduction in traditional dwellings and an increase in formal dwellings), toilet facilities, means of refuse disposal, source of water, and source of energy for cooking and lighting. However, as I will demonstrate, the improvements have not been spatially evenly distributed, contributing to considerable spatial inequality.

There has also been some research that suggests that improvements in basic services has measurable effects on other economic and social outcomes. For instance, Dinkelman reported that the electrification of poor rural areas of KwaZulu-Natal resulted in an increase in female labour force participation of 9 percentage points, and subsequent increases in earnings and consumption. These outcomes were achieved through female household members being able to reallocate some of their time through the use of time-saving electric appliances (Dinkelman, 2011).

The discussion of patterns and trends in inequality, poverty and deprivation provided above draws almost entirely on national level measures. Although some decomposition of inequality levels and poverty rates by urban/rural status has been undertaken by Leibbrandt et al (2010) and others, this says little about the spatial patterns of inequality, poverty and deprivation within South Africa. As I argued in Chapter 1, my contention in this thesis is that people's *lived experience of inequality* is influenced by where they live and where they carry out their routine daily activities. By extension, I contend that the spatial configuration of poverty and deprivation that people experience on a regular basis is an important determinant of their lived experience of inequality. In light of this, I will now review empirical research in South Africa that takes a sub-national approach.

2.3 Sub-national level poverty and deprivation within South Africa

As I have indicated earlier, income poverty and deprivation are not evenly distributed across South Africa. The various research highlighting the spatial patterning of poverty and deprivation gives a vivid visual indication of unequal spatial distribution. In general, these studies have provided a comparative account at a given point in time. The spatial distribution reflects historic colonial, segregation and apartheid racial policies which have effectively concentrated poor people in former homelands or peripheral townships (e.g. Barnes et al., 2009; Christopher, 1994; Noble et al., 2009a; Noble and Wright, 2013b; Noble et al., 2013). Despite the removal of internal migration restrictions at the beginning of the post-apartheid period, the resulting migration patterns have not had a *dramatic* impact on the spatial patterning of deprivation (e.g. Noble and Wright, 2013a; Wright and Noble, 2013). Helpful as it is to observe the spatial differentiation cartographically and conclude that there is indeed spatial inequality in the experience of poverty and deprivation, there is an obvious need for a rigorous empirical measure to quantify this unequal distribution, which is the focus of the thesis.

The main focus of Chapter 3 will be to re-analyse the data pertaining to one particular sub-national measure of deprivation, namely, the South African Index of Multiple Deprivation 2001 at Datazone level (SAIMD 2001) (Noble et al., 2009a). This will entail a detailed account of the literature relating to that particular index. However, it is important to set that study in context and to review more broadly the literature relating to the spatial distribution of poverty and deprivation across South Africa and this is the focus of this section of this chapter.

In recent years a number of analyses of poverty and deprivation have been undertaken at small area level in South Africa. These approaches include accounts of income poverty and unemployment, as well as measures which bring together a series of indicators at small area level to create multidimensional measures of deprivation at small area level.

One of the earlier studies to emerge during the democratic era was undertaken by Statistics South Africa and utilised the World Bank's poverty mapping methodology (Alderman et al., 2002; Alderman et al., 2000). The World Bank pioneered the expenditure based small area estimation techniques using household unit data (e.g. Elbers et al., 2003; Ghosh and Rao, 1994; Hentschel and Lanjouw, 1996). This approach uses nationally representative income/expenditure surveys and household Censuses to produce estimates of households falling below an expenditure poverty line and has been used by the World Bank and partners in a number of countries (e.g. Benson, 2002; Demombynes and Ozler, 2006; Hentschel and Lanjouw, 1996). Alderman et al (2002) used the 1996 census in combination with the South African Income and Expenditure survey 1995 and the South African October Household survey 1996, to model and map estimates of household expenditure (as a proxy for income) at various spatial levels. However, the estimates became less reliable at small spatial scales. Demombynes and Ozler (2006) also used the World Bank's small area estimation technique to construct estimates of per capita expenditure at police precinct level using data drawn from a combination of the 1996 Census, 1996 October Household Survey and 1995 Income and Expenditure Survey. They proceeded to calculate indices of inequality within and between police precincts using a General Entropy measure (specifically, the GE(0) measure). They employed their per capita expenditure estimates and their inequality indices as independent variables (alongside other variables such as population density calculated directly from the Census) to test for statistical relationships between crime-related outcomes and potential explanatory factors and found that burglary and vehicle theft were most likely to occur in precincts with high intra-precinct inequality and high mean expenditure, whilst violent crimes were more likely to occur in areas where inequality was high between neighbouring precincts[12].

At around the same time, and inspired by similar indices constructed in the UK by Peter Townsend (Townsend et al., 1988), McIntyre et al (2000) produced four alternative deprivation indices at magisterial district level using data from the 1996 Census in order to look at the relationship between deprivation and health inequalities in South Africa. The indices produced included a general index of deprivation containing a number of variables relating to socio-economic, demographic and physical household characteristics. They also produced an index of at risk groups and two other indices.

The HDI discussed above has also been produced at a sub-national level in South Africa, initially for the Western Cape province by the Social Research and Population Development Unit of the Department of Health and Social Services in the Western Cape (Department of Health and Social Services, 1999). They combined four indicators – income, employment status, literacy and water supply (each created from single or multiple 1996 Census indicators) with equal weight to form a composite index. Subsequently they further developed this work (Department of the Premier of the Western Cape, 2005) using the 2001 Census to produce a HDI at municipality level which combines variables relating to a long and healthy life (life expectancy) with

[12] Unfortunately, police precincts are problematic geographical units for this type of analysis because they vary in size considerably (from approximately 370 residents in the lowest population police precinct to approximately 370,000 residents in the highest population precinct, based upon population estimates I derived from the 2011 Census for a separate piece of research), meaning that police precincts do not allow areas to be compared on a like-for-like basis.

variables relating to ‘knowledge’ (adult literacy and gross school enrolment) and a decent standard of living (using mean household income). This 2005 work also generated a set of indices which are combined to form a City Development Index (CDI) which was presented at Census Main Place level[13].

More recently, the Centre for the Analysis of South African Social Policy (CASASP) at the University of Oxford developed a series of indices of multiple deprivation for South Africa. These were based on earlier work they had undertaken in developing models of multiple deprivation and in constructing national indices of deprivation at small area level within the UK (e.g. Noble et al., 2005; Noble et al., 2000a; Noble et al., 2001; Noble et al., 2000b; Noble et al., 2004; Noble et al., 2003). In this work, small area level multiple deprivation was conceptualised for the first time as a weighted combination of distinct dimensions or ‘domains’ of deprivation.

The first small area level index of multiple deprivation produced by the CASASP team for South Africa was the Provincial Indices of Multiple Deprivation 2001 (PIMD 2001) produced by Noble et al (Noble et al., 2006a). The PIMD 2001 was based on 2001 Census data and was constructed at ward level. Due to systematic differences in the size of ward populations between provinces, wards were ranked within provinces and separate indices produced for each of the nine provinces. The PIMD 2001 consisted of five component domains of deprivation measured at ward level using the Census data: income and material deprivation; employment deprivation; education deprivation; health deprivation and living environment deprivation. The PIMD 2001 allowed a ranking of wards within each province on each of these five component domains. The five domains were then brought together to generate an

[13] See Chapter 3 for a discussion of the various South African geographies, including Census Main Place.

overall Index of Multiple Deprivation for each ward, which was also ranked across wards within each province.

In response to the problem presented by the ward level population sizes, Avenell et al (2009) developed a new statistical geography for South Africa which they named Datazones. This new statistical geography, which consisted of areas of standardised population size across the whole of South Africa[14], facilitated the re-construction of the PIMD 2001 at datazone level resulting in a new nationally consistent South African Index of Multiple Deprivation 2001 (SAIMD 2001) (Noble et al., 2009a). As I will elaborate in Chapter 3, I use the SAIMD 2001 at datazone level as the basis of much of the empirical development work and analysis in this thesis. In light of this, I will not go into detail here on the results from the SAIMD 2011 at datazone level, other than to say that, broadly speaking, levels of deprivation are seen to be extremely high across many parts of South Africa, particularly within the largely rural former homelands and the urban and peri-urban townships (Noble et al., 2009a). However, in some parts of the country (e.g. Cape Town), there is evidence of a heterogeneous spatial mix of deprivation levels amongst proximate neighbourhoods, with very high deprivation neighbourhoods and very low deprivation neighbourhoods either directly bordering one another or being located very close to one another. Simply mapping the data from the SAIMD 2011 at datazone level provides a visual indication of the unequal spatial distribution of deprivation across South Africa, and I explore this in more detail in Chapter 3.

In addition to the PIMD 2001 and the SAIMD 2001, which both measured deprivation across the entire age range of the population, Barnes et al (2007b)

[14] See Chapter 3 for a discussion of the different type of geographical unit in existence in South Africa, including ‘statistical geographies’.

developed an index using the same conceptual model but focusing specifically on deprivation experienced by children. This child-centred index, termed the South African Index of Multiple Deprivation for Children 2001 (SAIMDC 2001), again used 2001 Census data but was restricted to presenting results at Municipality level due to difficulties in securing access to the necessary data access permissions from Statistics South Africa that were required for sub-Municipality level analysis. Wright et al (2009a), however, subsequently re-constructed this child-centred index on the datazone geography once the relevant data access permissions had been obtained. Barnes et al (2009) provide a valuable account of the spatial distribution of child deprivation across South Africa drawing upon the data generated through the two child-centred indices of deprivation described here.

Wright and Noble (2009) utilised the Community Survey 2007 to produce a Municipality level index of multiple deprivation for 2007, while Wright et al (2009b) used the Community Survey 2007 to produce a Municipality level index of deprivation for children for 2007. In both cases, comparable indices were also constructed for the 2001 time point to enable assessment of change over time at Municipality level. Both Wright and Noble (2009) and Wright et al (2009b) found that despite reductions in absolute levels of deprivation across all dimensions considered at national level, this was not the case for all Municipalities, with some areas getting worse on some measures.

A recent addition to the series of small area level multiple deprivation indices for South Africa consists of the SAIMD 2011 at ward level and an accompanying measure of income poverty also at ward level, both of which are based on the 2011 Census (Noble et al., 2013; Noble et al., 2014). The SAIMD 2011 follows the same broad methodological approach as the SAIMD 2011, but consists of four component domains: material deprivation; employment deprivation; education deprivation; and living environment deprivation. Whereas the SAIMD 2001 had included an ‘income and material deprivation domain’, in the SAIMD this had been replaced by one focused only on material deprivation, and the income poverty component was considered separately as a stand-alone measure (i.e. not included in the overall composite SAIMD 2011). The SAIMD 2011 and accompanying income poverty measures were restricted to ward level as it was not possible to reconstruct the datazone geography from the publicly available 2011 Census data and it was not possible to fit the 2011 Census data to the 2001 datazone boundaries. However, the 2011 ward rankings (which are ranked across the entire country, unlike the PIMD 2001) do provide an up-to-date account of the spatial distribution of deprivation and income poverty at sub-national level. As was the case with its predecessors, both the SAIMD 2011 and the ward level income poverty measures continue to show the stark spatial patterning evident in 2001. The highest rates of income poverty and the highest scores on the SAIMD 2011 continue to be apparent in the former homeland areas and in the townships abutting major metropolitan areas. In the case of the latter the deprivation/poverty tends to be worst in informal settlements and is often relatively spatially proximate to affluent areas (Noble et al., 2013).

As noted above, I use the SAIMD 2001 at datazone extensively in this thesis and, indeed, it forms the basis of the vast majority of the empirical development work and analysis. I do not use the more recent ward level SAIMD 2011 or 2011 income poverty measures in this thesis. There are two main reasons why I used the SAIMD 2001 at datazone level and not the more recent SAIMD 2011 at ward level: first, on a technical level, the datazone geography is better suited for my purposes than the ward level geography because datazones are a statistical geography; and second, on a practical level, I had completed my empirical development work and analysis using the SAIMD 2001 at datazone level prior to the publication of the SAIMD 2011 at ward level. However, as noted above, the general spatial pattern of deprivation in 2011 was much the same as was observed in 2001, and so the results and conclusions in this thesis, although referring to 2001 data, will still have academic and policy relevance today and into the future.

As noted briefly above, simply mapping the levels of deprivation at sub-national level across South Africa gives an effective *visual* indication of the unequal spatial distribution between and within different areas of the country. I explore this in greater detail using a more structured analytical approach in Chapter 3. I argue that this geographical patterning of deprivation provides the spatial context for people's experiences of socio-economic inequality as they go about their daily lives.

In order to test for associations between people's lived experience of inequality and other social outcomes it is necessary to move beyond the analyses of broad spatial patterns in deprivation and look instead at specific empirical measures that quantify the nature and extent of spatial inequality. The majority of spatial inequality research at neighbourhood level draws upon measures of 'residential segregation', and I now turn to introduce this particular literature.

2.4 Residential segregation

Residential segregation measures express the degree to which members of one population category are spatially segregated from those of another category (or categories). Indices of residential segregation typically quantify the extent to which population composition varies *between* neighbourhoods that lie within a larger geographical area (e.g. within a city, local municipality or some other geographical bounds). These indices typically utilise categorical measurement variables (such as a distinction between different racial groups). Categorical data are often available at small area level, whether from Census or administrative data sources (although, as noted above, currently administrative data in South Africa is limited), and so residential segregation measures are ideally suited to neighbourhood level analyses. In contrast, the classical income inequality measures, such as Gini coefficients, are typically based upon continuous data relating to individual or household incomes or expenditures collected through national surveys. Whilst national surveys are often designed to be statistically representative at national (or sometimes provincial) level, sampling error usually precludes their use for the production of small area level indicators. For my purposes therefore, given that I am interested primarily in how experience of inequality varies by place, measures of residential segregation are the most appropriate.

There is an extensive international literature on measures of residential segregation, which began with a focus on racial segregation in US cities in the 1950s (e.g. Duncan and Duncan, 1955), and has now expanded in both geographical breadth and methodological depth to cover numerous empirical developments and applications across many countries. One key development is that residential segregation indices have now been applied to study the extent and nature of spatial inequality in socio-economic outcomes (e.g. segregation by social class, Dorling, 2014a) as well as the traditional focus on racial segregation. For the purpose of this thesis I take one particular classic study of residential segregation as my starting point: Massey and Denton's (1988) article *The dimensions of residential segregation*.

Massey and Denton constructed and reviewed twenty indices of residential segregation using data on the racial composition of neighbourhoods in 60 metropolitan areas across the United States. Their analyses examined the degree to which the minority non-white population was spatially segregated from the majority white population. They identified five conceptually distinct dimensions of residential segregation into which they grouped their measures: evenness, concentration, centralisation, exposure and clustering. Based upon a combination of empirical testing and researcher judgement, they selected one measurement approach from each of the five dimensions which they recommended for adoption as principal techniques for research in the field of residential segregation. Although Massey and Denton's paper focused on the spatial distribution of racial demographic groups within urban areas in the United States, many of the techniques they discuss can be applied to the study of socio-economic factors across both urban and rural areas of other areas, such as South Africa.

More recently, Reardon & O’Sullivan (2004) have argued that four of the five dimensions proposed by Massey & Denton are, in fact, all variants of a single general measurement concept. They propose that residential segregation measures can be simplified into a distinction between just two dimensions: (i) exposure, and (ii) evenness/clustering. According to Reardon & O’Sullivan, centralisation and concentration can be regarded as specific subcategories of evenness/clustering. However, they acknowledge that “in some cases, centralization and concentration may be of sufficient theoretical interest to be considered distinct subdimensions” (Reardon and O’Sullivan, 2004, p.127). Although Reardon & O’Sullivan (2004) suggest that Massey and Denton’s five dimensions of residential segregation can perhaps be collapsed into just two, I argue here that, given the lack of existing research on spatial inequality in South Africa, there is merit in initially considering all five.

As I will elaborate in Chapter 4, Massey and Denton's classification of measures of segregation provides a valuable framework for considering which measure(s) might be suitable for reflecting the uneven spatial distribution of deprivation which I contend is an important basis for people's lived experience of inequality. In the consideration of each of the five dimensions I include a review of relevant empirical literature and, as such, I do not go into detail on that literature here.

In addition to the debate around conceptual distinctions between the proposed dimensions of segregation, it is also important to consider methodological distinctions in the way in which segregation indices can be constructed and presented with regard to spatial scale. Lloyd and Shuttleworth (2012) note that spatial scale can be handled in three different ways in the calculation of residential segregation indices. Standard global measures consider each neighbourhood separately (i.e. ignoring spatial inter-relationships between neighbourhoods) and generate a summary segregation statistic for the entirety of the larger geographical study area (e.g. for each local municipality). Geographically weighted measures do take account of the spatial inter-connectedness of neighbourhoods within the larger geographical study area but still generate a summary segregation statistic for the entirety of the study area. Local measures also take account of the spatial inter-connectedness of neighbourhoods but generate a summary segregation statistic for each individual neighbourhood rather than a single summary statistic for an entire study area. Massey and Denton present their indices of evenness, concentration, exposure and centralisation as global measures, while they present their indices of clustering as geographically weighted measures. As will become evident in Chapter 4, I initially follow Massey and Denton's approach in this regard, whilst recognising that many of the global measures that they produced can also be modified to generate geographically weighted and/or local variants (Feitosa et al., 2007). Indeed, as will become apparent in Chapter 4, upon selection of my chosen segregation index, I proceed to develop a particular localised variant using a novel approach.

Whereas the measures proposed by Massey and Denton all assess the level of segregation between two constituent population sub-groups, other commentators have developed multi-group segregation indices. For example, Morgan (1975) developed a multi-group measure of segregation between different social class groups in the UK, while Wong (1998) developed a spatial variant of the multi-group segregation index and Reardon and Firebaugh (2002) derived and evaluated a range of multi-group segregation measures. Whilst I acknowledge here the contribution that these multi-group indices make to the residential segregation literature, they are not relevant for my purpose due to the nature of the data upon which my analyses are based (which is discussed in more detail in Chapter 4), and so I do not go into detail on these particular measures.

Although measures of residential segregation have been applied widely in developed country contexts, there have so far been relatively few studies utilising these measures in South Africa. The studies that have used such measures have typically employed the Dissimilarity Index which falls under Massey and Denton's dimension of evenness (see Chapter 4 for a more detailed discussion of this). For example, in South Africa, Christopher (1988) initially calculated dissimilarity indices of racial segregation at around the time of 'Union'[15] using the data from the 1911 Census at enumeration tracts and then proceeded to calculate equivalent indices of racial segregation using data from the 1921, 1936, 1951, 1960, 1970 and 1985 Censuses (Christopher, 1989). Unsurprisingly given the nature of colonial, segregationist and apartheid rule,

[15] 1910 saw the unification of four previously separate British colonies ('Cape of Good Hope', 'Natal', 'Transvaal' and 'Orange River') to form the single administrative entity of the Union of South Africa.

Christopher found that levels of racial segregation were remarkably high by international standards and that the degree of segregation tended to increase over time as the effects of the legislative enactments became increasingly severe. However, Christopher also found that despite these trends, “Rather surprisingly racial exclusivity has not been attained in any South African town...It is evident that where economically possible, White householders are unwilling to forego the services of a resident servant” (Christopher, 1989, p.259). Other research concerning racial segregation in South Africa includes Horn’s (2005) study in which he assessed a range of indices for measuring racial segregation in South Africa using data from the 1991 and 1996 Censuses and concluded that Wong’s (1998) spatial index of multi-group segregation was most appropriate. More recently, Parry and Van Eeden (2015) used Census data from 1991, 1996, 2001 and 2011 and calculated Theil's entropy indices for grid squares of varying spatial sizes across Cape Town and Johannesburg. They concluded that racial segregation levels decreased over time at each geographic scale assessed in both cities but, despite these observed decreases, both cities remain highly segregated in racial terms.

Whilst residential segregation measures have, therefore, been applied in a limited number of studies in South Africa, the purpose of their application has primarily been to examine racial segregation. The purpose of this thesis is to assess the utility of residential segregation measures for the purpose of measuring people’s day-to-day ‘lived experience of inequality’ in a socio-economic sense rather than a racial sense. In order to achieve this objective I draw upon data at neighbourhood level on levels of social and economic deprivation, specifically, the South African Index of Multiple Deprivation 2001 at Datazone level (SAIMD 2001) which I introduced in the discussion above. As I will elaborate in Chapters 3 and 4, the SAIMD 2011 provides a valuable resource for assessing the spatial distribution of deprivation across South Africa at a detailed geographical level, and it represents a suitable dataset upon which to construct measures of residential segregation.

The following five chapters consist of empirical analyses, interpretation and discussion. In Chapter 3, I introduce the SAIMD 2001 in more detail and present an analysis of the unequal spatial distribution of deprivation across South Africa at datazone level. In Chapter 4, I proceed to use data derived from the SAIMD 2001 to construct a measure of residential segregation which I argue represents a measure of people's day-to-day 'lived experience of inequality'. In Chapters 5 and 6, I present a range of empirical analyses of my selected segregation measure, and in Chapter 7 I present one example of how my selected measure can be used to test for associations with other socio-economic outcomes.

Chapter 3: The unequal spatial distribution of deprivation across South Africa

3.1 Aims and objectives

The aim in this chapter is to address sub-question Q1: To what extent is deprivation distributed unequally across neighbourhoods in South Africa? As highlighted in Chapters 1 and 2, there has been an increase in policy attention on sub-national distributions of deprivation over recent years and this has been accompanied by an increase in empirical research. However, the research base is still relatively sparse, due in part to the limited data sources available for mapping and tracking deprivation at neighbourhood level in South Africa. Following the review of possible data sources presented in Chapter 2, I conclude that the best measure of deprivation at neighbourhood level in South Africa is the SAIMD 2001 at Datazone level (Noble et al., 2009a). This particular data source is valuable both for describing the variations in the spatial distribution of deprivation across the country (which is the focus of this chapter) and also for producing statistical measures of residential segregation for the purpose of measuring people's 'lived experience of inequality' (which is the focus of Chapter 4).

Before commencing analysis of the SAIMD 2001, I will first provide important contextual information concerning the basis upon which it was developed. Geography is a particularly important component to any sub-national measure of deprivation as the choice of geographical unit of analysis will affect the results obtained. This 'modifiable areal unit problem' (MAUP) has been the subject of a great deal of research, with

Openshaw's (1984) book of the same name regarded as a seminal contribution. A crucial feature of the MAUP is that, typically, as the population size of a given spatial unit decreases, the composition of the population *within* each given unit becomes more homogenous, while the composition of the population *between* areas becomes more heterogeneous (Openshaw, 1984). This, of course, is related to Tobler's 'first law of geography' which states that "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970, p.236) in the sense that people tend to live close to other people of similar deprivation status (due to a wide range of factors, including housing availability and affordability). The objective here is not to look in detail at the effects of the MAUP on the analyses, but rather to acknowledge the issue and to reference it in my consideration of the best geographical unit to utilise in this thesis.

In order to be 'fit for purpose' in this thesis, the geographical unit of analysis must ideally meet three criteria: (i) be of relatively standardised population size to permit comparisons between areas on a like-for-like basis[16]; (ii) be of small enough population size to permit detailed spatial analyses; and (iii) maximise homogeneity of population characteristics within each area such that each unit may be regarded as representing a 'meaningful' neighbourhood. In Section 3.2 below I discuss the various geographical layers available from Statistics South Africa, the inter-relationships between them and their respective strengths and weaknesses for my analytical purpose.

[16] If the constituent units of a geographical layer vary considerably in terms of population size then it may be the case that an area with a 'larger' population size may, in effect, contain the equivalent of multiple areas of 'smaller' population size. Any statistics calculated for the 'larger' area would, in effect, be averages of the results for the different parts of that larger area, which may be quite heterogeneous. My purpose in this thesis is to develop a measure of the lived experience of inequality that can be linked to other individual- and area-based socio-economic outcomes and so my goal is to develop my measures on a geography of standardised population size in order to permit spatial comparisons on a like-for-like basis.

I then proceed to discuss the development of the datazones and justify the decision to use these datazones as the chosen unit of analysis in this thesis.

In Section 3.3 I focus specifically on the SAIMD 2001 at datazone level and I describe the definitions of the indicators and domains and the methodology for building the overall composite index. I also detail the process through which I, as a CASASP researcher, was granted permission by StatsSA to work on the full 100% individual level 2001 Census within a secure data processing environment in Pretoria to perform the data development work for the SAIMD.

I present empirical analyses in Sections 3.4 to 3.7. In each of these sections the aim is to address the research question stated above, in terms of examining the extent to which deprivation is unequally distributed across South Africa. In Section 3.4 I provide a short overview of deprivation rates at national, provincial and municipal levels to set the context for the more detailed neighbourhood level analyses contained in Sections 3.5, 3.6 and 3.7. The focus within Section 3.5 is on assessing variation in neighbourhood levels of deprivation between and within provinces and municipalities, first looking at the overall composite SAIMD and then proceeding to concentrate on the four component domains of deprivation that are expressed as deprivation rates (as these four domains represent the measures which I utilise as input data for the residential segregation measures I develop in Chapter 4). The focus in Section 3.6 is on examining the extent of commonality or difference in neighbourhood characteristics between the most deprived neighbourhoods and the rest of the country in order to assess whether there are certain ‘types’ of neighbourhood where deprivation is most pronounced. The final empirical part of this chapter in Section 3.7 considers the extent to which provinces and/or municipalities exhibit a spatial mix in neighbourhood deprivation levels, in other words, assessing whether there are areas where neighbourhood deprivation levels are relatively homogenous or relatively heterogeneous. Section 3.8 contains a chapter conclusion in which I highlight the key findings that enable me to address the specified research question.

3.2 South African geographies

Internationally, there are typically three types of geographical units for which socio-economic data may be available at sub-national level: (i) political/administrative; (ii) Census enumeration/output; and (iii) statistical. Political/administrative geographies tend to consist of spatial units with relatively large populations as their primary purpose is to underpin democratic governance and service delivery. Census enumeration geographies, on the other hand, tend to consist of spatial units with relatively small populations as their primary purpose is to facilitate the process of collecting Census data from every household in the country which requires extremely localised logistical planning and monitoring. Finally, statistical geographies are those designed for the specific purpose of comparing areas on a like-for-like basis and, as such, tend to exhibit relatively small populations with relatively homogenous population characteristics. The three categories are not *necessarily* mutually exclusive, as it is potentially feasible for political/administrative and/or Census enumeration/output geographies to also satisfy the criteria for statistical geographies, although this is not the case in South Africa. The distinction between political/administrative and Census enumeration/output geographies is also not completely clear as Census statistics are often released for political/administrative regions. However, for my purpose here, I impose a classification which regards geographies as political/administrative if they were designed to serve that function, irrespective of whether they are also used for the publication of Census statistics or not.

Geographical boundaries relate to a particular point in time and, as such, are potentially subject to change as time progresses. Certainly, political and administrative geographies are subject to change in response to changing electoral and/or service delivery demands. Census enumeration geographies, on the other hand, are fixed at the time of the respective Census but may (and indeed do, in South Africa) change for each iteration of the Census. Statistics geographies should ideally remain fixed over time to permit temporal analysis of patterns and trends, but even these need to be revisited periodically to ensure that the criteria upon which they were designed still holds true in light of ongoing population change.

A further important consideration is the spatial relationships between geographies. Some geographies nest within other higher-level geographies, whilst other geographies are not coterminous with any other geography. The issue of spatial hierarchies and spatial nesting of geographies is an important consideration for the choice of analytical techniques when summarising neighbourhood level deprivation scores as ideally the neighbourhoods will nest within the higher-level geographies for which results are summarised[17].

For the purpose of this thesis, the focus is on geographies as at the time of the 2001 Census.[18] The empirical work undertaken for this thesis was completed prior to the publication of the 2011 Census and so this more recent resource was not available at that time.

[17] Where nesting does not occur it may be possible to assign neighbourhoods to higher-level geographies using a proportional area approach, although the validity of this approach is dependent upon the degree of overlap in the respective boundaries.

[18] Although I do utilise certain 2011 Census geographies for descriptive purposes in Chapters 5 and 6.

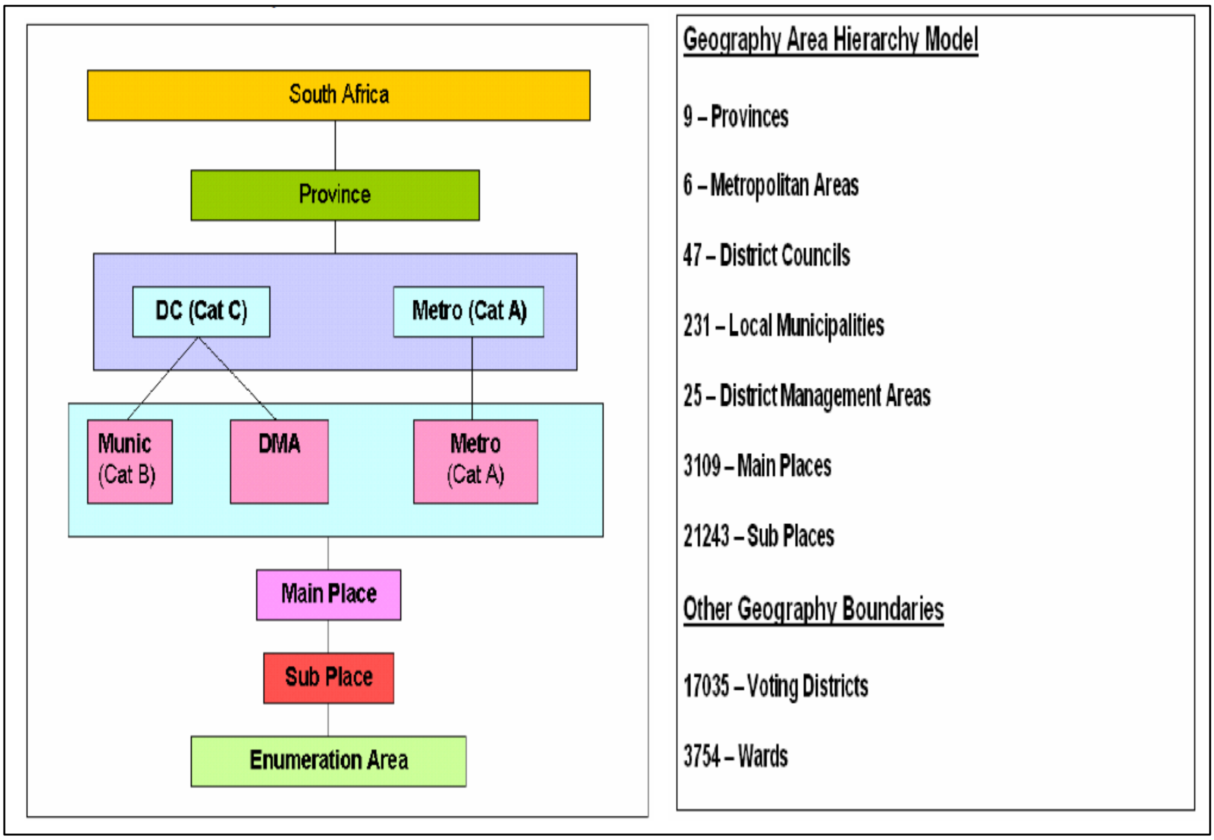
Figure 3.1 below shows the hierarchy of official South African geographies for which Census results are collected and/or made available by Statistics South Africa. Following on from the discussion above concerning distinctions between the three categories or geography, I regard province, district/local/metropolitan municipality and ward as political/administrative geographies, even though Census statistics are also released for these areas. I regard Census Main-Place, Sub-Place, Small Area Layer (SAL) and Enumeration Area (EA) as Census enumeration/output geographies because these areas were defined either for the explicit purpose of enumerating the Census (EAs) or based solely on the results of the Census (Main-Places, Sub-Places and SALs).

As would be expected, in 2001 the nine provinces nested perfectly within the national boundary. However, the District Councils and Local Municipalities did not always nest perfectly within the province boundaries, though that particular anomaly was resolved by 2011. Wards are the smallest political/administrative geography and these areas nest within municipalities but do not relate to Main place or sub place. In terms of the Census enumeration/output geographies, the EAs were designed to nest within municipalities and were defined prior to commencement of the Census data collection. The SAL, Sub-Place and Main-Place geographies were defined post-enumeration by grouping the EAs into progressively larger areas. There is a therefore a clear hierarchical nesting from EA through to Main-Place. The majority of Main-Places nest within municipal boundaries but there are some cases where Main-Places and Sub-Places straddle municipality boundaries. In these cases the areas have been sub-divided and given unique identifier codes to ensure that the sub-divided geographical units do nest with the higher-level geographies. Importantly, the boundaries of the wards are not coterminous with the boundaries of the Main-Places, Sub-Places or SALs or EAs.

None of the geographies listed in Figure 3.1 meet the criteria for classification as a 'statistical' geography. It is for this reason that Noble et al (2009a) were motivated to develop datazones to meet the need for a robust statistical neighbourhood geography.

Before discussing the development of the datazones in detail, I will first provide a short account of the background and functions of the various geographies listed in Figure 3.1 as this is of relevance to the discussion of the residential segregation measures in Chapter 4.

Figure 3.1: Statistics South Africa's Geography Area Hierarchy Model

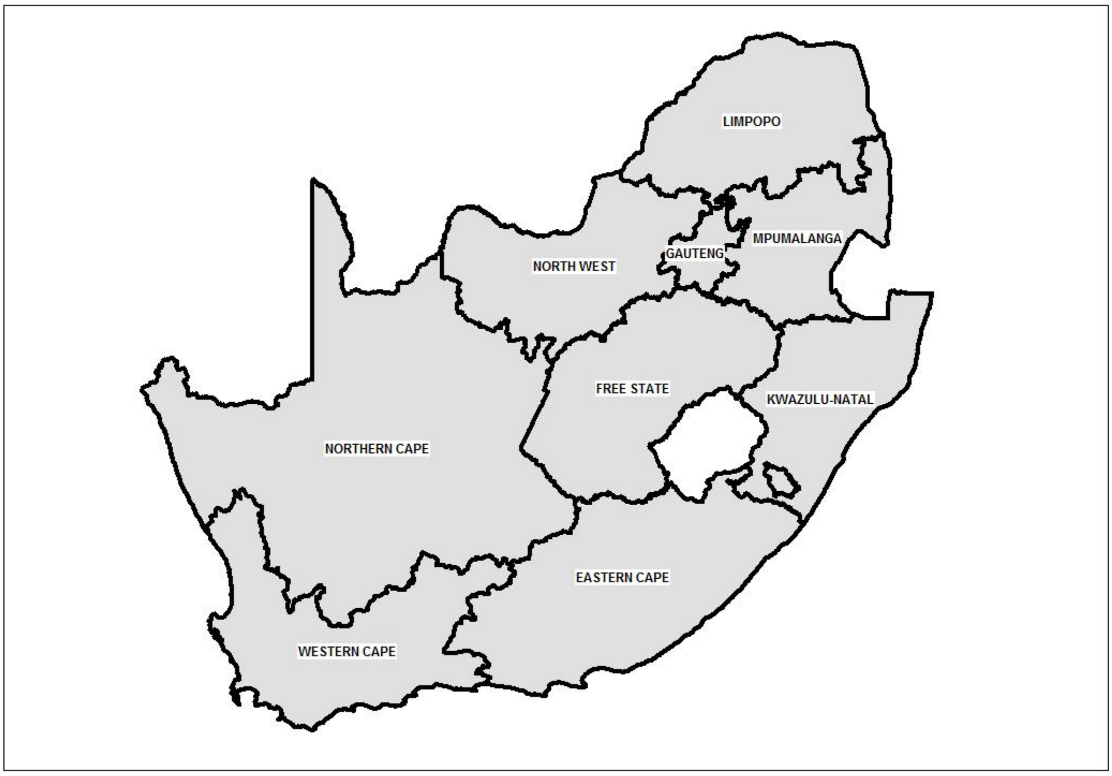


Source: (Statistics South Africa and HSRC, 2007, p.46)

Provinces

As of the time of the 2001 Census, South Africa was divided up into nine provinces, as shown in Figure 3.2. Present-day South Africa is also divided up into these nine provinces, although a number of small boundary changes have been implemented in the years since 2001.

Figure 3.2: The boundaries of the nine provinces in South Africa



Municipalities

At the time of the 2001 Census there were 6 metropolitan authorities - Cape Town, Johannesburg, Tshwane (Pretoria), Ekurhuleni (the East Rand), eThekwini (Durban) and Nelson Mandela Bay (Port Elizabeth). Outside of the metropolitan areas, the rest of the country was divided into 47 district municipalities and 25 District Management Areas (DMAs) which comprised special areas such as national parks that had low or no populations. The 47 district municipalities were themselves subdivided into a total of 229 local municipalities. By the time of the 2011 Census, a further two local municipalities had been classified as metropolitan areas: Buffalo City and Mangaung. Where the term 'local municipality' is used in analyses within this thesis, this refers to the group of metropolitan municipalities plus the non-metropolitan local municipalities.

Wards

As noted above, the smallest political/administrative geographical units in existence in South Africa in 2001 were electoral wards. These nest within local municipalities but are otherwise outside the main Census geography hierarchy. That is, they are not precise aggregates of EAs. There were 3,799 wards in South Africa in 2001 and these ranged in population size from those with less than 200 residents to those with more than 80,000 residents, with a mean population size of 11,475 and median population size of 9,136. As noted in Chapter 2, CASASP, with colleagues from the Human Sciences Research Council (HSRC) and Statistics South Africa (Stats SA), had produced ward level Provincial Indices of Multiple Deprivation (PIMD) for each province in South Africa (Noble et al., 2006a; 2006b). Electoral wards were the best sub-municipality level geographical unit available at the time. Their use was recommended by Statistics South Africa as, although their populations varied, the variation was much less than for Main Place or Sub Place. Furthermore they were a conventional tessellated geography whereas both Main Place and Sub Place could be configured in unusual ways especially in former homelands (Avenell et al., 2009). For example, one Sub Place might surround a number of other Sub Places – like a sea surrounding a group of islands. However, wards vary greatly by population size and so are not an optimal unit for this purpose, particularly for comparing areas across South Africa[19] (see Noble et al., 2006a). Having produced the PIMD, CASASP highlighted the need for a new *statistical* (not political/administrative) geography to better delineate pockets of deprivation and containing similar numbers of people so that each area could be compared to all other areas in the country.

Main Place, Sub Place, SAL & EA

EAs were an input geography. They were the basic units of collection of Census information – the areas covered by a single enumerator. Because of issues of disclosure, data were not published at that level but rather a new geography – the Small Area Layer (SAL) was created to allow limited publication of data. If an EA was of sufficient size then it qualified as a SAL in its own right – around 70% of EA's were SALs. The remaining SALs were created by amalgamating smaller EAs. EA's were also aggregated to form areas typically larger than SALs - Sub Places. In turn these sub places were aggregated to form the Main Places. Main Places nest within local municipality/metros.

[19] Indeed this was the reason why an index of multiple deprivation was separately created at ward level for each province rather than an overall index of multiple deprivation for the whole country, at the time.

Datazones

The datazones were developed as a specific response to the weaknesses in existing South African geographies. The current datazones use Census 2001 enumeration areas (EAs) as the building blocks (the units from which datazones could be constructed) to create a standard *statistical* geography. EAs were selected as the building blocks as it was considered that they gave greater flexibility for creating a higher level geography and would result in a better end product than other available options such as the small area level geography. A datazone comprises one or more contiguous EAs which share common characteristics. The creation of datazones involved complex geographical programming (Avenell et al., 2009). The various steps in the process of datazone construction were specified as a series of rules. The aims for the new geography were that datazones should:

i. have a similar population size, of between 1000 and 3000 and target size of 2000;
ii. delineate pockets of deprivation by maximising datazone social homogeneity and population density homogeneity (to prevent urban areas being merged with more rural areas that are adjacent);
iii. be a manageable geography, placing controls on datazone size (to prevent creation of very large datazones) and shape (to minimise complex shapes);
iv. nest within existing Census municipality and province geographies; and
v. form a continuous geography, constructed from contiguous lower level building blocks (EAs).

There are 22,846 datazones covering the entire country (compared to 3,799 wards). However, a small number of these datazones were judged to be unreliable for poverty/inequality analysis as they contained small residential populations (e.g. areas with large institutional populations or District Management Areas[20]). Removing these problematic datazones resulted in final set of 22,164 usable datazones across South Africa. The datazones nest within municipality boundaries as they existed at the time of the 2001 Census.

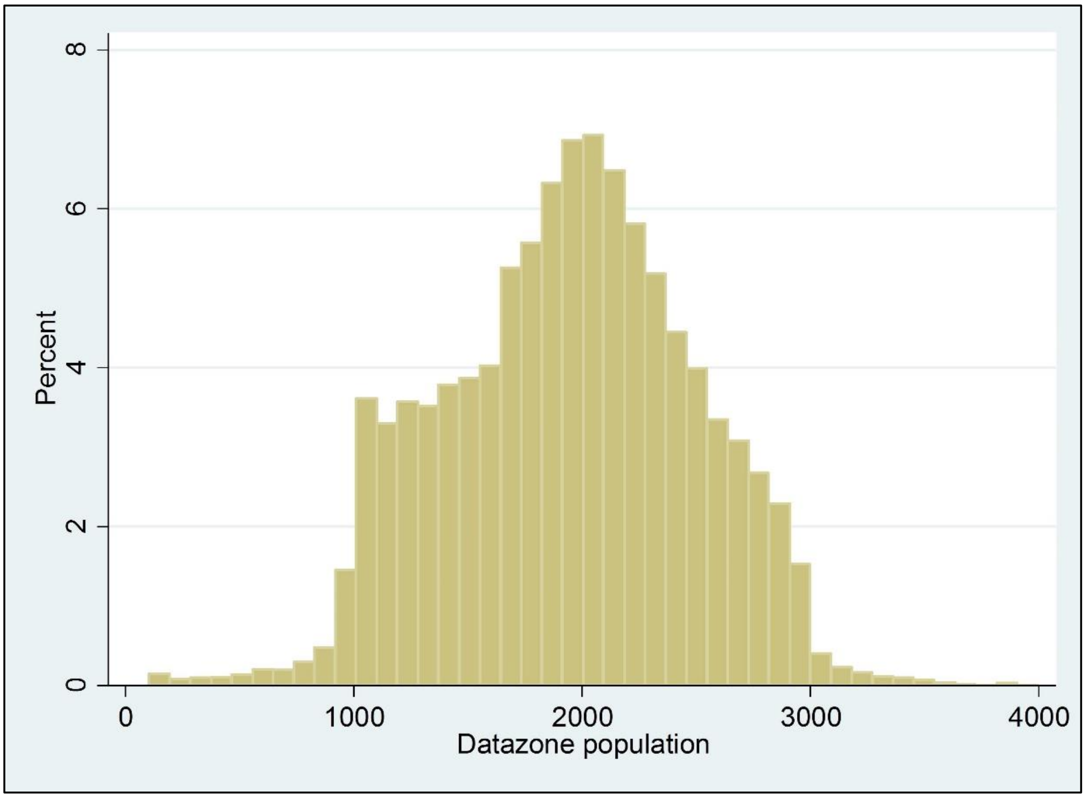
Figure 3.3 shows a histogram of datazone population sizes for the 22,164 areas retained for analysis. It is evident from this chart how standardised the population sizes are, thereby resulting in a statistical geography suitable for assessing deprivation levels on a like-for-like basis.

Figure 3.4 shows the datazone boundaries for the City of Cape Town, with Main Place boundaries overlaid[21]. It is evident from Figure 3.4 that the datazones provide a detailed geographical unit of analysis.

[20] District management areas are areas such as game reserves and mining complexes with small populations with special characteristics. They produce anomalous results and are customarily excluded by Stats SA from small area analyses.

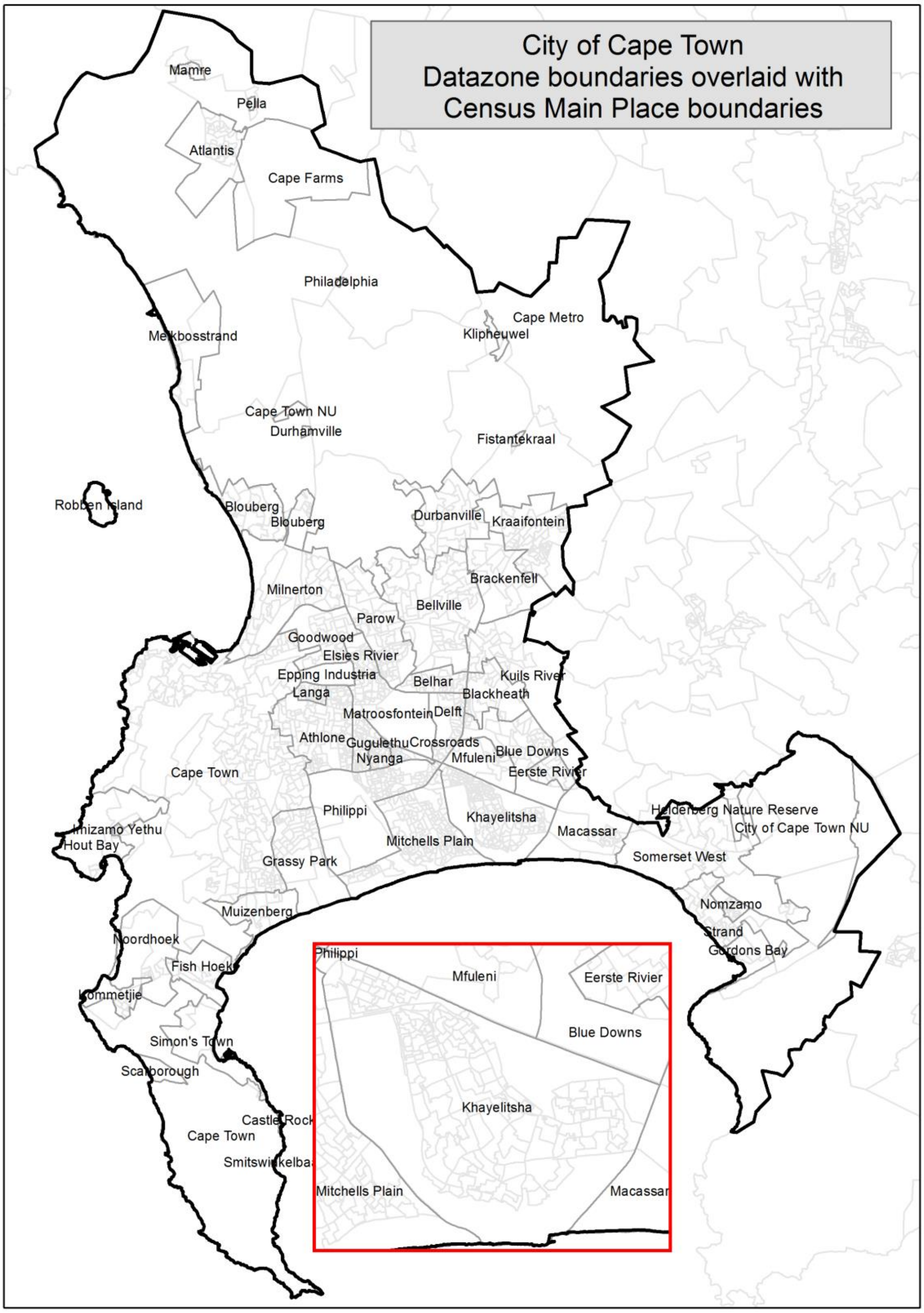
[21] For instance, the darker grey line around the Khayelitsha Main Place defines that boundary, and the lighter grey lines within Khayelitsha show the datazone boundaries.

Figure 3.3: Histogram of Datazone population counts, 2001



Note: the x-axis is capped at 4000 in this chart for clarity. However, 55 datazones had populations exceeding 4,000. These 55 datazones represented just 0.2% of the total datazones across the country. Of these 55 datazones, 35 had populations between 4,000 and 6,000, and the remaining 20 datazones had populations of more than 6,000 but less than 12,000.

Figure 3.4: The Cape Town metropolitan municipality, showing datazones boundaries overlaid with Census Main Place boundaries



In summary, therefore, the construction of the datazone geography was motivated by the need for a robust statistical geography at neighbourhood level which was not provided by the existing official geographies developed and supplied by Statistics South Africa (or indeed any other agency in South Africa). The datazones meet the criteria for a statistical geography and I therefore adopt this geography as the primary unit of analysis for the remainder of this thesis. The higher-level geographies discussed here are also utilised in the analysis throughout this thesis, but primarily as a means of grouping the datazones to look at patterns between and within larger areas.

I will now proceed to detail the primary data source underpinning the analyses in this chapter and in subsequent chapters, namely the South African Index of Multiple Deprivation 2001 at datazone level.

3.3 SAIMD domains, indicators and process of construction

I introduced the concepts behind the SAIMD 2001 in the discussion of small area level measures of deprivation towards the end of Chapter 2. Here, in Section 3.3, the focus is on providing detail on the definitions of the component indicators and domains, summarising the main stages in the methodology, and describing the actual physical process through which I developed the SAIMD dataset within a secure data setting at Statistics South Africa in Pretoria.

The five dimensions or domains of deprivation and component indicators in the SAIMD 2001 are as follows (derived from Noble et al 2009 Appendix 1 pp 35-39):

Income and material deprivation domain

- Number of people living in a household that has a household income (need-adjusted using the modified OECD equivalence scale) that is below 40% of the mean equivalent household income; or
- Number of people living in a household without a refrigerator; or
- Number of people living in a household with neither a television nor a radio.

Employment deprivation domain

- Number of people of working age (15-65) who are unemployed (using official definition); plus
- Number of people of working age (15-65) who are not working because of illness or disability.

Health deprivation domain

- Years of Potential Life Lost

Education deprivation domain

- Number of 18-65 year olds (inclusive) with no schooling at secondary level or above.

Living environment deprivation domain

- Number of people living in a household without piped water inside their dwelling or yard or within 200 metres; or
- Number of people living in a household without a pit latrine with ventilation or flush toilet; or

- Number of people living in a household without use of electricity for lighting; or
- Number of people living in a household without access to a telephone; or
- Number of people living in a household that is a shack; or
- Number of people living in a household with two or more people per room.

Therefore, with regards to the income and material deprivation domain and the living environment deprivations domain, individuals are counted as deprived if they meet *any one* (or more) of the constituent indicator definitions.

For the purposes of simplicity, the full domain names listed above will also be referred to in the following abbreviated form:

Income and Material Deprivation	=	INC
Employment Deprivation	=	EMP
Health Deprivation	=	HEA
Education Deprivation	=	EDU
Living Environment Deprivation	=	LIV

In order to construct the indicators at datazone level, special permission was obtained from Statistics South Africa for me, in my capacity as a CASASP researcher, to access the full 100% individual level 2001 Census data. The full 100% Census data were contained within four separate data files: (i) the 'person' file which contained answers relating to individual respondents; (ii) the 'household' file which contained answers relating to the household; (iii) the 'mortality' file which contained details of all deaths that occurred within the 12 months prior to census date; and (iv) the 'geography' file which linked each household to the standard census enumeration geographies. Each individual person had a unique 'person reference number' and each household had a unique 'household reference number'. Individuals were linked to the other members of their household using the household reference number. The household was spatially referenced by linking to the geography file using the household reference number[22].

The SAIMD 2001 data preparation work using the 100% Census was undertaken within the Statistics South Africa offices in Pretoria, Gauteng Province. Due to the size of the dataset it was necessary to use the SAS software package. Prior data testing and writing of SAS code had been undertaken by myself and my CASASP colleagues using the publicly available 10% anonymised sample of the Census. This enabled my work on the full 100% Census dataset to be completed within one week.

The INC, EMP, EDU and LIV domains were each constructed by creating binary flags within the individual level Census dataset to show whether or not each person was deprived on each domain. If a person was defined as deprived on a domain he/she would be coded as a '1' on that domain, whereas if the person was not defined as deprived on a domain then he/she would be coded as '0' on that domain. These binary flags would form the numerators for the domain rates. Another series of binary flags was created to show whether or not each person could *potentially* be deprived on each domain and this information was used for the domain rate denominators. For instance, it would not be appropriate to count children in the denominator for the EMP domain as they are not part of the working age population. A datazone code was assigned to each person according to the Enumeration Area in which they were recorded, using an EA to datazone look-up table. A series of numerator and denominator counts were then created by aggregating the respective binary flags from individual level to datazone level.

---

[22] People living in institutional living quarters were excluded from the SAIMD analysis.

Finally, each domain numerator was expressed in terms of the relevant denominator to generate domain rates which represented the proportion of the relevant dominator count that was deprived on the particular domain. Each domain rate therefore ranged from a possible low of zero (i.e. none of the relevant population denominator is deprived) to a possible high of one (i.e. 100% of the relevant population denominator is deprived).

The HEA domain was created using a different methodology. Here, the mortality table from the 2001 Census was used to create an age/sex standardised measure of 'years of potential life lost'. The age/sex mortality profile in each datazone was compared against the national age/sex profile to generate the datazone level age/sex standardised scores for this domain. This domain therefore relates to individuals who were not alive on census enumeration date and so are not captured anywhere else in the census data.

Empirical Bayesian estimation - commonly known as 'shrinkage' - was applied to all five domain scores at datazone level. Shrinkage involves moving all small area scores towards another more robust score (see, for instance, Noble et al., 2004), often relating to a higher level geography. Scores for small areas that have large standard errors (i.e. those that are less 'reliable') are moved more than those that have small standard errors (i.e. those that are more 'reliable'). For the SAIMD 2001, the datazone level domain scores were 'shrunk' towards the relevant local municipality level score. The technique therefore takes into account the extent of heterogeneity within the relevant municipality.

Once all five shrunk domain scores had been constructed at datazone level, the scores were standardised by ranking from least deprived to most deprived. The domain ranks were then transformed to a specified exponential distribution to control the cancellation between the domains[23]. Finally, the five exponentially transformed domains were combined using equal weights to produce the overall South African Index of Multiple Deprivation (SAIMD) score. The SAIMD score was then ranked from least deprived to most deprived. A full discussion of the concepts, data and methods employed in the construction of the SAIMD 2001 is provided in the published project report (Noble et al., 2009a).

As noted in Chapter 2, residential segregation measures typically require categorical input data (e.g. deprived/not-deprived) and so the INC, EMP, EDU and LIV domains are suitable but the HEA domain is not. For the same reason, the overall SAIMD is also not suitable for use as the basis of residential segregation measures. However, for the purpose of introducing the broad spatial patterns of deprivation across South Africa which is the primary purpose of this chapter, the composite SAIMD is valuable as an initial starting point as it provides a helpful single measure of multiple deprivation which can then be unpicked through the analysis of the component domains. As such, the analyses I present within this chapter do draw upon the overall SAIMD initially, despite this measure not being used in subsequent chapters in terms of residential segregation.

Sections 3.4 to 3.7 below contain empirical analyses of the SAIMD and component domains to illustrate the unequal spatial distribution of deprivation across South Africa. In Section 3.4 I present national, province and municipality level deprivation rates as initial context, before focusing on the datazone level results in Sections 3.5 to 3.7.

[23] The model of multiple deprivation adopted in the SAIMD 2001 includes a methodological control to ensure that a very high deprivation score on one domain cannot be fully cancelled out by a very low deprivation score on another domain. In other words, there is an acknowledgement that having a very high level of deprivation on any domain represents a considerable challenge to an area, and this feeds through into the overall Index of Multiple Deprivation.

3.4 Results: Levels of deprivation nationally, by province and by municipality

As discussed above, four of the domains of the SAIMD 2001 (Income and Material Deprivation; Employment Deprivation; Education Deprivation; and Living Environment Deprivation) are presented as simple rates (i.e. the percentage of the population experiencing the deprivation). Table 3.1 below presents the deprivation rates for each of the nine provinces and South Africa as a whole.

As can be seen, on average 76.0% of the population in South Africa is income/materially deprived, based on the definition used in the SAIMD 2001. The percentage is highest in Limpopo at 89.5% of population deprived, but a further five provinces also exhibit deprivation rates in excess of 80% on this measure. In contrast, both the Western Cape and Gauteng exhibit deprivation rates of less than 60% on this particular measure.

In terms of employment deprivation, 45.2% of working age people are employment deprived in South Africa. Over half of the working age population in the Eastern Cape, KwaZulu-Natal and Limpopo are deprived on this measure, while slightly less than 30% of the working age population in the Western Cape are deprived.

Across South Africa, 36.4% of adults have no secondary schooling and are therefore defined as deprived on the education domain. There is a clear distinction on this measure between the provinces of Gauteng (22.9%) and Western Cape (27.2%) and the other seven provinces, where rates range from 39.9% in KwaZulu-Natal to 45.7% in Limpopo.

Table 3.1: National and provincial rates of deprivation for four of the domains in the datazone level SAIMD 2001

	Income and Material Deprivation	Employment Deprivation	Education Deprivation	Living Environment Deprivation
Western Cape	53.3%	29.5%	27.2%	41.0%
Eastern Cape	87.9%	59.2%	45.2%	81.0%
Northern Cape	74.5%	40.6%	43.9%	53.6%
Free State	82.1%	46.8%	41.6%	66.9%
KwaZulu-Natal	80.4%	52.3%	39.9%	73.8%
North West	81.2%	48.1%	42.2%	74.4%
Gauteng	59.6%	38.3%	22.9%	43.9%
Mpumalanga	82.8%	45.6%	44.2%	74.5%
Limpopo	89.5%	54.0%	45.7%	89.0%
South Africa	**76.0%**	**45.2%**	**36.4%**	**66.8%**

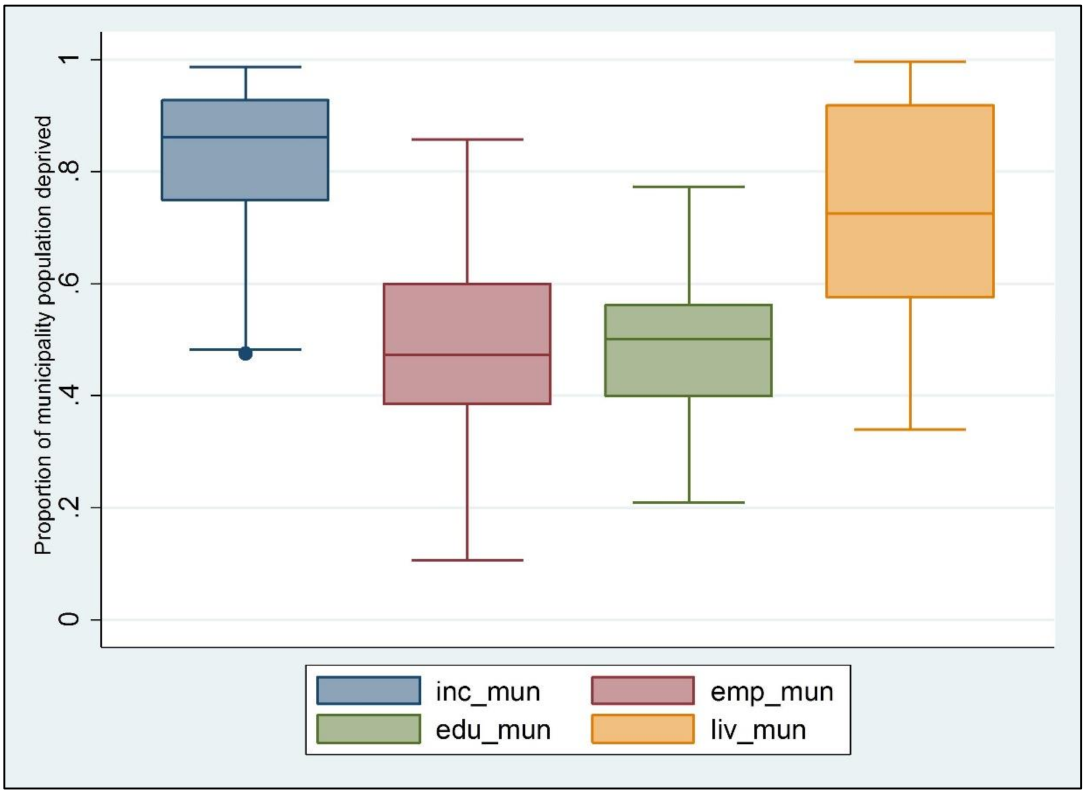
For South Africa as a whole, two-thirds of the population experience living environment deprivation (66.8%). The levels of deprivation on this measure are highest in Limpopo (89.0%) and Eastern Cape (81.0%), followed by the three provinces of Mpumalanga, North West and KwaZulu-Natal in each of which approximately three-quarters of the population are deprived. In contrast, less than half the population of the Western Cape and Gauteng are deprived on this measure.

In summary, levels of deprivation are typically lowest in Gauteng and the Western Cape, and highest in the Eastern Cape and Limpopo. With reference to the research question addressed in this chapter, these findings demonstrate an unequal distribution of deprivation between provinces on all four dimensions of deprivation presented here. However, it is important to highlight that, although the Western Cape and Gauteng are seen to be the least deprived provinces on all four domains domain, the *absolute levels* of deprivation (i.e. the proportions of population deprived) are still very high in these two provinces.

It is also possible to present deprivation rates for each local municipality (including the metropolitan municipalities) and these data are shown graphically in Figure 3.5. Along the x-axis, each box and whisker data series relates to one of the four domains of deprivation discussed above. The y-axis shows the proportion of population (of relevant denominator) that is deprived at local municipality level. The scores on each domain are calculated as simple proportions of population deprived per municipality. The horizontal lines within the boxes of the data series represent the median municipality score on each domain measure.

It is evident from Figure 3.5 that almost all municipalities in South Africa have average deprivation rates on the income and material deprivation domain of 50% deprived or more and, indeed, over half of the municipalities have deprivation rates of 80% or more on this domain measure. On the living environment domain, over half the municipalities have deprivation rates of 70% deprived or more, while on the education domain over half the municipalities have deprivation rates of 50% deprived or more. The median municipality value on the employment domain can be seen to sit just below 50% deprived. The chart shows relatively wide ranges in the overall distribution of municipality deprivation rates on each domain, but the interquartile ranges are typically around 20 percentage points, with the exception of the living environment domain where the interquartile range is approximately 35 percentage points.

Figure 3.5: Deprivation rates across the four SAIMD 2001 domains for local municipalities



Whilst these national, provincial and municipal averages provide a valuable insight into the very high levels of deprivation across South Africa, the real strength of a neighbourhood level measure such as the SAIMD and its constituent domains comes from the ability to examine variations in deprivation levels *between neighbourhoods*. In the next sub-section I therefore focus on the datazone level results and examine the degree to which deprivation rates at neighbourhood level are equally or unequally distributed across the country.

3.5 Results: Levels of deprivation at datazone level

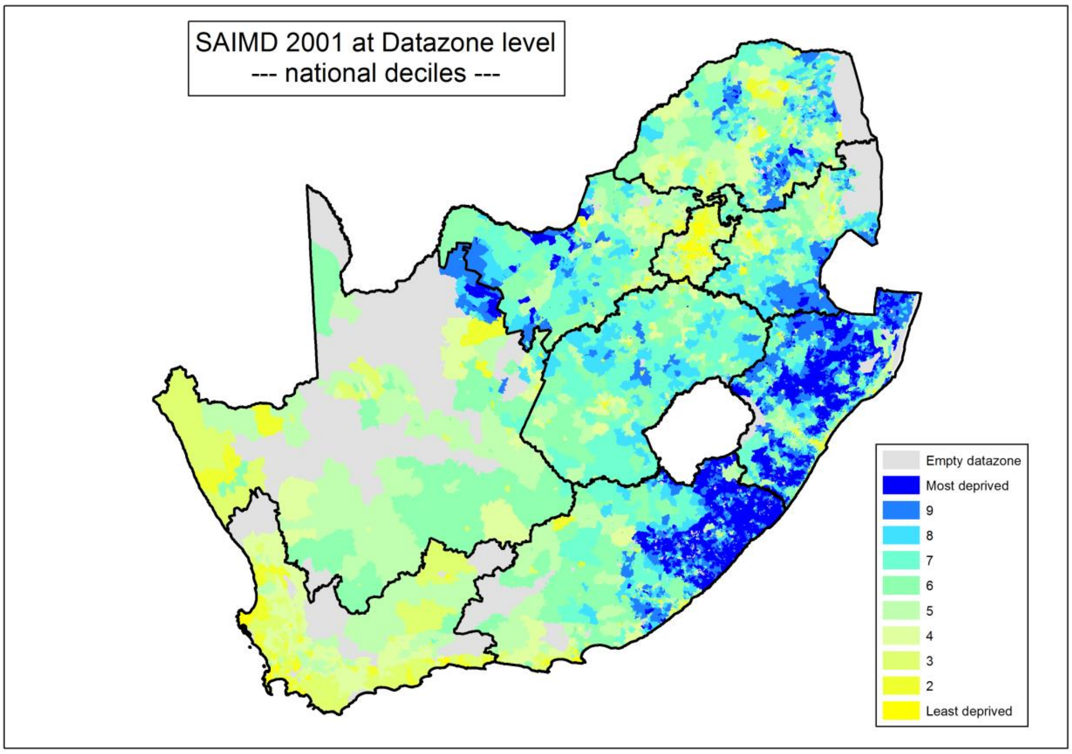
There are two parts to Section 3.5. In the first part I provide an overview of deprivation at neighbourhood level using the composite SAIMD. As I discussed above, although the SAIMD is not suitable for use as an input dataset into the calculation of residential segregation measures in Chapter 4, it is nevertheless a valuable indicator of multiple deprivation with which to set the context for the more detailed domain-specific analyses. In the second part of this section I turn to focus on the four component domains of deprivation from the SAIMD that are expressed as deprivation rates at neighbourhood level. The aim is to assess the extent to which deprivation is equally or unequally distributed across South Africa at neighbourhood level using the SAIMD and the four component domains.

SAIMD 2001: national deciles

As discussed above, every datazone in the country is assigned a rank across the country for each of the component domains and the overall SAIMD. The most deprived datazone for each measure is given a rank of 1 and the least deprived datazone is assigned a rank of 22,164. These ranks show how each datazone compares to all the other datazones in the country and, as such, are easily interpretable. In the analysis here these ranks are utilised to explore the levels of relative deprivation at datazone level across the provinces and across the metropolitan municipalities, in the context of (i.e. with reference to) all datazones in the country.

Figure 3.6 shows the national map of datazone deciles for the whole country. All 22,164 datazones in South Africa are shown on the map and, as such, the detail in densely populated urban areas is not readily apparent. However, the map does provide an immediate visual indication of the uneven spatial distribution of deprivation across the country.

Figure 3.6: National map of datazone level SAIMD 2001



The most deprived areas appear to be located in the more rural parts of the country, particularly in the Eastern Cape and KwaZulu-Natal. Indeed, the areas of greatest deprivation appear to be located within the former homelands. Separate analysis has shown that this is indeed the case (Noble and Wright, 2013b). The provinces of Western Cape and Gauteng seem to be characterised by predominantly bright yellow colour-coded datazones, indicating that these areas are amongst the least deprived in South Africa. However, in order to quantify this visual representation of the spatial pattern it is necessary to examine the data through a combination of tabular statistics and case study maps, because the detailed geographical nuances of the results can only be fully appreciated and assessed through focused geographical analysis using case study maps.

In order to unpick the cartographic information in Figure 3.6, Table 3.2 shows the provincial share of the most deprived decile (10 per cent) of datazones and the most deprived quintile (20 per cent) of datazones nationally. It is clear that almost all of the datazones that form the most deprived national decile on the SAIMD are located in either KwaZulu-Natal (44.7 per cent) or the Eastern Cape (46.8 per cent). A similar, if somewhat less extreme, pattern is also evident in terms of the most deprived quintile.

Table 3.2: Provincial share of the most deprived national decile and quintile of datazones on SAIMD 2001

Province	Share of most deprived decile of datazones	Share of the most deprived quintile of datazones
Western Cape	0.0	0.1
Eastern Cape	46.8	35.7
Northern Cape	0.0	0.4
Free State	1.2	3.2
KwaZulu-Natal	44.7	39.0
North West	3.6	6.2
Gauteng	0.1	0.9
Mpumalanga	1.4	4.2
Limpopo	2.3	10.2
Total	**100.0**	**100.0**

Table 3.3: Percentage of each province's datazones in the most deprived decile and the most deprived quintile of the SAIMD 2001 nationally

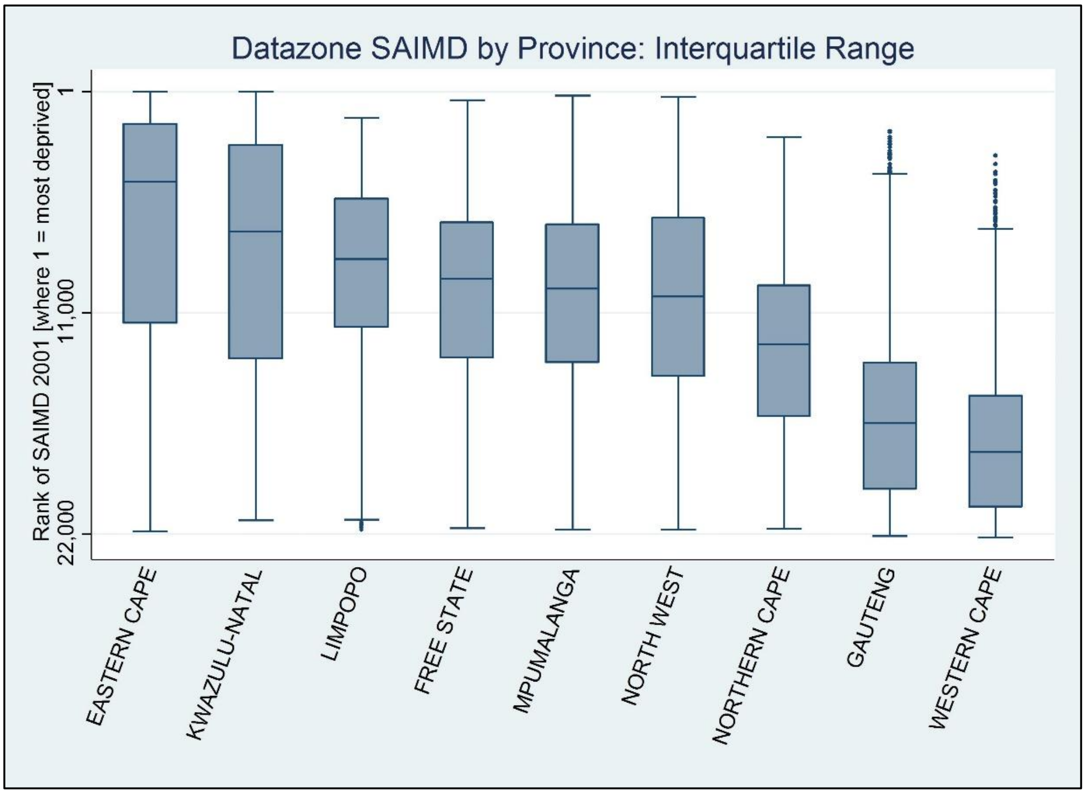
	N datazones in the province	N in most deprived decile nationally	N in most deprived quintile nationally	% in most deprived decile nationally	% in most deprived quintile nationally
Western Cape	2184	0	5	0.0	0.2
Eastern Cape	3181	1036	1583	32.6	49.8
Northern Cape	417	0	18	0.0	4.3
Free State	1373	27	143	2.0	10.4
KwaZulu-Natal	4663	991	1729	21.3	37.1
North West	1827	79	275	4.3	15.1
Gauteng	4280	2	40	0.0	0.9
Mpumalanga	1527	31	187	2.0	12.2
Limpopo	2712	50	452	1.8	16.7

Source: Noble et al. (2009a, p.24)

Table 3.3 shows the proportion of each province's datazones that are in the most deprived decile and quintile of datazones nationally. If the most deprived decile of datazones was distributed equally amongst the nine provinces, ten per cent of datazones in each province would be defined as being in the most deprived national decile. Similarly, if the most deprived quintile of datazones was distributed equally amongst the nine provinces, twenty per cent of datazones in each province would be defined as being in the most deprived national quintile. However, it is apparent from Table 3.3 that the most deprived areas are not equally distributed amongst the nine provinces. Almost a third (32.6%) of the datazones in the Eastern Cape are in the most deprived decile nationally, whilst over a fifth (21.3%) of the datazones in KwaZulu-Natal are similarly in the most deprived national decile. In other words, over three times as many datazones from the Eastern Cape and over twice as many datazones from KwaZulu-Natal are amongst the most deprived national decile than would have been observed had the most deprived decile been distributed equally (i.e. proportionally) across all nine provinces. Almost half of the datazones in the Eastern Cape are in the most deprived national quintile, which can be contrasted with the Western Cape and Gauteng where less than one per cent of the datazones fall within the most deprived national quintile. The clear message from these analyses is that the most severely deprived neighbourhoods in South Africa are not found in either the Western Cape or Gauteng, but rather in the more rural provinces of the Eastern Cape and KwaZulu-Natal.

Figure 3.7 shows the spread of deprivation in each province for the SAIMD 2001 using a 'box and whisker' plot. The vertical y-axis shows the respective datazones ranks, with rank 1 being at the top of the chart and representing the most deprived datazone in the whole of South Africa on this measure, and rank 22,164 being at the bottom of the chart and representing the least deprived datazone on this measure. The overall range of deprivation ranks within each province is illustrated by the respective vertical lines. So, KwaZulu-Natal's most multiply deprived datazone – dz_code: "531_239**"** in Ulundi municipality - is ranked 1, meaning it is the most deprived datazone in the whole of South Africa. Indeed, of the 100 most deprived datazones in South Africa, 54 are located in KwaZulu-Natal. In contrast, KwaZulu-Natal's least deprived datazone – dz_code: "572_2889" in eThekwini municipality - is ranked 21323 (where rank 22164 = least deprived). The boxes within the chart indicate the interquartile range[24]. The horizontal line within the box signifies the median datazone rank for each province, and the provinces are ordered from left to right in descending order according to this median deprivation rank.

Figure 3.7: Datazone SAIMD 2001 rank by province



For the Eastern Cape and KwaZulu-Natal the IQR boxes are relatively wide compared to the other provinces, indicating that there is a broader spread of values in Eastern Cape and KwaZulu-Natal which in turn indicates a broader range of deprivation levels across the datazones in these provinces. The IQR boxes for Eastern Cape and KwaZulu-Natal

[24] The interquartile range (IQR) is 'a measure of dispersion calculated by taking the difference between the first and third quartiles (that is, the 25th and 75th percentiles). In short, the IQR is the middle half of a distribution' (Vogt, 1999, : 143).

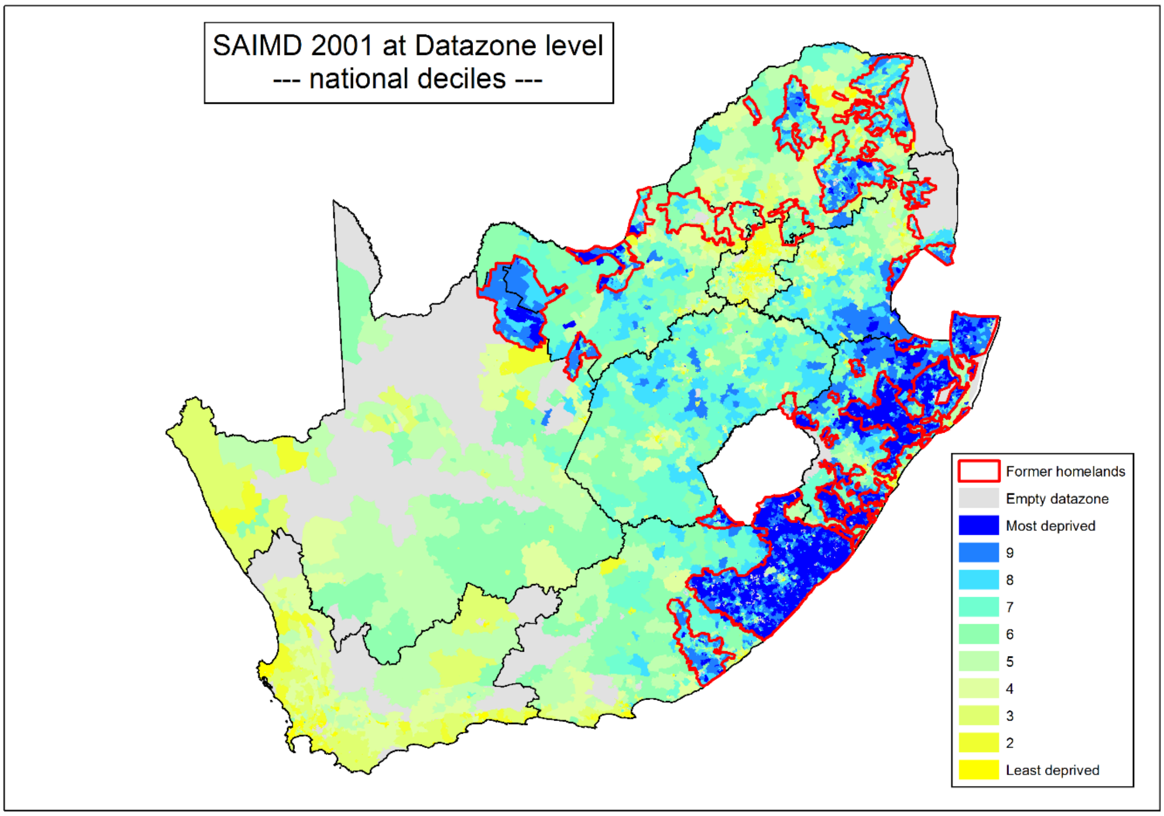
also sit towards the top of the chart, which shows that deprivation in these provinces is primarily concentrated in the more deprived part of the national distribution.

Gauteng and the Western Cape are seen to sit at the other end of the distribution, with substantially lower (i.e. less deprived) median ranks. However, both Gauteng and the Western Cape do also exhibit fairly long ‘whiskers’ which indicates that both these provinces do contain some datazones that rank relatively highly in the distribution. It is evident, for instance, that both Gauteng and Western Cape have some datazones that are placed within the ‘top’ (i.e. most deprived) 5000 datazones in the country.

The findings presented here provide an initial descriptive account of the unequal spatial distribution of deprivation across neighbourhoods in South Africa using the composite SAIMD measure. Whilst all nine provinces in South Africa contain a mix of high deprivation and low deprivation neighbourhoods, the results presented here indicate a clear spatial differentiation between the two provinces of Gauteng and Western Cape and the other seven provinces. The differentiation is particularly pronounced when comparing Gauteng and the Western Cape against the Eastern Cape and KwaZulu-Natal. A potentially important factor in this is the location of the former homeland areas which, as discussed in Chapter 1, were deliberately delineated by the apartheid government to consist of rural areas characterised by typically poor quality farm land. There are no former homelands within either Gauteng or the Western Cape, whereas the Eastern Cape contains the two large former homelands of Transkei and Ciskei, while KwaZulu-Natal contains the large former KwaZulu homeland.

Before moving on to examine the constituent domains of deprivation in detail, it is worthwhile comparing the spatial distribution of deprivation shown in Figure 3.6 with the locations of the former homeland areas. Figure 3.8 shows the same national decile distribution of the SAIMD 2001 as shown in Figure 3.6 above, but this time with the former homeland boundaries overlaid in red.

Figure 3.8: Multiple deprivation and the former homelands



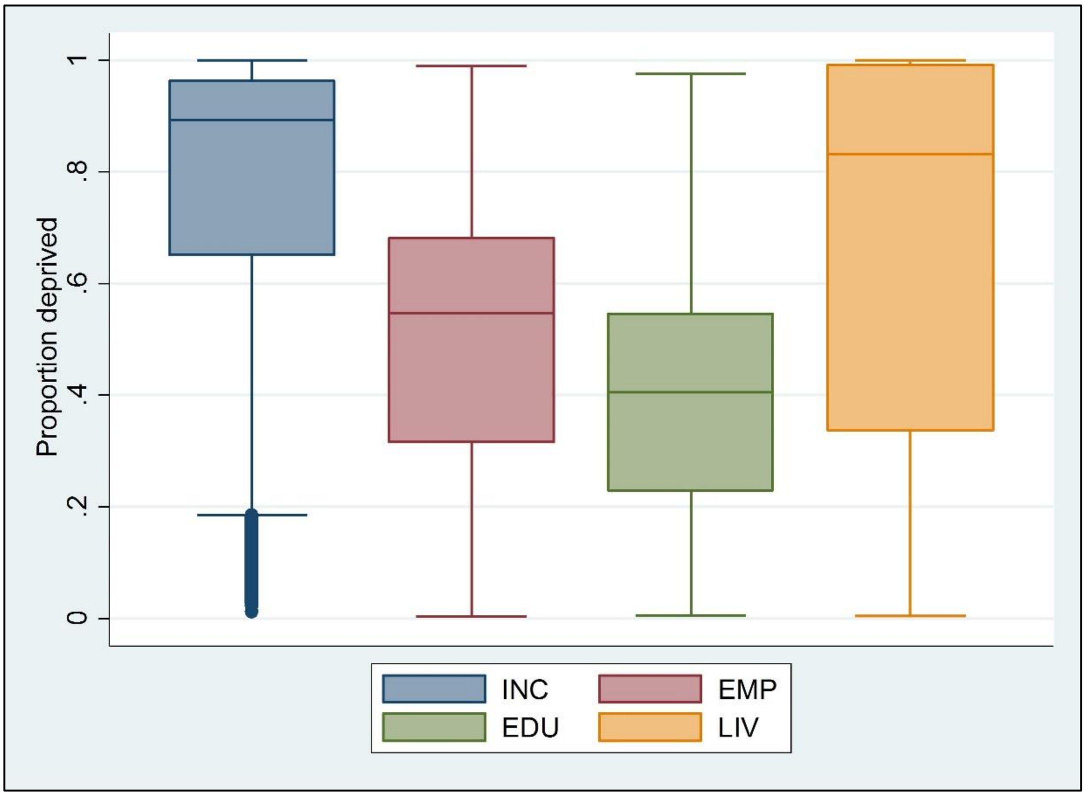
Performing this comparison gives a visual indication that many of the most deprived parts of the country are located in the former homelands. Again, it is noteworthy that Gauteng and the Western Cape, which were shown in Table 3.2, Table 3.3 and Figure 3.7 to have lower deprivation levels than the other seven provinces, are dominated by metropolitan municipalities and do not contain any former homeland areas.

In summary, therefore, whilst deprivation has been shown to be unequally distributed between and within provinces, perhaps the greatest contrast is between the metropolitan areas and the non-metropolitan former homeland areas. This juxtaposition will continue to be considered throughout the remainder of this thesis. In the next subsection I consider the unequal spatial distribution of deprivation according to the four domains of deprivation which are expressed as deprivation rates and which are therefore suitable for underpinning the construction of residential segregation measures in Chapter 4.

Domain analysis

In terms of the distribution of deprivation rates at datazone level across South Africa, Figure 3.9 demonstrates that wide disparities are evident across the country in all four domains. While the minimum and maximum datazone level scores are fairly similar across the four domains, the median values can be seen to vary considerably.

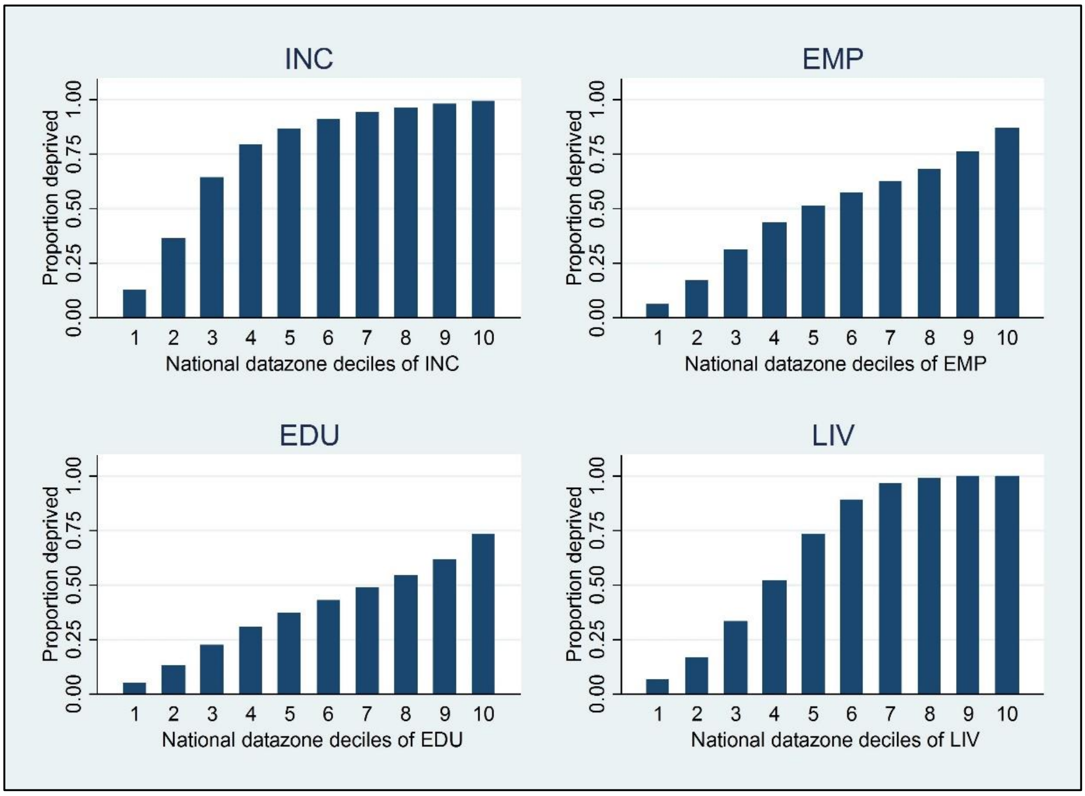
Figure 3.9: Datazone level deprivation rates for each of the four domains



The INC domain has the highest median datazone score, at 89.2%. This is followed by the LIV domain at 83.2%. The median values for both the EMP domain and the EDU domain, at 54.7% and 40.5% respectively, are notably lower than those for INC and LIV, yet still represent sizeable proportions of the denominator population.

The information in Figure 3.9 can be unpacked by examining the average rate of deprivation per decile of datazones, as shown in Figure 3.10. The decile groupings are based upon the relevant domain distribution. As such, the group of datazones that forms decile 10 on one domain is not identical to the group that forms decile 10 on each other domain[25].

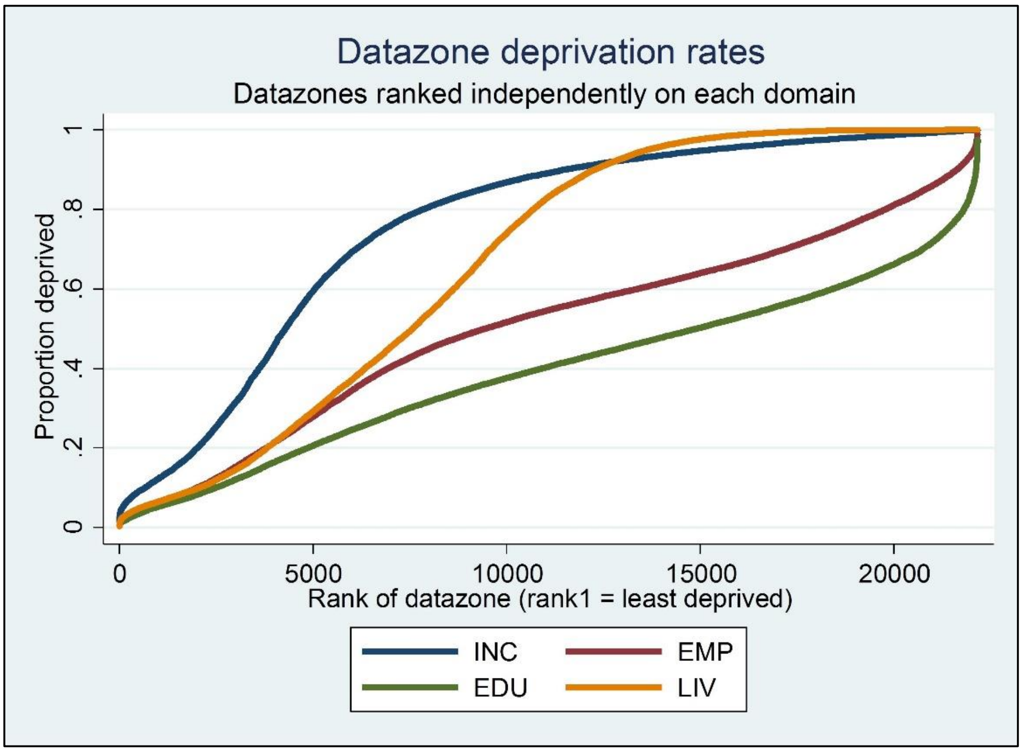
Figure 3.10: Average deprivation rates per decile of Datazones



[25] Deprivation rates are of course a reflection of the definitions of deprivation used and by ‘relaxing’ the definitions of EMP and EDU we would see distributions more similar to INC and LIV, and so it is not necessarily the case that INC and LIV are more ‘common’ then EMP and EDU; and vice versa.

It is evident from Figure 3.10 that there is a great deal of inequity between the least deprived and most deprived ends of the distribution on all four domains. For instance, the average rate of income deprivation in decile 10 of the INC domain distribution is 7.6 times higher than that of decile 1. The equivalent ratios for the EMP, EDU and LIV domains are 13.9, 14.1 and 14.8 respectively. The shapes of the four domain distributions shown in Figure 3.10 can also be seen to differ, with the EMP and EDU domains exhibiting a relatively gradual increase in deprivation rates as we move up the decile categories, whereas the INC and LIV domains exhibit a much steeper increase in rates across the lower half of the distributions, plateauing at extremely high rates of deprivation across the deciles in the upper half of the distribution. This particular feature is displayed more clearly in Figure 3.11 where the actual datazone level scores (rather than the decile averages) are plotted for each domain on a single chart.

Figure 3.11: Datazone level deprivation rates, ranked from least deprived to most deprived



It could perhaps be argued that the definitions of deprivation chosen for the INC and LIV domains do not permit sufficient discrimination between datazones at the most deprived end of the distribution. For instance, in 80.7% of datazones across South Africa, over half the datazone population is deprived on the INC domain. Furthermore, 48.2% of datazones have INC domain rates of 90% or more, 31.1% of datazones have INC domain rates of 95% or more, and 7.6% of datazones have INC domain rates of 99% or more. Similarly, in 65.8% of datazones, over half the population is deprived on the LIV domain, 44.7% of datazones have LIV domain rates of 90% or more, 38.5% of datazones have LIV rates of 95% or more, and 26.1% of datazones have LIV domain rates of 99% or more. Whilst these data provide an accurate reflection of the level of deprivation *as measured by these indicators*, they are somewhat problematic when used as the basis for ranking datazones from least deprived to most deprived or splitting datazones into decile groups. This is because there is very little absolute difference between rates of Deprivation at the most deprived end of the distributions and therefore even a very small change in deprivation rate - which could simply be due to random 'noise' in the data - could result in a datazone being placed in say decile 8 rather than decile 10. As the purpose of a conventional Index of Deprivation approach, which is the purpose for which these four domains are constructed, is to rank areas from least deprived to most deprived then the issue with the INC and LIV domains is important to recognise. However, for the purpose of constructing residential segregation measures in Chapter 4 of this thesis, the lack of discrimination between areas at the upper-end of the deprivation distribution does not cause any problems.

Table 3.4 and Table 3.5 below show, respectively, the proportion of datazones per province that rank within the most deprived and least deprived decile on each of the four domains of deprivation.

Table 3.4. Proportion of Datazones per province falling in the most deprived decile per domain

Province	Total number of Datazones in the province	Proportion of Datazones per province in the most deprived decile			
		INC	EMP	EDU	LIV
EASTERN CAPE	3,181	29.6%	27.8%	21.7%	25.2%
FREE STATE	1,373	5.1%	1.0%	9.3%	3.1%
GAUTENG	4,280	2.7%	0.1%	0.4%	2.9%
KWAZULU-NATAL	4,663	15.3%	17.4%	14.8%	17.0%
LIMPOPO	2,712	7.1%	13.1%	6.4%	10.4%
MPUMALANGA	1,527	4.3%	2.6%	10.3%	2.2%
NORTH WEST	1,827	4.3%	5.0%	14.1%	6.7%
NORTHERN CAPE	417	1.7%	1.2%	14.1%	0.5%
WESTERN CAPE	2,184	1.6%	0.3%	1.9%	0.8%

Table 3.5. Proportion of Datazones per province falling in the least deprived decile per domain

Province	Total number of Datazones in the province	Proportion of Datazones per province in the least deprived decile			
		INC	EMP	EDU	LIV
EASTERN CAPE	3,181	3.9%	3.6%	6.1%	4.9%
FREE STATE	1,373	5.8%	5.6%	6.1%	7.7%
GAUTENG	4,280	21.0%	18.4%	19.9%	19.6%
KWAZULU-NATAL	4,663	7.2%	5.0%	6.9%	7.5%
LIMPOPO	2,712	1.7%	5.3%	1.6%	2.1%
MPUMALANGA	1,527	5.4%	6.8%	5.2%	5.4%
NORTH WEST	1,827	4.9%	5.8%	6.5%	7.3%
NORTHERN CAPE	417	9.1%	15.1%	7.7%	9.4%
WESTERN CAPE	2,184	24.1%	26.9%	22.4%	20.8%

A degree of consistency across domains is evident with regard to the location of the most deprived decile of datazones. The Eastern Cape has the highest proportions on all four domains, while KwaZulu-Natal has the second-highest proportion on all four domains. In contrast, less than 2% of datazones in the Western Cape and less than 3% of datazone in Gauteng are in the most deprived decile on any domain. Turning to look at the spatial distribution of the least deprived decile of datazones, it is apparent that the Western Cape contains the highest proportions on each domain and Gauteng contains the second-highest proportions on each domain. The figures for these two provinces are far larger than those for the other seven provinces.

The main message emanating from these domain-specific analyses is that the unequal patterns of deprivation observed when looking at the composite SAIMD can largely be seen to hold true when looking at these four component domains. These findings all represent new contributions to the evidence base concerning the unequal spatial distribution of deprivation across neighbourhoods in South Africa and provide an important contextual backdrop to the analyses of residential segregation measures in Chapters 5 and 6 of this thesis.

Whilst the findings above suggest a degree of commonality across the four domains concerning the location of the most and least deprived decile of datazones, they do not reveal whether it is the *same datazones* within each province that rank highly or lowly on multiple domains. In order to tackle this question I will first examine datazone correlations across the four domains before moving on to consider scatterplots of relationships between domains and finally considering the numbers of datazones that do fall within the most and least deprived deciles on multiple domains.

Table 3.6 shows a Pearson correlation matrix for the four domain scores and Table 3.7 shows a Spearman correlation matrix for the four domain ranks.

Table 3.6: Correlations between the deprivation rates on the four domains at Datazone level

	INC	EMP	EDU	LIV
INC	1.000			
EMP	0.783	1.000		
EDU	0.806	0.579	1.000	
LIV	0.879	0.684	0.795	1.000

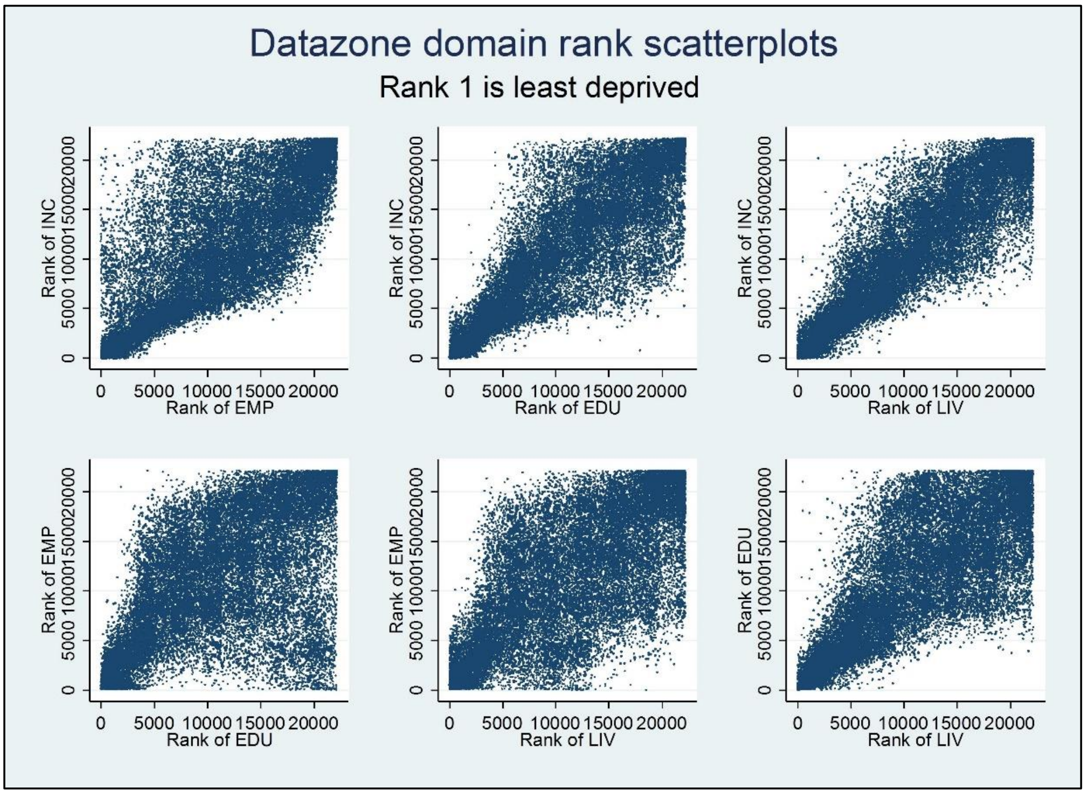
Table 3.7: Spearman rank correlation between the deprivation rates on the four domains at Datazone level

	INC	EMP	EDU	LIV
INC	1.000			
EMP	0.772	1.000		
EDU	0.827	0.564	1.000	
LIV	0.888	0.698	0.773	1.000

It is evident from Table 3.6 and Table 3.7 that relatively strong positive correlations exist between the INC, EDU and LIV domains. In terms of correlations between these three domain scores, as shown in Table 3.6, the coefficients range from a low of 0.795 to a high of 0.879. The correlations between the EMP domain and the other three domains are somewhat lower, ranging from a low of 0.579 to a high of 0.783. A similar pattern is observed in terms of the correlations between the domain ranks in Table 3.7. It may be the case that the conceptual linkage between employment and the other three domains does not hold as strongly in South Africa as initially envisaged (in other words, employment is not as dependent upon education, and employment does not necessarily result in a reduction in income deprivation). Alternatively (or in addition), it may be that there is an issue with the definition of employment deprivation used in the SAIMD (in other words, the definition is not capturing employment deprivation as fully as originally intended).

A series of scatterplots are presented in Figure 3.12 representing the relationships between each pair of domains. These scatterplots add important context to the summary correlation coefficients presented in Table 3.6 and Table 3.7. The scatterplots display the ranks of each pair of domains. Note that on these charts the datazone ranked at position 1 is the least deprived and the datazone ranked at position 22,164 is the most deprived.

Figure 3.12: Scatterplots of domain ranks



It is interesting to note that there are very few datazones in which employment deprivation is very high and income deprivation is very low. This is perhaps unsurprising given the important role that the labour market plays in determining income levels (Leibbrandt et al., 2010). In contrast, there is a sizeable number of datazones in which employment deprivation is very low but income deprivation is very high. This could be due to a situation where people living in these datazones are in very low paid employment and/or unstable employment (Neves and Du Toit, 2010) and, as such, are defined as income deprived despite not being defined as employment deprived at the time of the 2001 Census. This finding would be consistent with other research in this field which finds that having a job does not necessarily lift an individual out of income poverty (Finn, 2015).

An important question to explore further empirically is whether the datazones that constitute the most deprived decile on one domain also constitute the most deprived decile on the other domains. In other words, is it the same group of datazones that are most deprived on all domains or is there variation in composition of the most deprived decile between domains? This question can equally be asked of the least deprived decile.

Of the 22,164 datazones in South Africa analysed here, only 372 datazones fall in the most deprived decile on all four domains. The 372 datazones are spread across five of nine provinces: Eastern Cape (210 of the 372 datazones), KwaZulu-Natal (147), Limpopo (6), North West (6) and Mpumalanga (3). These figures demonstrate that there is not a sizeable group of datazones that are in the most deprived decile on all four domains. In addition to the 372 datazones that are in the most deprived decile on all four domains, a further 753 datazones fell in the most deprived decile on any three domains. Of these 753 datazones, 82 were in the most deprived decile on all except the INC domain, 163 were in the most deprived decile on all except the EMP domain, 242 were in the top decile on all except the EDU domain, and 266 were in the top decile on all except the LIV domain.

At the other extreme of the distribution of deprivation, there is a group of 1,135 datazones that are in the least deprived decile on all four domains. These 1,135 datazones are spread across all nine provinces: Gauteng (424 of the 1,135 datazones), Western Cape (336), KwaZulu-Natal (125), Eastern Cape (79), Free State (47), Mpumalanga (40), North West (43), Northern Cape (19) and Limpopo (22).

The existence of a relatively small group of datazones that are in the most deprived decile on all four domains and a relatively large group of datazones that are in the least deprived decile on all four domains is perhaps not that surprising given the analyses discussed above in relation to Figure 3.10, Table 3.6, Table 3.7 and Figure 3.12. Two key factors were identified from those analyses.

Firstly, in terms of the INC and LIV domains, there is little absolute difference (in percentage point terms) between rates of deprivation amongst those datazones at the more deprived end of the distributions. For instance, the average rate of income deprivation amongst those datazones in the eighth decile of the INC domain was over 95% of the population deprived, while the average rate of living environment deprivation amongst those datazones in the eighth decile on the LIV domain was over 99% of the population deprived. As a sizeable proportion of the total datazones in South Africa are therefore separated at the most deprived end of the distribution by only a very small range in deprivation rates, it is difficult to meaningfully discriminate between these areas. As such, focusing on just the most deprived decile becomes less appropriate than in the cases of the EMP and EDU domains where it was seen in Figure 3.10 that there is a more graduated range of deprivation rates by decile and where there are indeed notable 'tails' at the most deprived end of the distributions where more extreme levels of deprivation are observed.

Secondly, and linked to the above point, it is evident that there is a tighter clustering at the least deprived end of the distributions on all the domain-by-domain scatterplots indicating that being at the least deprived end of the distribution on any one particular domain increases the likelihood of being at the least deprived end of the distribution on the other domains.

In summary, the results from the domain-specific analyses have demonstrated an unequal spatial distribution of deprivation across neighbourhoods on all four component domains considered here, with a degree of commonality observed between all four domains but with the EMP domain diverging most from the other three domains. The most deprived areas tend to be concentrated in the more rural provinces whilst the least deprived areas tend to be found in the more urbanised provinces. Those datazones that rank within the very least deprived nationally on any one domain tend also to rank amongst the least deprived on the other three domains. This cross-domain correspondence is less evident at the most deprived end of the distribution. This is perhaps what we would expect given the skewed distributions of the underpinning domain scores (i.e. deprivation rates) on the INC and LIV domains, with large numbers of datazones recording extremely high deprivation rates thereby resulting in very little discrimination between areas at the most deprived ends of the distributions. Despite this issue with national rankings (and by extension the deciles), there remains clear evidence that many neighbourhoods across South Africa suffer from very high rates of deprivation across multiple domains and many of these areas are located in provinces that used to contain large former homelands. In contrast, the two provinces that are dominated by major metropolitan municipalities (Western Cape and Gauteng) tend to contain greater concentrations of the less deprived neighbourhoods. In order to examine this potential contrast between urban and rural areas in more detail, and to pay particular attention to the case of the former homelands, I will now proceed to consider deprivation levels by area type.

3.6 Results: Deprivation levels by area typology

In this section I compare levels of deprivation between urban datazones and rural datazones utilising the urban-rural classification from the 2001 Census. Every SAL in South Africa was categorised as either 'urban', 'rural' or 'mixed' by StatsSA. For this thesis it was therefore necessary to utilise the SAL categorisation to derive a datazone categorisation. This was achieved by coding all datazones that consisted entirely of urban SALs as 'urban' and all datazones that consisted entirely of rural SALs as 'rural'. Those datazones that consisted entirely of mixed SALs, and those datazones that consisted of a combination of urban, rural or mixed SALs were classified as 'mixed'. Of the 22,145 datazones across South Africa, 11,755 are urban, 9,671 are rural, and 1,175 are mixed.

Figure 3.13, Figure 3.14, Figure 3.15 and Figure 3.16 show the percentage of datazones per urban and rural category that fall within each of the deciles on each of the four domains. The data for the mixed category are not shown. So, for instance, Figure 3.13 shows that 18% of the urban datazones across South Africa are classified as being within the least deprived national decile of datazones in terms of INC deprivation rates. It is immediately apparent that the distributions for the INC, EDU and LIV domains are very similar and consist of the rural datazones being concentrated at the most deprived end of the decile distribution and the urban datazones being concentrated at the least deprived end. On these three domains the distribution of urban datazones across the deciles reflects a roughly inverse distribution to that of the rural datazones. The distributions of urban and rural datazones across the EMP deciles are again somewhat different to the other three domains. Here we see again that the most deprived areas are overwhelmingly located in the rural areas, but the graduated distribution observed for urban datazones on the other three domains is not apparent here on the EMP domain. Finally, it is evident that between 1% and 3% of urban datazones are in the most deprived deciles on each of the four domains, whereas only 0.01% of the rural datazones are in the least deprived decile on the INC, EDU and LIV domains.

Figure 3.13: Urban and rural Datazones by INC deciles

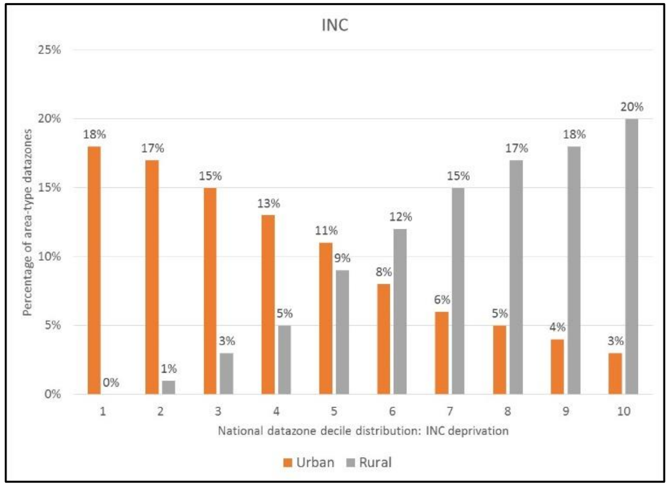


Figure 3.14: Urban and rural Datazones by EMP deciles

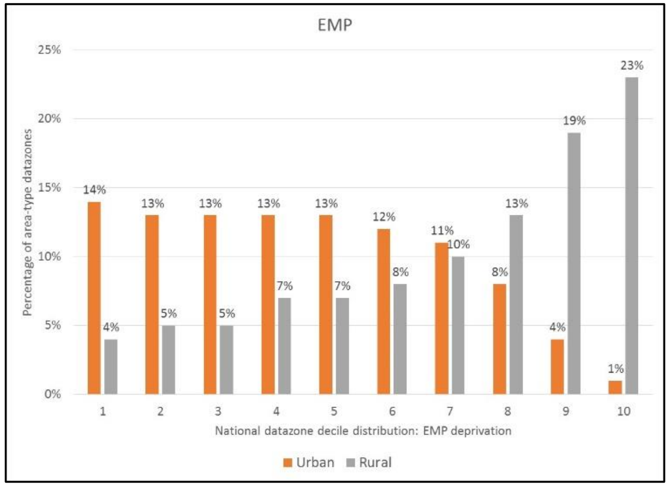


Figure 3.15: Urban and rural Datazones by EDU deciles

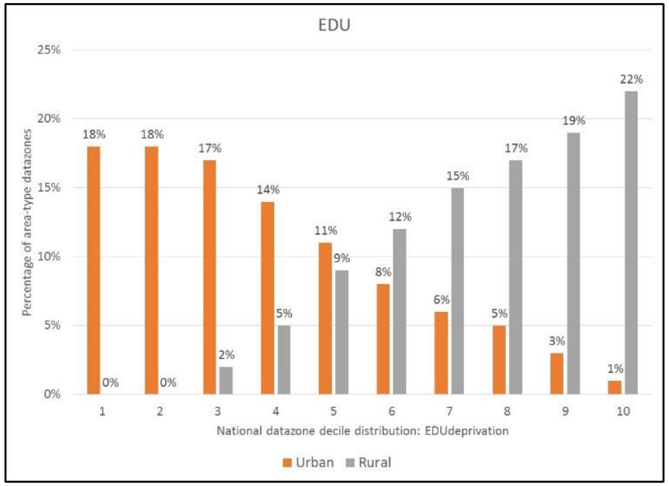
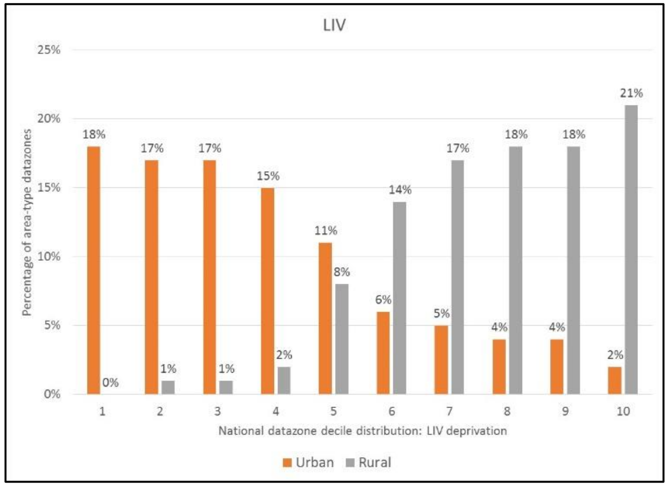
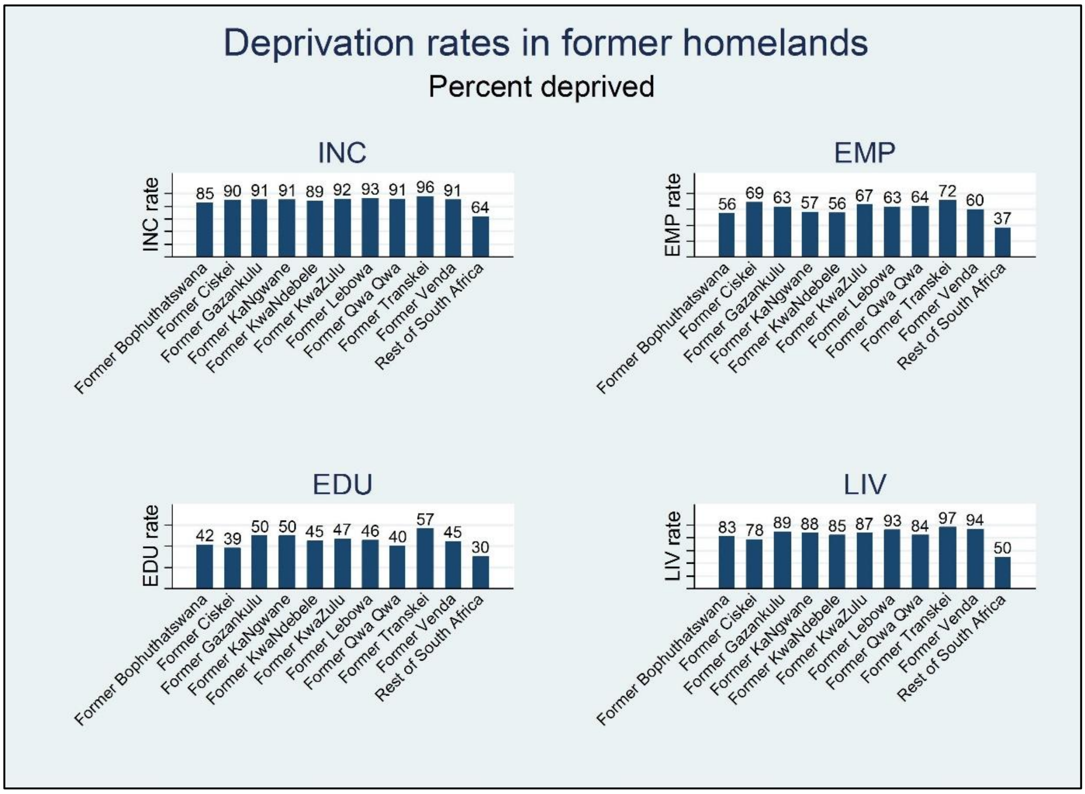


Figure 3.16: Urban and rural Datazones by LIV deciles



In Figure 3.8 above it was evident that the former homeland areas contained many of the most severely deprived areas of South Africa according to the composite multiple deprivation measure from the SAIMD 2001. Further analysis here shows that the ten former homeland areas are also characterised by extremely high levels of deprivation across all four domains. Figure 3.17 shows the rates of deprivation per former homeland for each domain,

Figure 3.17: Rates of deprivation in the former homeland areas



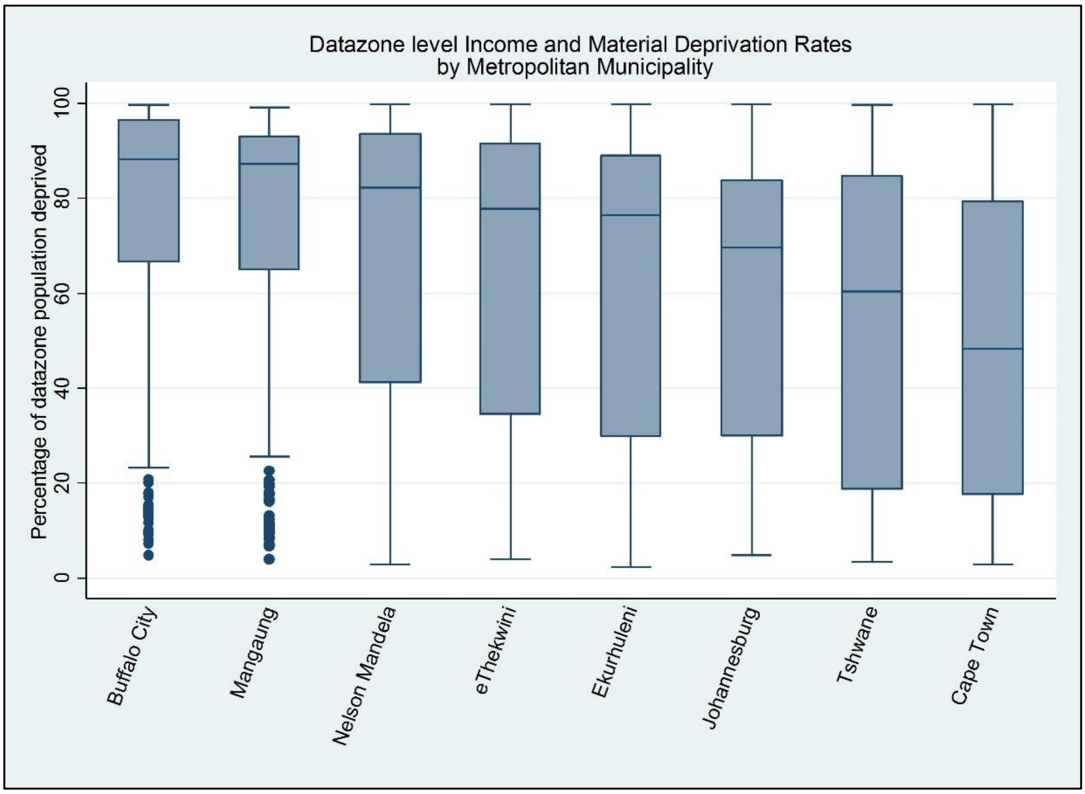
It can be observed that the level of deprivation on the INC domain ranges from 85% of the population in the former Bophuthatswana to 96% of the population in the former Transkei. On the EMP domain the values range from 56% in the former Bophuthatswana and former KwaNdebele, up to 72% in the former Transkei. In terms of education deprivation, values range from 39% of the relevant population in the former Ciskei to 57% of the relevant population in the former Transkei. Finally, on the LIV domain, the values range from 78% of the population in the former Ciskei to 97% in the former Transkei.

The results presented and discussed so far in this chapter demonstrate an unequal distribution of deprivation between provinces, municipalities and datazones on the different deprivation measures utilised here. A key advantage of the datazones is that they permit analysis of unequal distributions of deprivation *within* as well as between these larger geographical areas. In the final empirical section of this chapter I examine the spread of deprivation values within the metropolitan municipalities and use Cape Town as a particular case study in order to illustrate the considerable variations in levels of deprivation that can be observed within a single municipality.

3.7 Results: Spatial mix

Figure 3.19 shows the datazone level distribution of income and material deprivation *rates* within metropolitan municipalities in South Africa. For the purpose of this chart, the definition of 'metropolitan municipality' includes the six municipalities classified as metropolitan at the time of the 2001 Census, plus Buffalo City and Mangaung which were subsequently afforded metropolitan status. The eight municipalities are ordered from left to right on the chart according to the median datazone level value in the municipality.

Figure 3.18: Datazone level rates of Income and Material Deprivation by Metropolitan Municipality



It is clearly apparent from Figure 3.19 that the vast majority of datazones across these eight metropolitan municipalities have very high rates of deprivation on this domain. Indeed, from analyses of the data underpinning this chart it is evident that approximately two-fifths of the group of metropolitan datazones have deprivation rates exceeding 80% deprived on this measure. The municipalities of Buffalo City and Mangaung have relatively narrow IQRs positioned towards the more deprived end of the spectrum, and median datazone deprivation rates of over 80% deprived. In contrast, Cape Town's IQR is much wider and more evenly spans the more- and less-deprived parts of the distribution. Cape Town is adopted as a case study here in light of the broad spread of values on this deprivation measure.

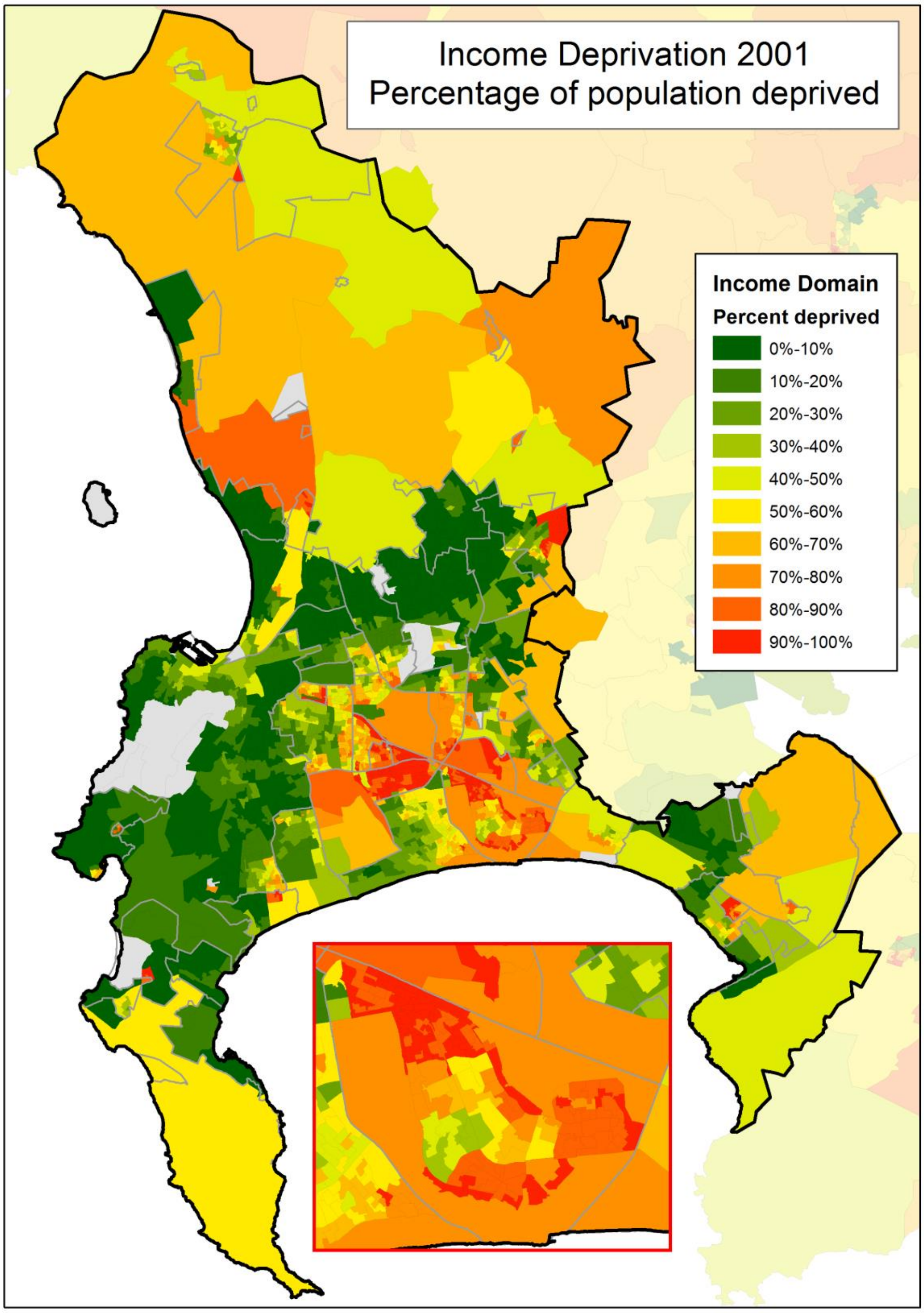
Cape Town is a large metropolitan municipality in the Western Cape. Its population at the time of the 2001 Census stood at approximately 2.8 million, with these people spread across 1,388 datazones. According to the SAIMD 2001, 48.5% of the population of Cape Town was deprived on the income and material deprivation domain, resulting in Cape Town being ranked as the third *least* deprived municipality nationally on this measure (behind Stellenbosch at 47.5% deprived, and Saldanha Bay at 48.2% deprived). The median datazone deprivation rate in Cape Town on this domain is 48.3%, meaning that in approximately one half of Cape Town's datazones less than half of the datazone population is deprived, and in the other half of Cape Town's datazones' more than half of the population is deprived. As is the case in many municipalities, the average deprivation rate across the entire municipality masks considerable variation in deprivation rates amongst the constituent datazones within the municipality. The highest absolute rate of deprivation on this domain within Cape Town is seen in datazones "171_4309" and "171_4301" (in Bloekombos, Kraaifontein), where every single person enumerated in the 2001 Census was defined as deprived on this measure. In contrast, the lowest absolute rates of deprivation on this domain within Cape Town are in datazones "171_4068" (in Arauna, Brackenfell) and "171_3399" (in Edgemead, Milnerton), where less than 3% of the total datazone population was defined as deprived on this measure.

To illustrate this intra-municipality variation cartographically, Figure 3.20 shows the deprivation rates on the income and material deprivation domains of the SAIMD 2001 for all datazones within Cape Town using ten colour-coded intervals of fixed range. For instance, those datazones colour-coded dark green are areas where less than 10% of the resident population are deprived on this measure, while those datazones colour-coded bright red are areas where over 90% of the resident population are deprived on this measure. As such, the colour-coding relates to absolute deprivation rate values, in contrast to the national decile maps shown earlier in this chapter.

We can see low levels of income and material deprivation across much of the southern suburbs and northern suburbs, with many areas exhibiting deprivation rates of between 0% and 10% (i.e. the dark green areas). This can be contrasted against the areas along the N2 highway, such as Langa, Nyanga, Philippi, Gugulethu and Khayelitsha, where there are some datazones in which between 90% and 100% of the population are deprived on this measure. Some of the outer-lying townships, such as Witsand, Masiphumelele and Doornbach are also seen to be colour-coded bright red, as is Bloekombos where we know that every single person enumerated in the 2001 Census was deprived on this measure.

Within Khayelitsha, the northern, southern and eastern sections are typically characterised by deprivation rates exceeding 80% deprived, whilst the central section of Khayelitsha is characterised by somewhat lower rates, with some datazones registering less than 50% of the population deprived on this measure.

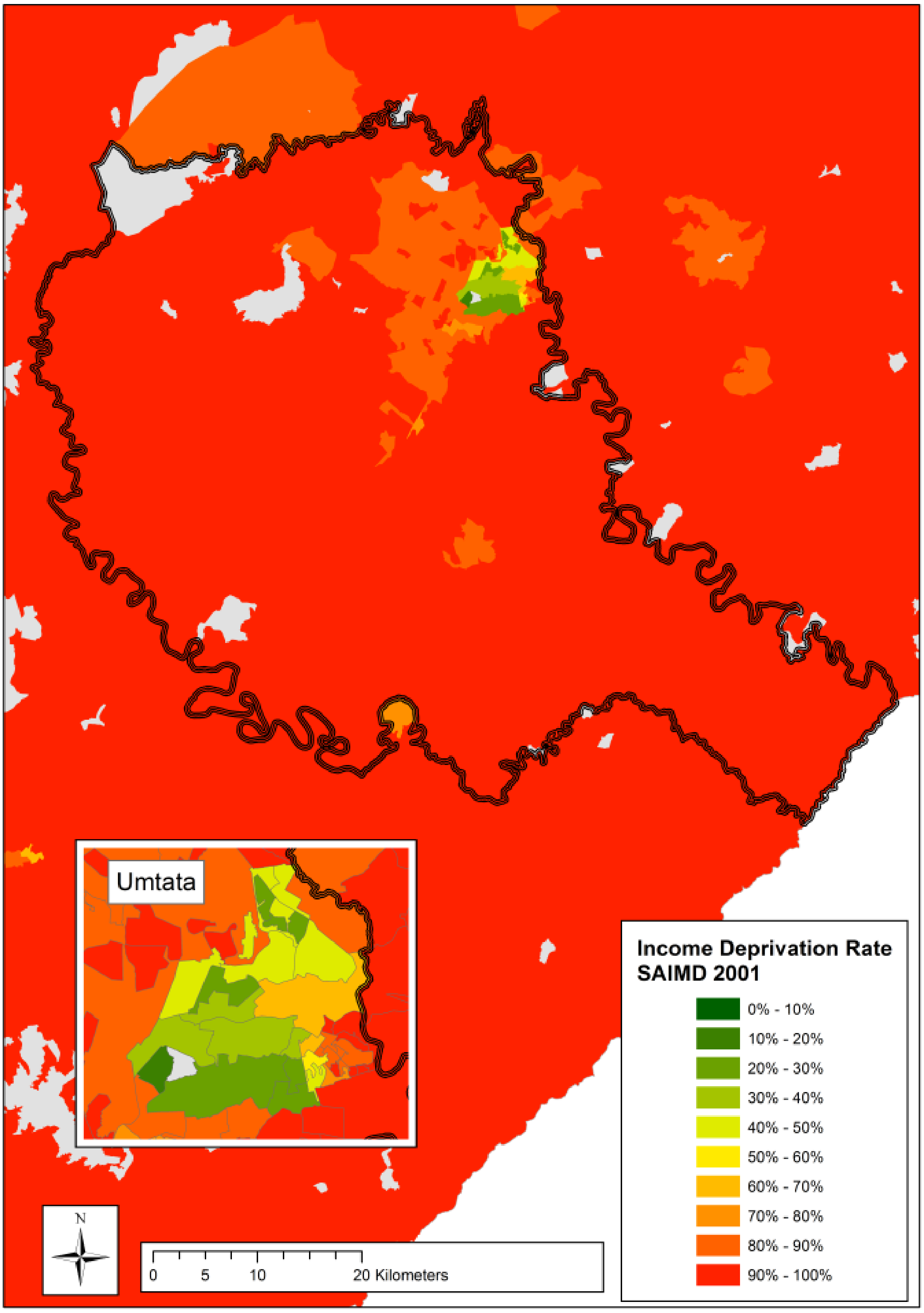
Figure 3.19: Datazone level map of Income Deprivation Domain rates for Cape Town



These cartographic analyses provide an additional indication of the degree to which deprivation rates are unevenly spread within municipalities as well as between municipalities. Whilst there are geographical concentrations of relatively low deprivation datazones (e.g. in the southern suburbs) and geographical concentrations of relatively high deprivation datazones (e.g. in the Cape Flats, including much of Khayelitsha), there are also notable areas where high deprivation and low deprivation datazones lie in close proximity. These analyses therefore provide a strong visual indication that certain parts of South Africa are characterised by considerable spatial inequality in levels of deprivation.

The picture for Cape Town can be contrasted with that of a mainly rural municipality in a former homeland area. As an example of this, Figure 3.21 shows the deprivation rates (using the same colour-coding scheme as Figure 3.20) for datazones within and around the King Sabata Dalindyebo municipality in the Eastern Cape (which also lies within the former Transkei homeland area). Whereas Cape Town was characterised by a considerable spatial mix in deprivation rates, King Sabata Dalindyebo is overwhelmingly characterised by extremely high rates of deprivation in all parts of the municipality other than a small pocket of less deprived datazones within the town of Umtata.

Figure 3.20: Datazone level map of Income Deprivation Domain rates for King Sabata Dalindyebo municipality in the former Transkei homeland, Eastern Cape Province



As discussed in Chapters 1 and 2, and as will be developed further in Chapter 4, the spatial configuration of deprivation at neighbourhood level represents the socio-spatial context in which people live their lives and which, I argue, forms the basis for shaping people's lived experience of inequality.

3.8 Conclusion

The aim of this chapter has been to address sub-question Q1: To what extent is deprivation distributed unequally across neighbourhoods in South Africa? In addressing this research question I sought to achieve two main objectives. First, I sought to make a new contribution to the empirical evidence base concerning spatial patterns of deprivation at neighbourhood level. Second, I sought to use these analyses as a means to introduce the deprivation dataset and base geography which I intend to use to construct and review residential segregation measures in Chapter 4.

Through the analysis of the constituent domains and overall composite index of multiple deprivation from the SAIMD 2001 at Datazone level, I have demonstrated that deprivation is unequally distributed across South Africa when measured at the level of province, municipality and neighbourhood. In terms of unequal distributions of deprivation between provinces, the analyses revealed that Gauteng and the Western Cape exhibited notably lower levels of deprivation than the other seven provinces, with deprivation typically highest in the Eastern Cape and KwaZulu-Natal. However, the levels of deprivation in Gauteng and Western Cape cannot be described as 'low' in an absolute sense, especially in the context of South Africa's classification as an 'upper middle income country' by the World Bank.

The analyses of deprivation rates at datazone level across the four constituent domains of deprivation revealed fairly strong positive correlations between the domains of income and material deprivation, education deprivation and living environment deprivation. The correlations were also positive, although somewhat weaker, with the employment deprivation domain. These findings demonstrate the importance of the choice of deprivation measure, as the results on the employment deprivation domain are less representative of the broad spatial distributions of deprivation seen across the other the domains. This finding informs the analytical approach later in this thesis as I continue to utilise all four domains of deprivation in my analyses of residential segregation.

The analysis of the spatial distribution of deprivation using the datazone level deprivation scores revealed a clear urban/rural pattern, with the most severely deprived datazones in the country being largely in rural areas, with particular concentrations within the former homeland areas. Furthermore, there were very few instances of the least deprived datazones in the country being located in rural areas. In contrast, urban areas tend to contain a more heterogeneous mix of more and less deprived datazones. This was demonstrated graphically with a focus on eight major urban areas (consisting of the six metropolitan areas from 2001 plus the additional two areas classified as metropolitan by 2011). Using Cape Town as a case study, I showed how deprivation can be distributed unequally between neighbourhoods *within* a municipality. Although there was evidence of spatial concentrations of less deprived datazones in some parts of the municipality and concentrations of more deprived datazones in other parts of the city, there were also areas of interface where high deprivation and low deprivation datazones were located in very close proximity.

In light of the contention throughout this thesis that people's local socio-spatial environment helps to shape their lived experiences of inequality, the findings in this chapter make an important contribution to the overall thesis research question. However, in order to operationalise these deprivation data for the purpose of measuring people's lived experiences of inequality it is necessary to go beyond the analyses here and develop new empirical measures of inequality that reflect the unequal spatial distribution. This is the focus of Chapter 4.

Chapter 4: Building a new spatial measure of exposure to inequality

4.1 Aims and objectives

Through the analyses in Chapter 3 I demonstrated a high degree of spatial inequality in levels of deprivation across neighbourhoods in South Africa. I was able to show that the level of inequality between proximate neighbourhoods varies according to where within South Africa a neighbourhood is located. For instance, I showed that there is far greater heterogeneity in deprivation levels between neighbourhoods within the City of Cape Town than between neighbourhoods in the King Sabata Dalindyebo municipality. Flowing on from this, I concluded that a person living in Cape Town may have a very different 'lived experience' of inequality than a similar person living in King Sabata Dalindyebo, for instance. The aim in this chapter is to try to quantify the extent to which a person's lived experience of inequality varies according to where they live and therefore I seek to specifically tackle Sub Q2: "Can an empirical measure be developed that reflects people's lived experience of inequality?"

As discussed in Chapter 2, measures of spatial inequality at neighbourhood level typically take the form of residential segregation indices and Massey and Denton's seminal work on identifying, categorising and critiquing various measures of residential segregation provides a valuable starting point for my research here (Massey and Denton, 1988).

In this Chapter I review each of Massey and Denton's five dimensions of residential segregation (evenness, concentration, centralisation, exposure and clustering) and consider their relevance to South Africa for the purposes of developing a small area level measure of lived experience of inequality. The suitability (or fit-for-purpose) assessment of the measures is inevitably based on researcher judgement taking into account the principal characteristics of the phenomenon I am endeavouring to measure – that is the degree to which people at one end of the socio-economic spectrum come into contact with people or visit places that are at the other end of the socio-economic spectrum.

In order to be 'fit for purpose' I argue that a small area level measure of spatial inequality must satisfy three criteria:

1. it must permit comparisons of inequality between all neighbourhoods in the country on a consistent basis;
2. it must reflect the experience of inequality from an individual's perspective;
3. data to calculate the measure must be available at the sub-municipality level.

I begin by considering the dimension of evenness, before progressing to consider the dimensions of concentration, centralisation, exposure and clustering. Where data availability permitted I constructed the particular statistical measure that was recommended by Massey and Denton as the best measure of that dimension in order to empirically test the suitability of each dimension for my purpose. For clarity within this thesis, I present some of the equations and empirical workings within Appendix B. I conclude by identifying the measure which I argue best suits the purpose of this thesis, but caveat this with a number of limitations concerning that chosen measure. I then present a detailed account of how I take the chosen measure and modify it in order to address some of the limitations in the conventional measure, resulting in the development of my final measure of the lived experience of inequality.

4.2 Massey & Denton's five dimensions of residential segregation

As discussed in Chapter 2, Lloyd and Shuttleworth (2012) raise the distinction between 'global', 'spatial' and 'local' measures of residential segregation. 'Global' measures are those in which neighbourhood level segregation scores are computed for each neighbourhood independently, with no consideration of spatial inter-connectedness between proximate neighbourhoods. These neighbourhood level segregation scores are then averaged over a larger geographical area (e.g. a municipality) to produce a single residential segregation statistic for that larger geographical area. 'Spatial' measures are those in which neighbourhood level segregation scores are computed for each neighbourhood by taking into account the spatial inter-connectedness of proximate neighbourhoods, and then averaged over a larger geographical area (e.g. municipality) to produce a single residential segregation statistic for that larger geographical area. 'Local' measures are those in which neighbourhood level segregation scores are computed for each neighbourhood by taking into account the spatial inter-connectedness of proximate neighbourhoods, and then averaged over a *bespoke* larger geographical area surrounding the target neighbourhood, so that each neighbourhood in the country is given its own residential segregation statistic based upon its own bespoke larger surrounding geographical area. As is discussed in more detail below, both 'spatial' and 'local' measures entail the use of a specified distance decay function to generate the empirical weights that reflect the inter-connectedness of respective neighbourhoods. The weights are used in the calculation of the final residential segregation statistic for the given target area (which might be a local municipality for the 'spatial' measure and the individual neighbourhood for the 'local' measure).

The statistical measures Massey and Denton presented under their dimensions of evenness, exposure, concentration and centralisation were constructed as 'global' measures, while the measures they presented under their dimension of clustering were constructed as 'spatial' measures. For the purpose of this thesis I initially follow Massey and Denton's approach in this regard, whilst acknowledging that other research in this field has sought to develop spatial and/or local variants of the traditional global measures.

For the purpose of testing each segregation measure I used the Income and Material Deprivation Domain from the SAIMD 2001 at datazone level as the input dataset. As discussed in Chapters 2 and 3, it is possible to derive from this measure the number of people who are defined as deprived and the number of people who are defined as not deprived according to the particular definition of Income and Material Deprivation. As shorthand and for clarity, I refer to the deprived population as 'poor' and the not deprived population as 'non-poor'.

Evenness

According to Massey and Denton, the dimension of evenness refers to the differential proportional distribution of two social groups among neighbourhoods within the specified broader geographical extent, such as within a city or other local government administrative area. A group is said to be segregated if it is unevenly distributed over the constituent neighbourhoods relative to another group (Blau, 1977; Massey and Denton, 1988). With regards to the poor/non-poor categorisation, the condition of maximum evenness occurs when all neighbourhoods have the same *relative* number of poor and non-poor members as the broader geographical area as a whole. On the other hand, the condition of maximum segregation occurs when no poor people share a neighbourhood with non-poor people, and vice versa. Massey and Denton recommend using the dissimilarity index when examining the dimension of evenness.

The dissimilarity index is perhaps the most commonly used empirical measure in the residential segregation literature (Christopher, 2001; Dorling and Rees, 2003; Duncan and Duncan, 1955; Duncan and Lieberson, 1959; James and Taueber, 1985; Lichter et al., 2012; Logan and Stults, 2011; Massey and Denton, 1988; Massey and Denton, 1989; Massey and Fischer, 2003; Massey et al., 2009) and Massey and Denton recommend this particular measure as their preferred index of evenness.

Although the dissimilarity index is widely used in social segregation research, its value for my purpose is somewhat limited. The dissimilarity index is composition invariant, meaning that results do not depend on the relative size of the groups being compared (Lieberson, 1981). For instance, two hypothetical municipalities, one with a poverty rate of 40% and the other with a poverty rate of 4%, could quite feasibly have the same score on the dissimilarity index. As Lieberson and Carter (1982) note, in many applications of segregation indices in the literature this composition invariance is regarded as a strength as it enables comparisons of evenness between different cities (even if the proportional composition varies between cities) or the same city at different points in time (even if the proportional composition varies over time). However, for my purpose, this feature of the dissimilarity index is a weakness because, I would argue, people’s perception of their relative socio-economic status – and hence potentially their lived experience of inequality – may well be related to the numbers (and proportions) of similar people in their neighbourhood and wider social area.

I therefore reject the dissimilarity index as it fails to meet the criteria listed above concerning the ability to compare areas across the country on a consistent basis and the need to reflect experience of inequality from an individual's perspective. Other measures of evenness considered by Massey and Denton include spatially referenced versions of the 'classical' measures of (income) inequality such as the GINI coefficient, General Entropy measures, and the Atkinson index (Massey and Denton, 1988). These measures fail to meet the same two criteria as the dissimilarity index and are therefore also not considered further here.[26]

Concentration

Concentration indices measure the extent to which a certain population sub-group occupies a small or large proportion of the actual physical space within a geographical area. A group can be said to be residentially concentrated if it occupies a small proportion of the overall physical space (Hoover, 1941; Massey and Denton, 1988). Massey and Denton construct a series of concentration indices, some of which measure concentration in an absolute sense (i.e. comparing the total land area inhabited by a certain group with the minimum and maximum possible land areas that could be inhabited by that group) and some which measure concentration in a relative sense (i.e. the extent to which a group is disproportionately concentrated into small geographical areas compared to another group), and conclude that the Relative Concentration Index (RCO) should be adopted as the optimal measure of residential concentration (Massey and Denton, 1988). The RCO was used in a number of studies concerning the

[26] In addition to spatially referenced measures of evenness, it is of course also possible in theory to construct a GINI for each neighbourhood, using individual-level data, however such data is only captured by the Census and is not publicly available.

concentration of certain racial groups in major cities of the United States and was one of the measures that Massey and colleagues used to identify patterns of 'hypersegregation' of the black population within those urban areas (Denton, 1994; Massey and Denton, 1989; Massey and Denton, 1993).

The RCO has been criticised on methodological grounds. For example, Egan et al (1988) demonstrated a number of flaws in the index which they conclude invalidates its use as a reliable indicator of residential concentration. They highlight that the RCO is based upon the average geographic size of neighbourhoods without actually considering population density (Egan et al., 1988). They also revealed a problem with the supposed bounding between -1.0 and 1.0, highlighting scenarios calculated through their own work where RCO scores exceeded these bounds, often by quite considerable amounts. In expressing this particular criticism, Egan et al make reference to the concentration of racial 'minority' groups (group *X*) compared to racial 'majority' groups (group *Y*) in exactly the same way that Massey and Denton present and describe the RCO in their 1988 paper:

> "Inspection of the equation for RCO shows that this index will remain bounded so long as the minority [X] group is not more sparsely concentrated than the majority [Y] group when it is as sparse as possible, and as long as the majority [Y] group is not more densely concentrated than the minority [X] group when it is as densely concentrated as possible. It may [be?] that Massey and Denton assumed that the minority [X] group would always be more densely concentrated than the majority group and overly relied on this inequality." (Egan et al., 1988, p.1120).

A further, more minor, criticism from Egan et al is that in Massey and Denton's equation, the cumulative populations associated with *n1* and *n2* may not (and indeed are unlikely to) equate to discrete numbers of entire areal units, which can introduce a margin of error into the calculations. Egan et al's recommendation, based upon their exploration of the RCO is that the RCO should not be used as a measure of residential concentration.

In response to Egan et al's criticisms, Massey and Denton (1998) acknowledge the issue with bounding, particularly that theoretically the RCO has no lower bound (although they remained confident that the upper bound is a value of 1.0). They state that:

> "whenever the number of X members is very small and the areas in which they live are very large, a range exception occurs and the index goes below -1, sometimes by a considerable margin" (Massey and Denton, 1998, p.1126-1127).

They give a number of reasons for not identifying this issue in their earlier work, such as focusing previously on typically quite large minority *X* groups and focusing on urban areas which led them to impose certain a priori restrictions on the areal units that they included in the calculation, including dropping very large and very sparsely populated areas. They offer two possible options to deal with the issue of boundedness: (i) simply bottom-code the RCO to -1.0 whenever the index falls below this value; or (ii) impose a group size threshold before undertaking the computation of RCO to avoid the situation occurring where the size of a group falls to a sufficiently low level to take the index value below -1.0. They accept that neither solution is ideal but they argue that either is better than completely discarding the RCO altogether. Massey and Denton also acknowledge Egan et al's third criticism that the total population of group *X* or *Y* rarely, if ever, coincide with *n1* or *n2*, the critical cut points in the cumulative distribution of areal units. Massey and Denton adopted the simplistic approach of including the entire population of a spatial unit that intersected the *n1* or *n2* cut points, although they accept that a more sophisticated approach might be to estimate by interpolation what share of that unit's area is inhabited by group *X* or group *Y*.

Egan et al's criticism that a concentration index should use population density rather than areal size of geographical units is more fundamental as it affects index values in all places and all scenarios. In response to this criticism, Massey and Denton point out that in most analyses that use Census geographies as the unit of analysis (such as US Census tracts used by Massey and Denton and Egan et al), there is a clear negative correlation between areal size and population density because in the US Census geographies are often designed to contain roughly similar *numbers* of people. In other words, Census units that are small in terms of areal size are typically fairly densely populated, and units that are large in areal size are typically fairly sparsely populated. In light of this, RCO values computed based upon areal size should be positively correlated with RCO values computed based upon population density (Massey and Denton, 1998). Indeed, Massey and Denton do find this to be the case using their dataset on racial segregation in the US, although they find that an area-based measure of concentration displays a much wider range of variation than the density-based measure of concentration, which they suggest provides greater explanatory power when testing relationships between concentration and other demographic, social or economic outcomes.

For the analyses undertaken for this thesis I followed Massey and Denton's recommendation to compute the RCO based upon areal size rather than population density because datazones in South Africa were specifically designed to be homogenous with respect to population size. I do however reject Massey and Denton's simplistic approach with regard to the cut points *n1* and *n2* in favour of a more sophisticated interpolation approach in order to estimate the share of a unit's area that is inhabited by group *X* or group *Y*. I refrain from applying any recoding of values that fall outside the preferred -1.0 to 1.0 range in order to first assess the extent and magnitude of unboundedness.

The RCO analyses (presented in Appendix B) revealed that multiple municipalities in South Africa breached the -1.0 to 1.0 bounds when constructed for my purpose. I therefore conclude that, on a methodological basis, the RCO is not suitable for the purpose of measuring residential concentration of poor or non-poor population in South Africa at datazone level.

In addition to the methodological challenges posed by the RCO, it is also important to consider the conceptual basis upon which concentration might be deemed to be a relevant measure of people's lived experience of inequality. I would argue there is no clear link between population concentration and people's experience of inequality in the South African context. For example, in predominantly urban municipalities poor people often live in areas of high population density (such as townships) whereas the more affluent population often live in the less densely populated neighbourhoods. In this case, poor people may understandably feel that they are disadvantaged compared to the rich due to the unequal distribution of space (among other things). However, in the predominantly rural municipalities the more affluent typically live in and around the main town(s) which are characterised by higher population densities than the typically more deprived rural areas surrounding the town(s). In this case, poor people in the sparsely populated areas are unlikely to feel disadvantaged by lack of physical space available to them, but they are arguably much more likely to feel disadvantaged by a lack of access to services compared to the residents of the town(s). As such, although the dimension of concentration would be relevant if this thesis was concerned only with urban areas within South Africa, because I am also including rural areas I conclude that concentration is not appropriate here as it does not permit comparison of inequality between all neighbourhoods (urban and rural) in the country on a consistent basis.

Centralisation

Centralisation refers to the extent to which members of a particular population sub-group are residentially located close to the centre of an urban area. Massey and Denton review three measures of centralisation and conclude that the Absolute Concentration Index (ACE) is the best overall measure of this particular dimension of segregation.

The ACE can be interpreted as the proportion of *X* members required to change residence to achieve a uniform distribution of population around the central business district. The scores can range between -1.0 and 1.0, with a score of zero meaning that the *X* group has a uniform distribution around the central business district, positive values between zero and 1.0 indicate a tendency for *X* members to reside close to the central business district, and negative values between zero and -1.0 indicate a tendency for *X* members to reside in more outlying areas. Although Massey and Denton are not explicit about how they operationalise this measure in practice, Galster (1984) developed a similar centralisation index in which he defined concentric spatial rings around the central business and used these rings to judge proximity to the core.

Conceptually, the dimension of centralisation may well be a potential determinant of people's lived experience of inequality within South Africa. Unlike the historical pattern in the US and many European countries, where minority racial groups often found themselves living in poor quality housing close to the central business district through the selective outmigration of more affluent population to the suburbs, in South Africa access to housing close to the central business district was often restricted to the majority white (and typically more affluent) population during the colonial, segregationist and apartheid periods (Christopher, 1994). So whereas in the US and Europe living close to the central business district was (until the fairly recent emphasis on urban regeneration) often a sign of disadvantage, in South Africa the opposite might be said to be true. In South Africa, living close to an urban core typically increases access to services and opportunities for employment. So it is conceivable that people who live a considerable distance away from an urban core may have an acute sense of the level of inequality in access to services and job opportunities compared to people who live much closer to the urban core.

Unfortunately, whilst this dimension of residential segregation is potentially of relevance for this thesis, in practice it has proved impossible to operationalise given the current availability of data in South Africa. Whereas in the US studies, such as Massey and Denton's, each major city that they assessed had a well-defined central business district, in South Africa no such spatially referenced dataset exists. It is also debatable whether a suitable CBD location dataset could *ever* exist in practice, as many cities have multiple commercial/industrial nodes rather than a single CBD. For instance, in Johannesburg there is an area recognised as the central city but this has undergone a degree of urban decline since the advent of democracy whilst the suburb of Sandton has seen considerable urban growth and is now home to many multinational businesses.

Exposure

The fourth dimension, exposure, refers to the likelihood of potential contact and possible interaction between members of two different groups of the population. The standard exposure indices discussed by Massey and Denton quantify the extent to which one group (e.g. the poor) is exposed to the other group (e.g. the non-poor) by virtue of living in the same neighbourhood (Massey and Denton, 1988). Massey and Denton state that "exposure indices attempt to measure the *experience* of segregation" as felt by the average member of a particular group (Massey and Denton, 1988, p.287). As such, exposure indices are of particular relevance to my thesis objective of measuring the 'lived experience of inequality'.

Massey and Denton recommend the use of the so-called $P*$ indices of exposure that had been adopted by Lieberson (1981), consisting of the ${}_xP*_y$ and ${}_yP*_x$ interaction indices and their counterparts the ${}_xP*_x$ and ${}_yP*_y$ isolation indices (not considered here). The subscripts x and y denote the two groups of population being compared, such that for my purpose, x would represent the poor group and y would represent the non-poor group. In this context, the ${}_xP*_y$ index represents the exposure of a member of the 'poor' population to members of the 'non-poor' population, while the ${}_yP*_x$ index represents the exposure of a member of the 'non-poor' population to members of the 'poor' population.

The standard notation for the $P*$ interaction indices is as follows:

$$xP^*y = \sum_{i=1}^{n} \left[\frac{x_i}{X}\right]\left[\frac{y_i}{t_i}\right] \qquad (1)$$

$$yP^*x = \sum_{i=1}^{n} \left[\frac{y_i}{Y}\right]\left[\frac{x_i}{t_i}\right] \qquad (2)$$

where, for my purpose, x_i, y_i and t_i are the number of poor, non-poor, and total population (i.e. sum of poor plus non-poor) , respectively, in the neighbourhood i, and X and Y represent the total number of poor and total number of non-poor, respectively, in the wider geographical area. As described here, the ${}_xP*_y$ interaction index can be interpreted as the likelihood of a poor person sharing the same residential neighbourhood as a non-poor person, while the ${}_{y}P^{*}{}_{x}$ interaction index can be interpreted as the likelihood of a non-poor person sharing the same residential neighbourhood as a poor person. The respective neighbourhood interaction probabilities are averaged across all neighbourhoods within a larger geographical area to produce overall ${}_{x}P^{*}{}_{y}$ and ${}_{y}P^{*}{}_{x}$ indices for the larger geographical area, such as each of the US cities that were examined by Massey and colleagues (e.g. Massey and Denton, 1988; Massey and Denton, 1989; Massey and Eggers, 1993; Massey and Fischer, 2000; Massey and Fischer, 2003; Massey et al., 1996). They could similarly be created for each local municipality in South Africa. In this sense, and as was the case for the measures of evenness, concentration and centralisation considered above, the standard exposure measure recommended by Massey and Denton can be regarded as a 'global' measure of segregation following Lloyd and Shuttleworth's (2012) categorisation.

A key feature of the P^{*} exposure indices is that they are affected by the population composition of the broader geographical area (e.g. the proportion poor compared to the proportion non-poor in the city or the local municipality). As such, they address the criticism of compositional invariance that was levelled at the dissimilarity index above. As Lieberson and Carter (1982, p.299) argue, "Insofar as population composition actually affects the interaction between groups with the same indices of dissimilarity, for some research problems it is not desirable to ignore this factor". Furthermore, whereas the dissimilarity index is symmetrical (i.e. there is only one dissimilarity index value between poor and non-poor), the P^{*} exposure indices are asymmetrical, acknowledging that the exposure of poor to non-poor is rarely the same as the exposure of non-poor to poor (Lieberson and Carter, 1982; Massey and Denton, 1988).

In the analyses that follow, $P*$ exposure indices have been constructed using datazone level counts of 'poor' and 'non-poor' and these datazone level input data have been averaged over local municipalities to produce exposure scores for all local municipalities in South Africa. Figure 4.1 and Figure 4.2 show the results on the ${}_xP*_y$ and ${}_yP*_x$ exposure indices respectively. The municipalities are divided into national deciles, so that the highest exposure decile of municipalities is colour-coded dark blue, while the lowest exposure decile of municipalities is colour-coded bright yellow.

Figure 4.1: Exposure index $_{x}P^{*}_{y}$ based on data derived from the income and material deprivation domain of the SAIMD 2001

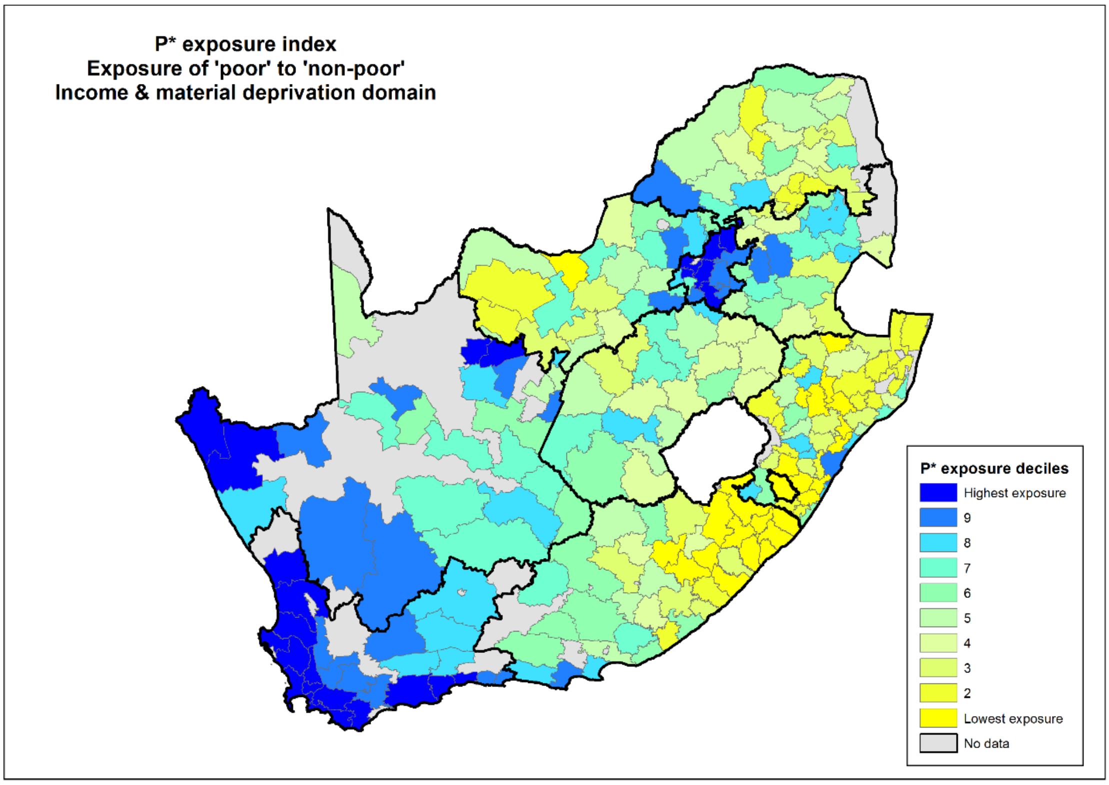
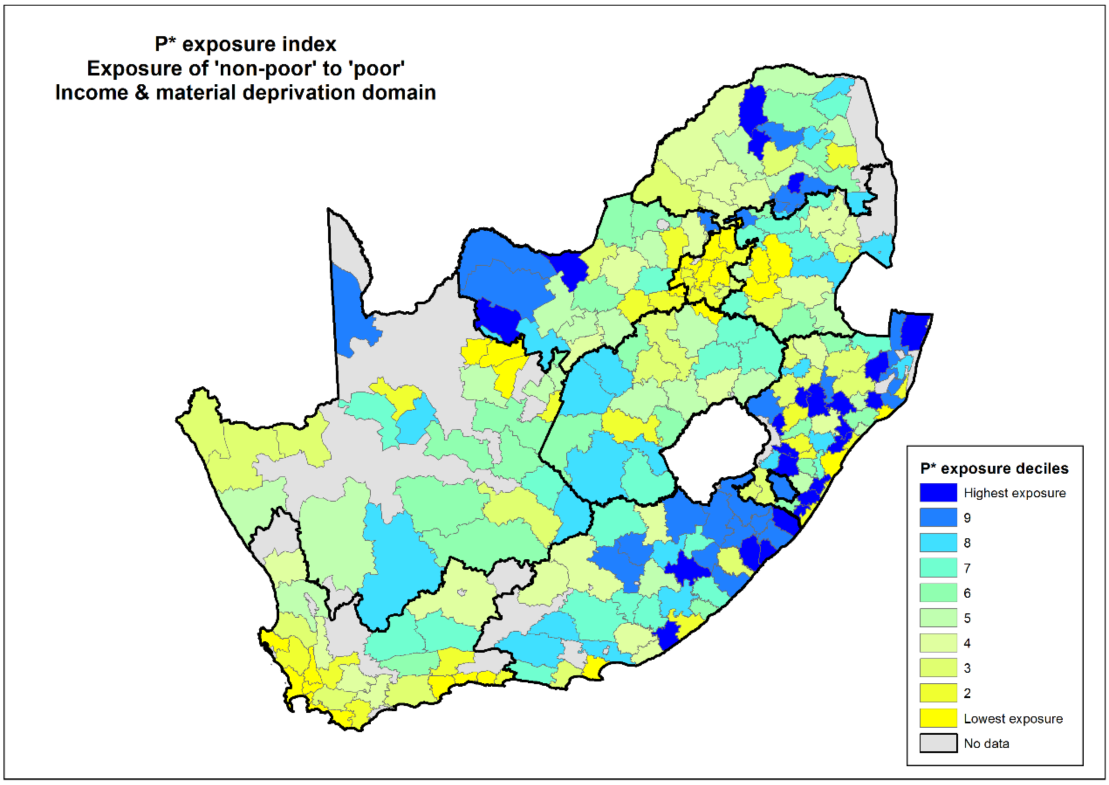


Figure 4.2: Exposure index $_{y}P^{*}_{x}$ based on data derived from the income and material deprivation domain of the SAIMD 2001



With regard to the pattern of exposure experienced by the poor population (i.e. ${}_xP^*{}_y$), it is evident from Figure 4.1 that the highest exposure municipalities are mainly located in the Western Cape, Northern Cape and Gauteng, whilst the lowest exposure municipalities are mainly located in the Eastern Cape and KwaZulu-Natal. In contrast, when looking at the pattern of exposure experienced by the non-poor population (i.e. ${}_yP^*{}_x$) shown in Figure 4.2, it is evident that the highest exposure municipalities are mainly located in the Eastern Cape, KwaZulu-Natal, North West and Limpopo, while the lowest exposure municipalities are mainly found in and around Gauteng and in the Western Cape. However, across both maps there is evidence of variation in municipality level exposure within provinces as well as between provinces.

However, a major limitation of the conventional exposure indices is that exposure is only measured within the confines of the home neighbourhood, and therefore any potential exposure that might take place away from the home neighbourhood is ignored. Whilst there is evidence that people tend to undertake their routine daily activities in a relatively confined geographical space close to their home (Johnston and Pattie, 2011), applying a restriction to limit people's potential exposure to only within their home neighbourhood does not reflect the full reality of people's actual activity patterns. Fortunately, for my purpose, Massey and Denton also identified a variant on the basic P^* exposure indices that does take into account people's potential exposure in neighbourhoods other than their home neighbourhood. However, Massey and Denton classified this variant under their 'clustering' dimension, which I shall discuss next.

Clustering

The fifth dimension, clustering, refers to the degree to which neighbourhoods populated by a particular group of the population are geographically contiguous or proximately located. Clustering therefore concerns the spatial distribution of neighbourhoods with respect to one another i.e. taking into account the spatial inter-relationships between neighbourhoods (Massey and Denton, 1988). For instance, where high-poverty neighbourhoods are contiguous or closely packed within a certain part of a city or local municipality, the degree of spatial clustering is high. Conversely, where high poverty neighbourhoods are dispersed in and amongst low poverty neighbourhoods the degree of clustering is low. White (1983, p.1010) refers to this situation of spatial clustering/dispersion as the "checkerboard problem". The fundamental principle of the clustering dimension is that the measures take account of the spatial relationships *between* neighbourhoods. In the conventional *P** measures of exposure discussed above, the spatial proximities between neighbourhoods are ignored. Only in the clustering dimension of Massey and Denton's classification is the spatial interconnectedness of neighbourhoods taken into account.[27]

Although Massey and Denton's chosen measure of segregation under the 'clustering' dimension was White's Index of Spatial Proximity, they also referenced a particular variant of the basic *P** exposure index that I discussed above: the *DP** distance-weighted exposure indices proposed by Morgan (1983). These distance-weighted exposure indices are an adaptation of the conventional *P** exposure indices discussed above to incorporate people's potential exposure in neighbourhoods other than their home neighbourhood. The *DP** exposure indices therefore better reflect the reality

[27] It should be noted however that other measures such as the dissimilarity index can be constructed to have spatially weighted or local variants, but Massey and Denton did not do this.

that individuals in one group (say, the poor) may be exposed to members of another group (say, the non-poor) in both their own neighbourhood and in a range of other proximate locations as they go about their daily lives.

The distance function within the *DP** indices can be specified in a variety of ways to reflect the expected likelihood of a person from one neighbourhood visiting each other neighbourhood within a given geographical bound, and is usually derived through the construction of a neighbourhood-to-neighbourhood distance matrix (White, 1983). Massey and Denton placed these indices within the clustering dimension of their classification because they utilise information on the spatial interconnectedness of neighbourhoods. However, I argue that they straddle the dimensions of exposure and clustering.

White (1983) constructed four versions of his *DP** exposure indices, varying in the specification of the distance function (including linear and exponential variants) and concludes that:

> "In the absence of some theory which determines the exact nature of the [distance decay] function, PE [i.e. standard exponential distance decay] is suggested as the "best" measure. It is defined at all distances...and it declines faster than distance itself" (White, 1983, p.1015).

Massey and Denton (1988) follow White's suggested approach and also adopt a standard exponential distance decay approach. The formula for the DP_{xy}* and DP_{yx}* measures as proposed by Massey and Denton are as follows:

$$DP_{xy}{}^{*} = \sum_{i=1}^{n} \frac{x_i}{X} \sum_{j=1}^{n} K_{ij} \left(\frac{y_j}{t_j}\right) \quad (3)$$

$$DP_{yx}{}^{*} = \sum_{i=1}^{n} \frac{y_i}{Y} \sum_{j=1}^{n} K_{ij} \left(\frac{x_j}{t_j}\right) \quad (4)$$

where

$$K_{ij} = \frac{\exp(-d_{ij})t_j}{\sum_{i=1}^{n} \exp(-d_{ij})t_j} \tag{5}$$

and where d is the distance between area i and area j. For each area i, the probability of a poor person being exposed to a non-poor person is a function of the likelihood of a poor person from area i visiting area j combined with the likelihood of a poor person meeting a non-poor person once within area j.

In the analyses that follow, DP^* exposure indices have been constructed using datazone level counts of 'poor' and 'non-poor' and these datazone level input data have been averaged over local municipalities to produce distance-weighted exposure scores for all local municipalities in South Africa. Figure 4.3 and Figure 4.4 below show the results of the DP_{xy}^* and DP_{yx}^* exposure indices, respectively. Again, the municipalities in South Africa have been ranked based upon the relevant measure and divided into decile groups, each of which has been assigned a different thematic colour, ranging from dark blue for the highest exposure decile to bright yellow for the lowest exposure decile. It is evident from Figure 4.3 that the highest exposure municipalities on the DP_{xy}^* measure tend to be in the Western Cape and Gauteng. This is as we might expect given the analyses in Chapter 3 where I demonstrated that these two provinces contained particular concentrations of less deprived neighbourhoods. For a poor person living in municipalities within the Western Cape or Gauteng, they are likely to experience inequality regularly as they go about their daily lives. In contrast, a poor person living in the mainly rural municipalities in the Eastern Cape is much less likely to be exposed to non-poor people as they go about their daily lives.

Figure 4.3: Exposure index DP_{xy}*: exposure of poor to non-poor

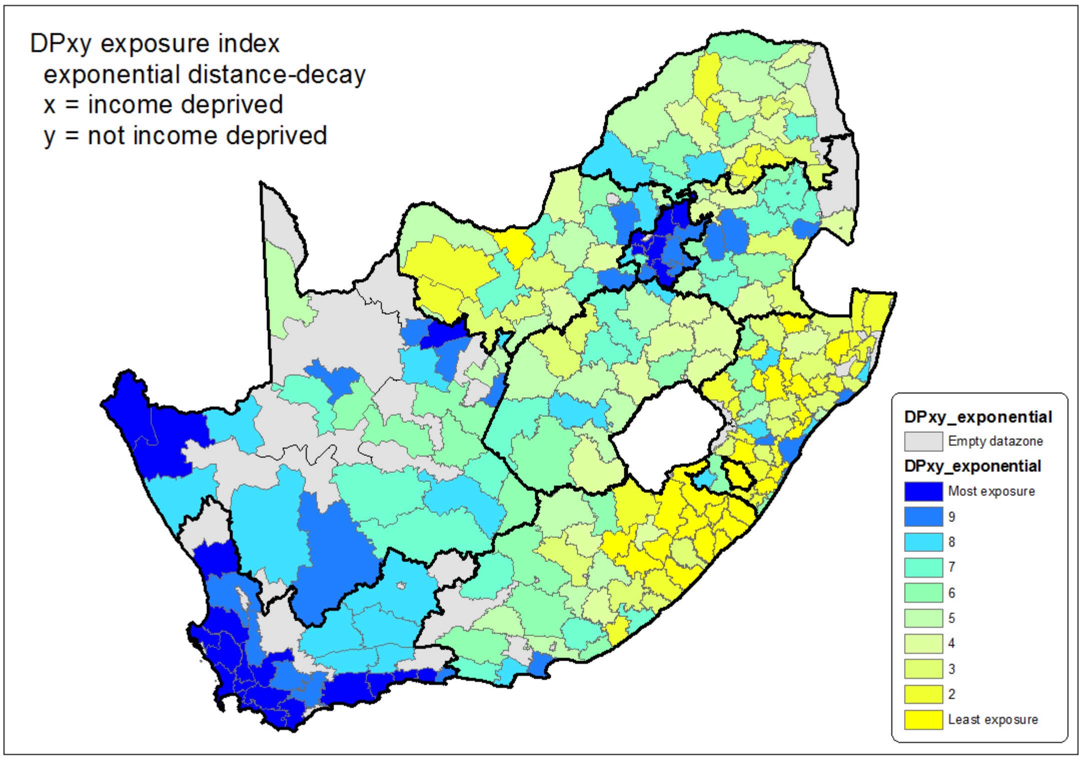
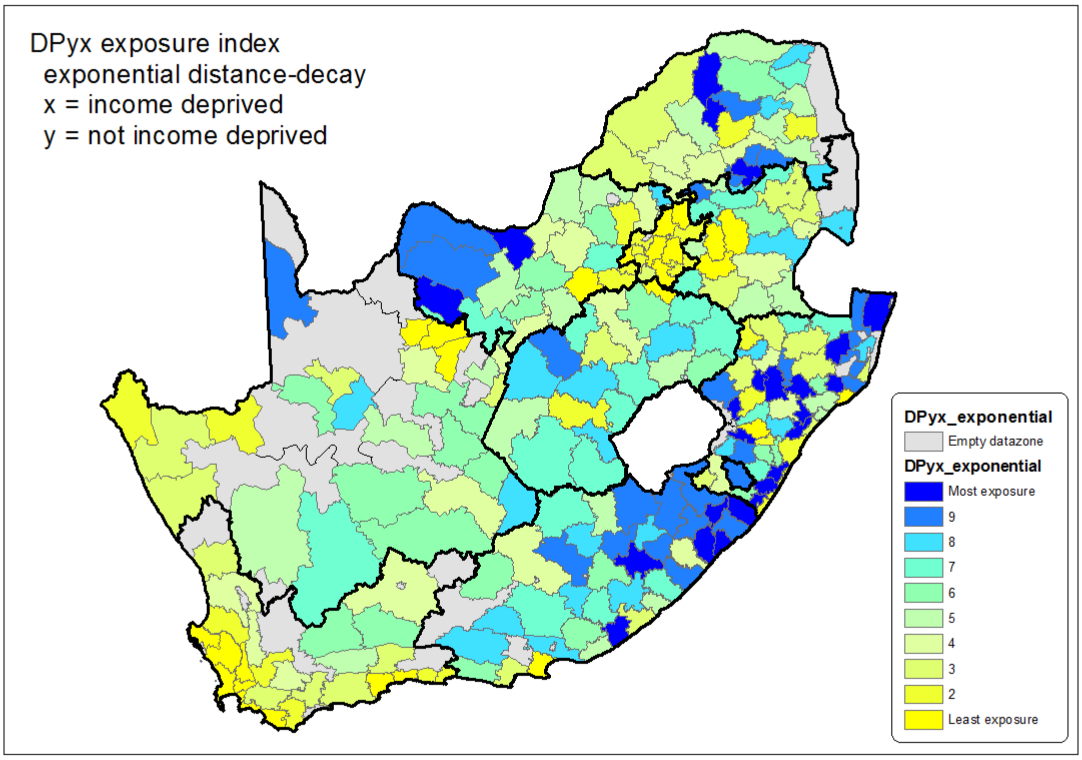


Figure 4.4: Exposure index DP_{yx}*: exposure of non-poor to poor



The picture on the DP_{yx}* index is roughly the opposite to that observed for the DP_{xy}* index, in that the highest exposure municipalities are typically those which we saw in Chapter 3 had the highest levels of deprivation, such as within the former homelands of the Eastern Cape and KwaZulu-Natal. In this sense, a non-poor person living in an area such as a former homeland would likely be exposed to poor members of the population regularly as they went about their daily lives due to the extremely high rates of deprivation in these areas.

The review of possible methodologies summarised above suggests that the distance-weighted *DP** exposure indices, DP_{xy}* and DP_{yx}*, are the most appropriate starting point for my purpose. The indices measure the level of exposure to inequality experienced first-hand by people as they go about their daily lives, with exposure occurring not only within their home neighbourhood but also throughout the wider geographical area in which they carry out their routine activities. Referring back to Lloyd and Shutleworth's (2012) distinction between different forms of segregation measure, the *DP** indices are a form of 'spatial' measure whereas the basic *P** indices are a form of 'global' measure.

However, a major limitation of the *DP** indices in its 'spatial' form (which is also a one of the limitations of the 'global' *P** indices) is that each *DP** index consists of a single summary statistic for each municipality which means that exposure is assumed to be uniform across the municipality irrespective of where within that municipality a person lives. For my purpose in this thesis I wish to develop a more geographically nuanced measure of the lived experience of inequality which varies according to the neighbourhood in which a person lives, not simply the municipality in which they live. As Lloyd and Shuttleworth (2012) discuss, this requires a 'local' form of residential segregation measure. In the remainder of this chapter I will proceed to develop a local form of the distance weighted exposure index which I argue better reflects the lived experience of inequality for residents of South Africa and which varies at neighbourhood level.

4.3 Developing a local deprivation-adjusted distance-weighted *P** exposure index.

The ‘local’ variant of the distance weighted exposure index is preferred because the aim is to measure inequality experienced by individuals and I argue that people's level of exposure to inequality will vary spatially within a city or other large area according to where within that large area (i.e. in which neighbourhood) they live. Local measures of exposure give a more nuanced insight into the patterns of segregation at a detailed local level than is provided through spatial (geographically weighted) measures of exposure (Lloyd, 2010; Lloyd and Shuttleworth, 2012; Wong, 2002).

The local distance weighted exposure indices, represented here by the notation $LDP_{xy}{}^*$ and $LDP_{yx}{}^*$, are in fact equivalent to equations (3) and (4) above, but without the final population weighted aggregation to a higher spatial level. They can therefore be expressed as:

$$LDPxy_i{}^* = \sum_{j=1}^{n} K_{ij} \left(\frac{y_j}{t_j}\right) \quad (6)$$

$$LDPyx_i{}^* = \sum_{j=1}^{n} K_{ij} \left(\frac{x_j}{t_j}\right) \quad (7)$$

where the *K* distance decay function can again be specified in a variety of different ways.

As outlined above, I used data on income and material deprivation at neighbourhood level in South Africa, taken from the 2001 South African Index of Multiple Deprivation (Noble et al., 2009a). For each datazone it is possible to quantify the number of people who are deprived in terms of income and material deprivation (i.e. ‘poor’) and the number of people who are not deprived in terms of income and material deprivation (i.e. ‘non-poor’). The total population of the datazone is the sum of those deprived and those not deprived (i.e. deprivation status is a straight binary classification). For the purpose of describing the methodology in this chapter I will refer only to the poor/non-poor classification based upon the Income and Material Deprivation Domain but, as will become evident through the analyses contained within Chapters 5 and 6, I also subsequently constructed equivalent exposure measures using the three other appropriate domains of Employment, Education and Living Environment.

Specifying the parameters

Before embarking on the development of this new measure, two major methodological issues required consideration. The first relates to the spatial bounds within which potential exposure is judged likely to occur. The second relates to the spatial weights assigned to the neighbourhoods to reflect the likelihood of a person from one neighbourhood actually visiting each other neighbourhood within the defined spatial bounds. These are discussed in turn as they represent important components of the exposure measure that I present in the following section.

Spatial bounds

The question of how to define the spatial bounds within which a person is likely to experience inequality first-hand is of central importance to the measure of exposure developed here. It is, of course, intrinsically linked to the second methodological consideration relating to the derivation of spatial weights to represent the likelihood of a person from one neighbourhood visiting every other neighbourhood. The spatial bounds should only encompass those neighbourhoods which a person is likely to visit, and then the spatial weights should reflect the actual likelihood of visiting each of these selected neighbourhoods. This would require a bespoke spatial bound around each separate neighbourhood demarcating the geographical limits of potential exposure of that neighbourhood's residents, and reflecting the reality that people's travel patterns when outside their home neighbourhood are neither random nor uniform, but instead determined by a combination of needs (e.g. access to services), opportunities (e.g. access to jobs) and barriers (e.g. travel time, cost and perceived difficulty of integration) (Schnell and Yoav, 2001).

In practice, three approaches to setting spatial bounds around each neighbourhood have typically been used in the literature concerning local measures of segregation: (i) selecting all neighbourhoods that lie within the administrative area in which the target neighbourhood is located; (ii) selecting all neighbourhoods that lie within a set distance radius of the target neighbourhood; and (iii) selecting a pre-specified number of neighbourhoods using a 'k nearest neighbours' approach based upon distance from the target neighbourhood (Anselin, 1995).

In the spatial (geographically weighted) measures of exposure presented under Massey and Denton's clustering dimension, the spatial bounds are defined as the city within which the person lives. In other words, each resident of a given city is exposed to the inequality apparent within the bounds of that city but they are oblivious to any inequality outside the city bounds. For the analysis in this paper, because I am concerned with inequality experienced in both urban and rural areas, the equivalent spatial units would be local municipalities and (where relevant) metropolitan areas (Statistics South Africa, 2004a). I argue here that choosing local municipalities/metropolitan areas as the spatial bounds of potential exposure would have one major advantage, counteracted by two major disadvantages. The advantage would be that every area, even if rural, will contain at least one town (often the administrative capital) which will exhibit 'pull factors' to attract people from the surrounding parts of the municipality (e.g. access to service hubs and increased employment opportunities) and which will be relatively well connected to the local transport infrastructure. This is important because in South Africa the rural areas are typically characterised by higher levels of deprivation than the urban areas and so within a rural municipality the town is most likely to contain concentrations of non-poor population. As such, poor people living in the rural areas are perhaps most likely to experience inequality first-hand through their visits to towns. The first disadvantage to choosing local municipalities as the spatial bounds of potential exposure would be that this effectively considers each municipality as having 'closed borders' which prevent people from one municipality from visiting any other municipality, which is wholly inappropriate, particularly for residents of neighbourhoods located close to municipal borders. The second disadvantage is that local municipalities vary greatly in terms of population size and in terms of geographical size, meaning the bounds of potential exposure will vary greatly from person-to-person depending on the municipality in which they live.

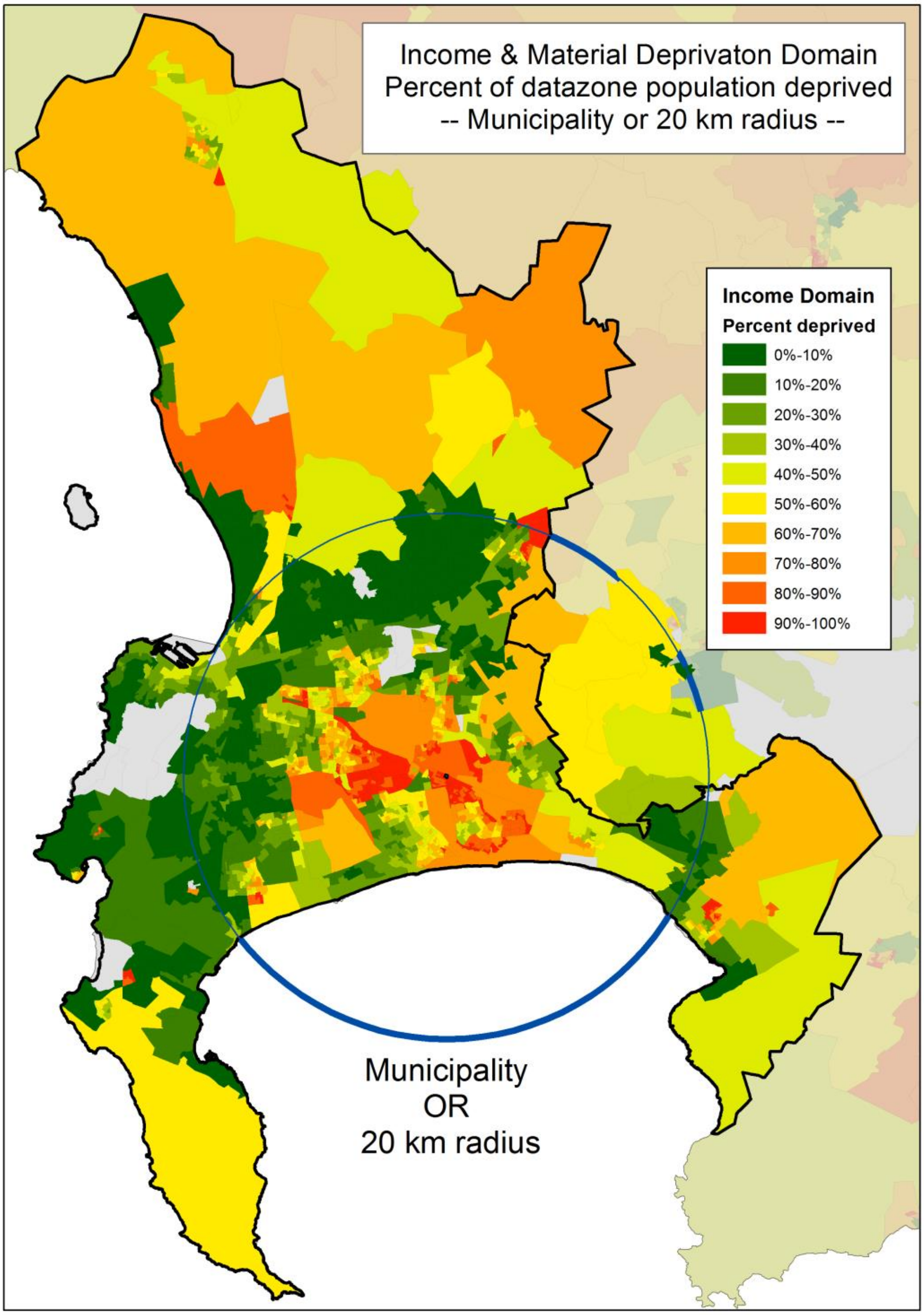
Other applications of local segregation indices use either the set distance radius approach or the *k* nearest neighbours approach, neither of which is dependent upon the often fairly arbitrary definition of municipal boundary lines. This independence from municipal boundary definitions is a major advantage underpinning both approaches as it reflects the reality that people living close to municipal boundaries are likely to be exposed to inequality in more than one local municipality. However, the set distance radius and *k* nearest neighbours approaches are also dependent upon the fairly arbitrary specification of parameters (in terms of the distance threshold or *k* value, respectively). Fixing the distance threshold results in widely varying *k* neighbours, while fixing the *k* value results in widely varying distances. As such, neither approach offers the ideal solution for my purpose. A further disadvantage of both the set distance radius and the *k* nearest neighbours approaches for my purpose is that both specify the spatial bounds without reference to the *types* of neighbourhoods that are being included within the bounds. In both cases, the final geographical extent of the spatial bounds around any given target neighbourhood will be roughly circular because both approaches utilise simple Euclidean distance, usually measured in all directions from the centroid of the target neighbourhood to the centroid of the proximate neighbourhoods, as the basis for selection. They do not reflect the reality that people's travel patterns are contoured by a combination of needs, opportunities and barriers which may not radiate out in a uniform circular way around the home neighbourhood.

Each of the approaches to defining spatial bounds discussed above was developed and tested in order to inform the selection of the most suitable approach for a measure of the lived experience of inequality. The approach I selected is a combination of the local municipality administrative geography and the set distance radius approach. For a resident of any given target neighbourhood, I assume that potential exposure to inequality may occur within that same neighbourhood, plus any other neighbourhood within the same local municipality, plus any neighbourhood that lies within a 20km distance radius of the target neighbourhood, irrespective of municipal boundaries[28]. My chosen approach to defining the spatial bounds around each neighbourhood draws upon the strengths of the local municipality geography, particularly that each municipality will include at least one administrative town, and the strengths of the set distance radius approach , particularly that spatial bounds can cross municipal boundaries. The choice of 20km as the set distance radius is arbitrary, but based upon what might be regarded as a manageable daily commute for most residents of South Africa.

Figure 4.5 shows the extent of the spatial bounds around a selected datazone within the Khayelitsha township in Cape Town. A resident of this particular datazone is assumed to encounter potential exposure to inequality within the brightly colour datazones that fall within the municipality boundary and/or within the 20 km radius buffer around the Khayelitsha datazone. The areas that fall outside these bounds have been dimmed to indicate that they are outside the geographical bounds of potential exposure for residents of this particular datazone. The detail of the thematic colour-coding is included in this map simply for context, and serves no other purpose than giving a sense of the types of areas that a resident of this Khayelitsha datazone might potentially visit as they go about their daily lives.

[28] Note that the distance calculation is based on Euclidean distance between the respective datazone centroids. These calculations were undertaken in Stata.

Figure 4.5: An example of the spatial bounds within which exposure is assumed to potential occur for a resident of the highlighted datazone in Khayelitsha, Cape Town



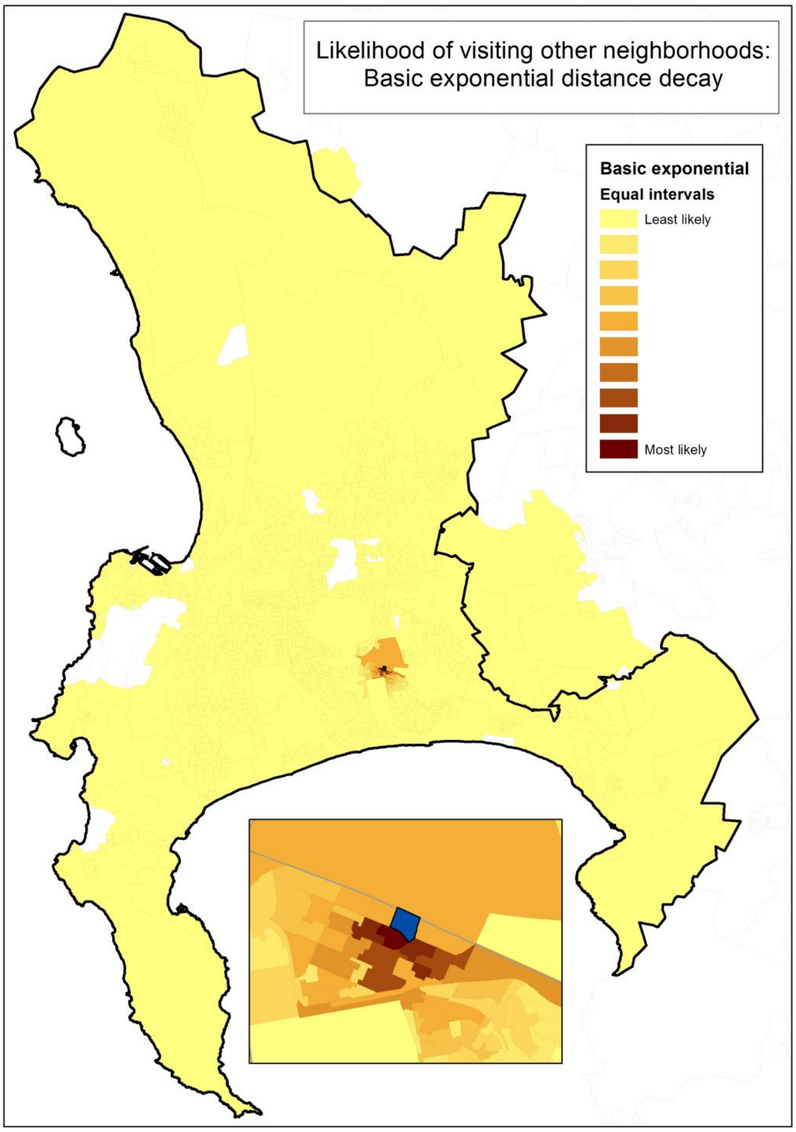
Spatial weights

The second important methodological consideration relates to the derivation of spatial weights assigned to the neighbourhoods to reflect the likelihood of a person from one neighbourhood actually visiting each other neighbourhood within the defined spatial bounds. In Massey and Denton's distance-weighted *DP** exposure indices they followed White’s (1983) recommendation and used an exponential distance decay function, radiating out evenly in all directions from the target neighbourhood to every other neighbourhood within their specified spatial bounds (US cities). Despite White’s (1983) recommendation to use the exponential distance decay function for K_{ij}, I argue here that the choice of distance decay approach should be considered on a case-by-case basis depending upon the specific research challenge.

Figure 4.6 below shows how the spatial weights would be configured for residents of the same Khayelitsha datazone that was discussed above in relation to setting the spatial bounds if the spatial weights were defined using exponential distance decay.

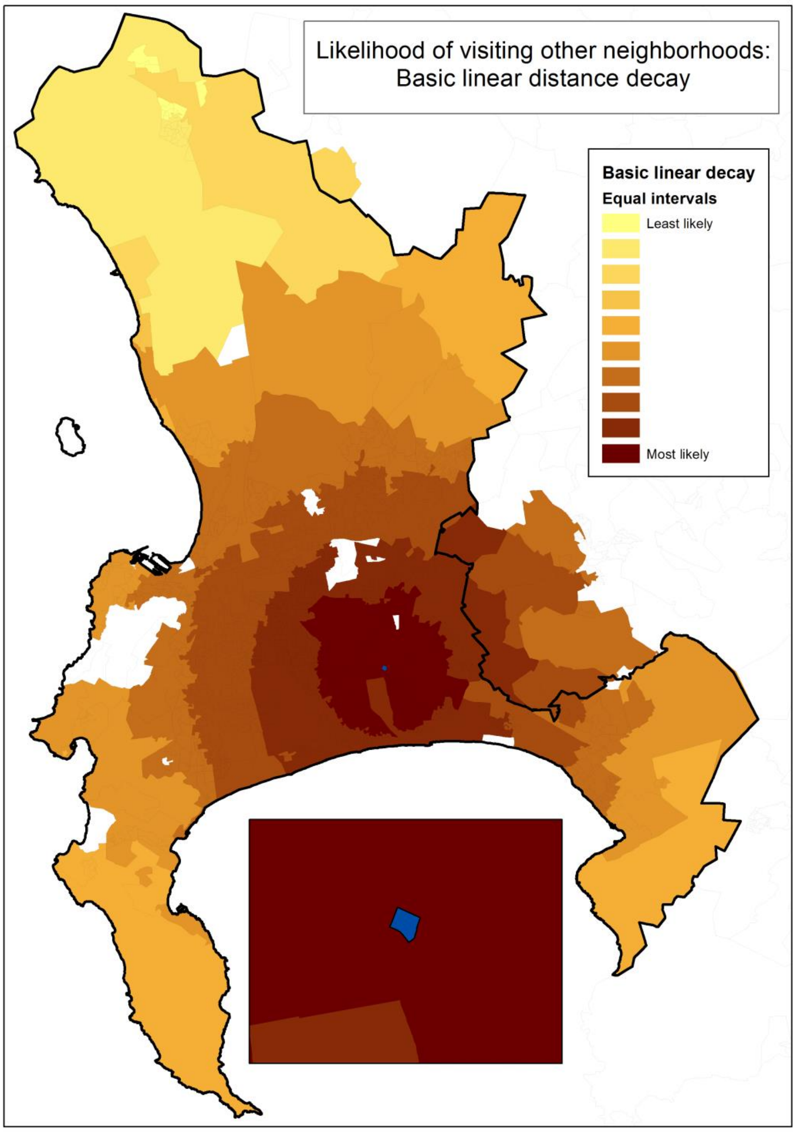
I decided not to use the exponential distance decay function for the analyses presented in this thesis because it places too great an emphasis on the immediate vicinity around the target neighbourhood which does not reflect the reality in South Africa that people often have to travel quite considerable distances (such as from a township on the periphery of a urban area into the central business district) in order to access services or employment opportunities.

Figure 4.6: Example of exponential distance decay weights



I therefore begin with a linear distance decay function rather than an exponential distance decay function. The linear function still attributes greatest weight to the neighbourhoods closest to the target neighbourhood, but the weights do not decrease as quickly with increasing distance as is the case with the exponential. Given the fact that in South Africa many people have to travel considerable distances to work or to seek work, the linear approach is judged more appropriate for my purpose than the exponential. Figure 4.7 shows the configuration of spatial weights for residents of the selected Khayelitsha datazone.

Figure 4.7: Example of linear distance decay weights



A major criticism of the basic linear distance decay weight approach (which also applies to the basic exponential approach) is that the decay function is based solely on the parameter of Euclidean distance. However, I argue that in reality people's travel patterns are contoured by a combination of needs and opportunities, whilst being constrained by travel times and costs (which can be proxied by geographical distance) (Schnell and Yoav, 2001). I derived a set of adjustment factors which I applied to the weights generated using the basic linear distance decay approach and which I propose results in a more realistic set of distance weights The major assumption underpinning the adjustment factors is that the more prosperous neighbourhoods will exert a stronger set of 'pull factors' than the less prosperous neighbourhoods because, in the South African case, the more prosperous areas tend to be closer to key service access points and tend to offer the greatest opportunities in terms of potential employment prospects. Conversely, and particularly in urban areas, the less prosperous neighbourhoods are often *perceived* (whether correctly or incorrectly) to be places of disorder, crime and violence (and such perceptions are supported by my own analysis of spatial patterns of homicide rates at police precinct level which demonstrates that the highest crime precincts are often also highly deprived). My proposition is that the likelihood of a person from one neighbourhood visiting another neighbourhood is based upon a combination of the distance between the respective neighbourhoods and the relative levels of prosperity in the respective neighbourhoods, with more prosperous areas being more attractive as a day-to-day travel destination.

Although no data currently exist to measure levels of prosperity, affluence or wealth at small area level in South Africa, the proportion of population experiencing income and material deprivation at small area level is likely to be strongly negatively correlated with levels of prosperity, affluence or wealth. I therefore use deprivation differentials to adjust the linear distance decay weights to produce the final set of weightings.

I generated the spatial weight adjustment factors as follows: first, I calculated the proportion of the population in the target neighbourhood that is *not* defined as experiencing income and material deprivation (i.e. a value of 0 means every person in that neighbourhood is deprived and a value of 1 means that no one in that neighbourhood is deprived). Second, I repeated this calculation for every other neighbourhood within the limits of the spatial bounds around the target neighbourhood. Third, I took each of the neighbourhoods within the spatial bounds in turn and calculated a deprivation differential by subtracting the target neighbourhood's non-deprived proportion from the equivalent non-deprived proportion of each respective neighbourhood. This generated a distribution that can range from a minimum of -1 (in the case where no one in the target neighbourhood is deprived and everyone in the proximate neighbourhood is deprived) to a maximum of +1 (in the case where everyone in the target neighbourhood is deprived and no one in the proximate neighbourhood is deprived). If the target neighbourhood and proximate neighbourhood have identical proportions of population deprived, the deprivation differential value will be zero. Fourth, I multiplied the deprivation differential value by the basic linear distance decay weight calculated above to produce an adjustment factor. Fifth, I added the adjustment factor to the basic linear distance decay weight value to apply the adjustment. Sixth, and finally, I scaled the spatial weights so that they sum to 1 across all neighbourhoods within the spatial bounds, thus generating the final set of deprivation-adjusted spatial weights.

Equation (8) provides the notation for steps 1 through to five described in this paragraph, while equation (9) provides the notation for the final step of scaling the weights so that they sum to one.

$$W_{ij} = \frac{(\max(d_{ij}) - d_{ij})}{\sum_{j=1}^{n} \llbracket \max(d_{ij}) - d_{ij})} + \left(\frac{\max(d_{ij}) - d_{ij}}{\sum_{j=1}^{n} \llbracket \max(d_{ij}) - d_{ij})} \times \left(\frac{y_j}{t_j} - \frac{y_i}{t_i} \right) \right) \tag{8}$$

$$Z_{ij} = W_{ij} \times \frac{1}{\sum_{j=1}^{n} W_{ij}} \qquad (9)$$

where d_{ij} is the distance between area i and area j, $max(d_{ij})$ is the maximum distance from area i to any other area j within the specified spatial bounds (i.e. within the local municipality and within 20km of area i), y and t are the numbers of non-poor population and total population, respectively, in areas i and j, W_{ij} is the pre-scaling deprivation-adjusted distance weight between areas i and j, and Z_{ij} is the final scaled deprivation-adjusted distance weight between areas i and j.

The end product of the deprivation-adjustment process is a set of spatial weights which I argue better reflects the likely routine activity patterns of individuals in South Africa than is provided by a simple distance decay function alone, meaning that the deprivation-adjusted spatial weights are a more realistic measure of the likelihood of an individual from one neighbourhood visiting various other neighbourhoods where he or she may be exposed to inequality[29].

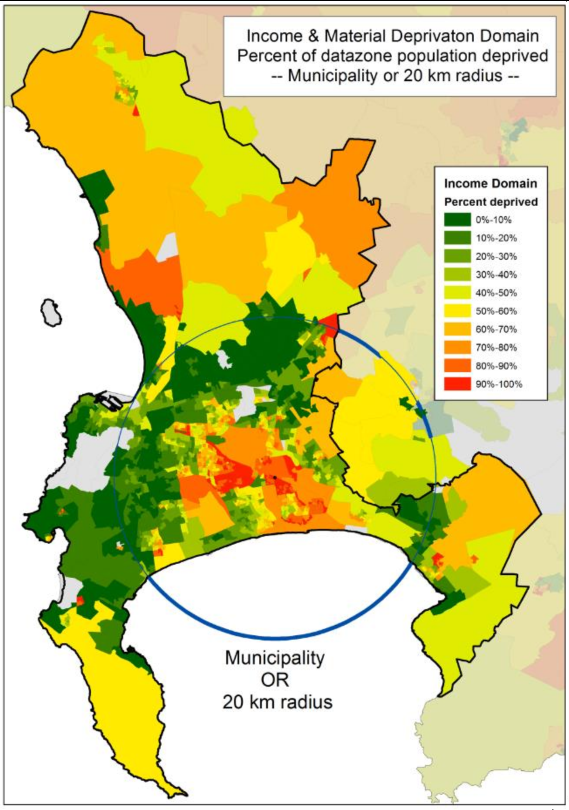
Figure 4.8 and Figure 4.9 below relate to two selected datazones within Cape Town which highlight the results of the spatial bounds and spatial weights approaches adopted here. Panel A of Figure 4.8 repeats the map shown in Figure 3.20 above but is presented here again for context and to allow easy comparison with Panel B of Figure 4.8, which shows the final spatial weights for all datazones that fall within the defined spatial bounds around the selected datazone in Khayelitsha. The distance decay function of the spatial weights is still evident, with the more distant datazones typically having

[29] As noted earlier in this Chapter, during the methodological development process for this thesis I tested a series of different approaches to defining the spatial bounds and associated weighting functions. I acknowledge that the alternative options that I rejected will each produce a different set of exposure scores. However, my review of the results from the respective tests (not presented here in the thesis) indicate that the same broad patterns of exposure are observed if the parameters are varied within a reasonable spectrum.

lower weights than more proximate datazones, but the deprivation-adjustment applied to the distance decay weights has resulted in a more nuanced configuration whereby the more prosperous (or rather, the less deprived) areas are exerting stronger pull factors than the more deprived areas of equal distance away.

Figure 4.8: Example of deprivation-adjusted linear distance decay weights: highlighted Khayelitsha datazone

Panel A



Panel B

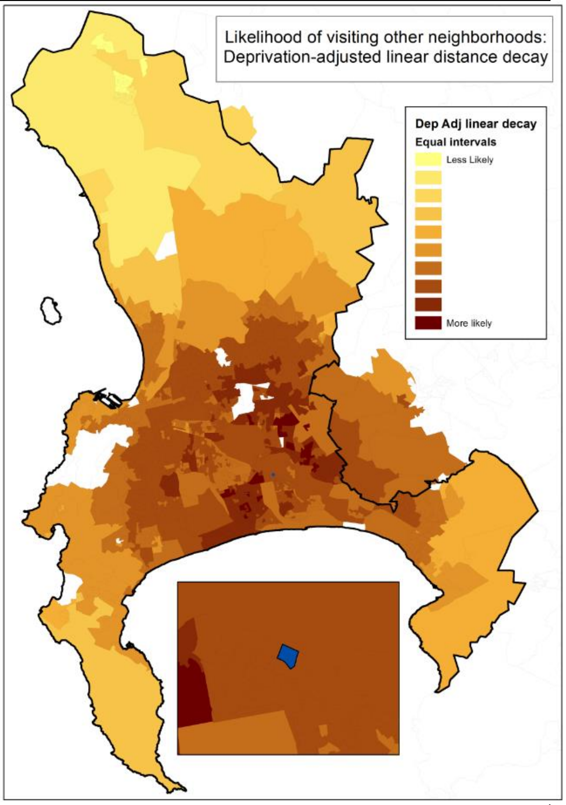
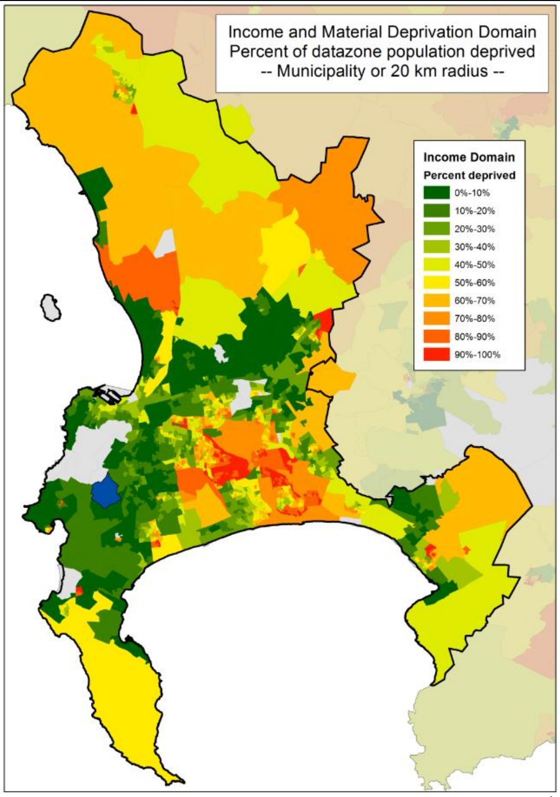


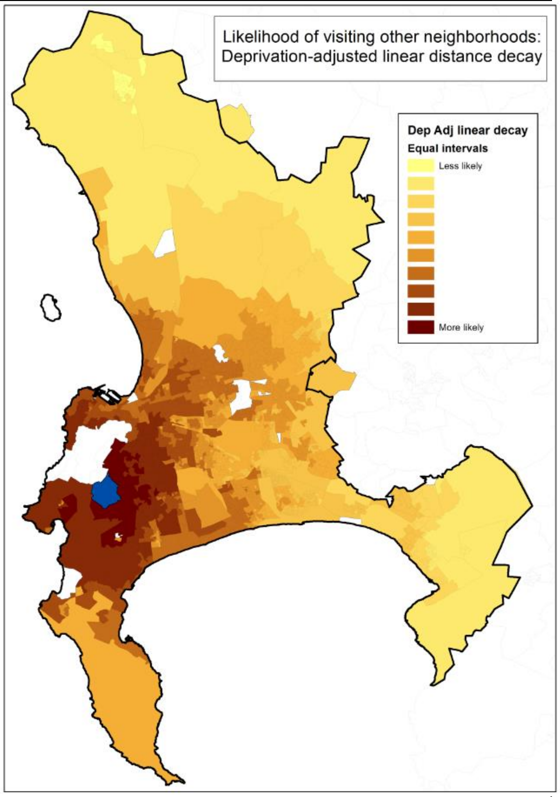
Figure 4.9 shows equivalent panels of data but for a selected datazone in Constantia, which is one of the most prosperous areas of South Africa. Here we see that the areas of potential exposure to inequality for residents of Constantia are largely in and around the southern suburbs of Cape Town, with much lower likelihood of visiting the deprived outer-lying areas such as Khayelitsha.

Figure 4.9: Example of deprivation-adjusted linear distance decay weights: highlighted Constantia datazone

Panel A



Panel B



In summary, my local deprivation-adjusted distance-weighted exposure indices for each area *i* can be depicted as follows:

$$aLDPxy_i{}^* = \sum_{j=1}^{n} Z_{ij}\left(\frac{y_j}{t_j}\right) \tag{10}$$

$$aLDPyx_i{}^* = \sum_{j=1}^{n} Z_{ij}\left(\frac{x_j}{t_j}\right) \tag{11}$$

where Z_{ij} is specified as shown in equation (9), x, y and t are the numbers of poor population, non-poor population and total population, respectively, in areas i and j, and $aLDPxy_i*$ and $aLDPyx_i*$ are the final local deprivation-adjusted distance-weighted exposure indices for each area i.

4.4 Conclusion

The aim in this chapter was to address Research Sub Q2: Can an empirical measure be developed that reflects people's lived experience of inequality? I contend that my deprivation adjusted distance weighted local exposure index satisfies the requirements for a spatial measure of the lived experience of inequality in South Africa.

The first criterion was that my chosen measure should permit comparisons of inequality between all neighbourhoods in the country on a consistent basis. My exposure measure is designed in such a way that relevant methodological parameters vary on an area-by-area basis in order to reflect the socio-spatial configuration of neighbourhoods. This results in an exposure measure that is equally applicable across all parts of South Africa irrespective of area type or geographical location.

The second criterion was that my chosen measure should reflect the experience of inequality from an individual's perspective. My exposure measure is designed in such a way that the neighbourhood exposure value reflects the likelihood of a given person from that neighbourhood interacting with people from the other end of the socio-economic spectrum as they go about their daily lives. Furthermore, my measure consists of two parts: the *aLDPxy** index measures exposure to inequality from the perspective of a 'poor' person, whilst the *aLDPyx** index measures exposure to inequality from the perspective of a non-poor person. As such, this measure acknowledges that people's socio-economic status may be an important determinant in their potential exposure to inequality. For instance, whereas an affluent suburban city dweller may be able to live his or her life without ever needing to visit the extremely poor townships within the city, a poor township resident may have no option but to travel to the more affluent areas in order to work, seek work or carry out other necessary activities. My exposure measure is designed to accommodate the differential magnitudes of push/pull factors linked to differential financial resources.

Finally, the third criterion was that data to calculate the measure must be available at the sub-municipality level. My exposure measure is designed to utilise appropriate small area level datasets that consist of binary categorisations between ‘poor’ and ‘non-poor’ and for my purpose here I have developed the measures using input data drawn from the SAIMD 2001 at datazone level. In the process of initially developing the measures described in this chapter I utilised the Income and Material Deprivation Domain of the SAIMD 2011 at datazone level but, as will become apparent in Chapter 4, I also then repeated the exposure measures using the Employment, Education and Living Environment Domains.

My exposure indices therefore satisfy all three criteria and, as such, I conclude in relation to Sub Q2 that it is indeed possible to construct an empirical measure that reflects people’s lived experience of inequality. However, I also acknowledge that certain components of the methodological basis for my exposure measures are based on researcher judgement in the absence of any suitable empirical data. I discuss this further in Chapter 8 and offer recommendations for future research that would help to validate the existing researcher judgements and/or provide a basis for refining the approaches taken.

Chapter 5: Spatial patterns of exposure to inequality

5.1 Aims and objectives

The new measures of exposure to inequality developed in Chapter 4 provide a geographically nuanced picture of how people's exposure to socio-economic inequality is influenced by where they live and where they carry out their routine daily activities. The purpose of Chapter 5 is to present analyses of the results of the exposure measures in order to demonstrate where within South Africa exposure is highest and lowest, and how patterns of exposure vary both within and between local municipality areas. With regards to the specific research questions, the aim in this chapter is to address Sub Q3: Where within South Africa is people's lived experience of inequality highest and lowest?

The primary analytical interest within this thesis is on exploring the levels of exposure to inequality experienced by the poor population of South Africa, and this is the explicit focus of this current chapter. However, I do also present a smaller set of analyses of the measures of exposure to inequality experienced by the non-poor in Appendix C.

In Chapter 4, the dataset used to operationalise the construction of the exposure measures was drawn from the Income and Material Deprivation Domain of the SAIMD 2001 at datazone level (Noble et al., 2009a) and the analyses in the first half of this chapter focus exclusively on this measures of exposure. However, as noted in Chapters 3 and 4, three other domains from the SAIMD 2001 at datazone level are also suitable as input datasets for the construction of exposure measures: Employment, Education, and Living Environment. In the second half of this chapter I therefore broaden the analyses to consider employment-based, education-based and living environment-based measures of exposure.

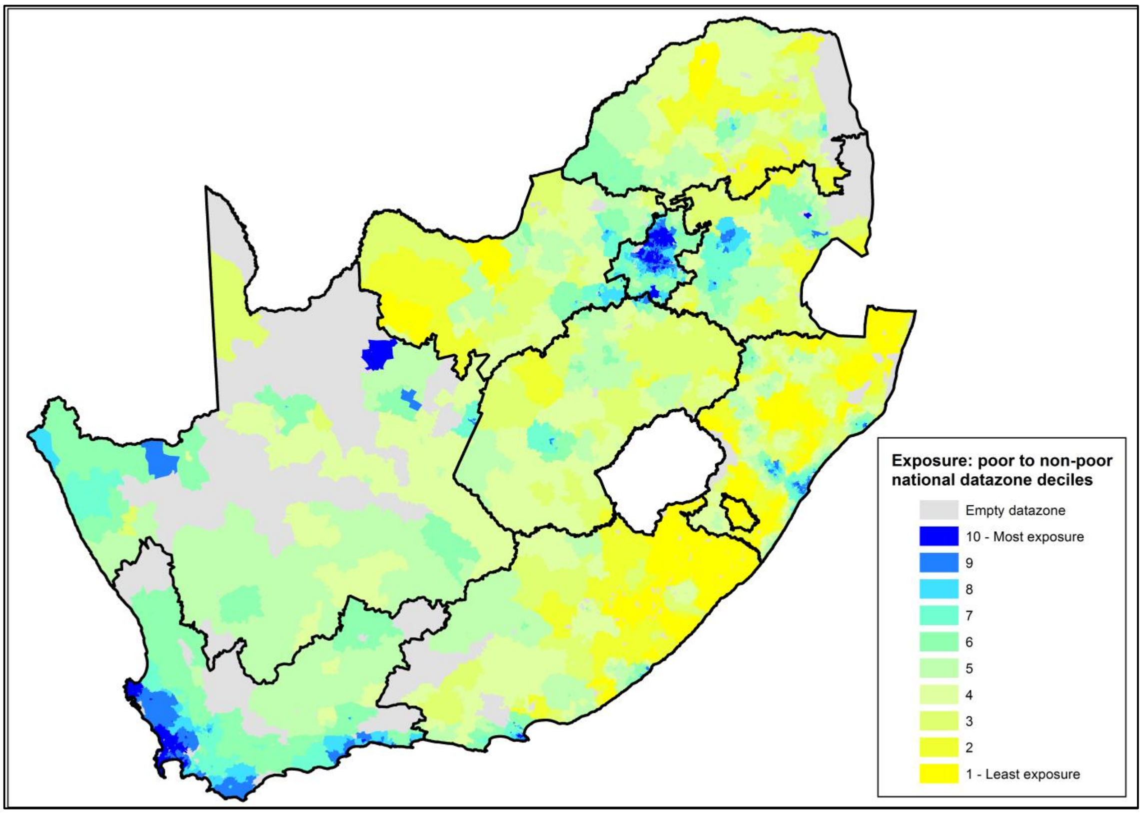
As will be shown below, the analyses at the beginning of this chapter reveal notable differences in the results between metropolitan areas and non-metropolitan areas. In the latter parts of this chapter, I therefore present some focused analyses looking only at the metropolitan areas.

5.2 Results: Spatial patterns of exposure to inequality across South Africa

I begin the analyses by mapping the datazone-level exposure scores based upon the INC domain across the whole of South Africa, before focusing on two contrasting municipalities to illustrate more clearly the geographical patterns of exposure at neighbourhood level. The aim here is to highlight how exposure levels can vary between neighbourhoods within municipalities as well as between municipalities. I then proceed to examine the differential distributions of exposure scores between metropolitan and non-metropolitan areas of the country, before focusing on the highest exposure and lowest exposure datazones and assessing their geographical location and the degree of spatial variation in exposure scores around these particular datazones. To reiterate, the focus of the analyses in this Chapter is on exposure to inequality from the perspective of the poor population (with analyses for the non-poor population provided in Appendix C).

In the analyses that follow, $aLDPxy_i$* exposure indices have been constructed using datazone level counts of 'poor' and 'non-poor' and these datazone level input data have been averaged over the bespoke spatial bounds around each separate datazone to produce exposure scores for every individual datazone in South Africa. Figure 5.1 shows the $aLDPxy_i$* scores for all datazones in South Africa. The datazones have been grouped into national deciles according to their scores on the $aLDPxy_i$* distribution. Those areas colour-coded bright yellow are the 10% of datazones with the *lowest* levels of exposure on this measure, whilst those areas colour-coded bright blue are the 10% of datazones with the *highest* level of exposure on this measure. The other eight colour-coded categories represent the other eight national deciles of datazones when ranked on this measure.

Figure 5.1: National Datazone deciles of $aLDPxy_i$* exposure measure

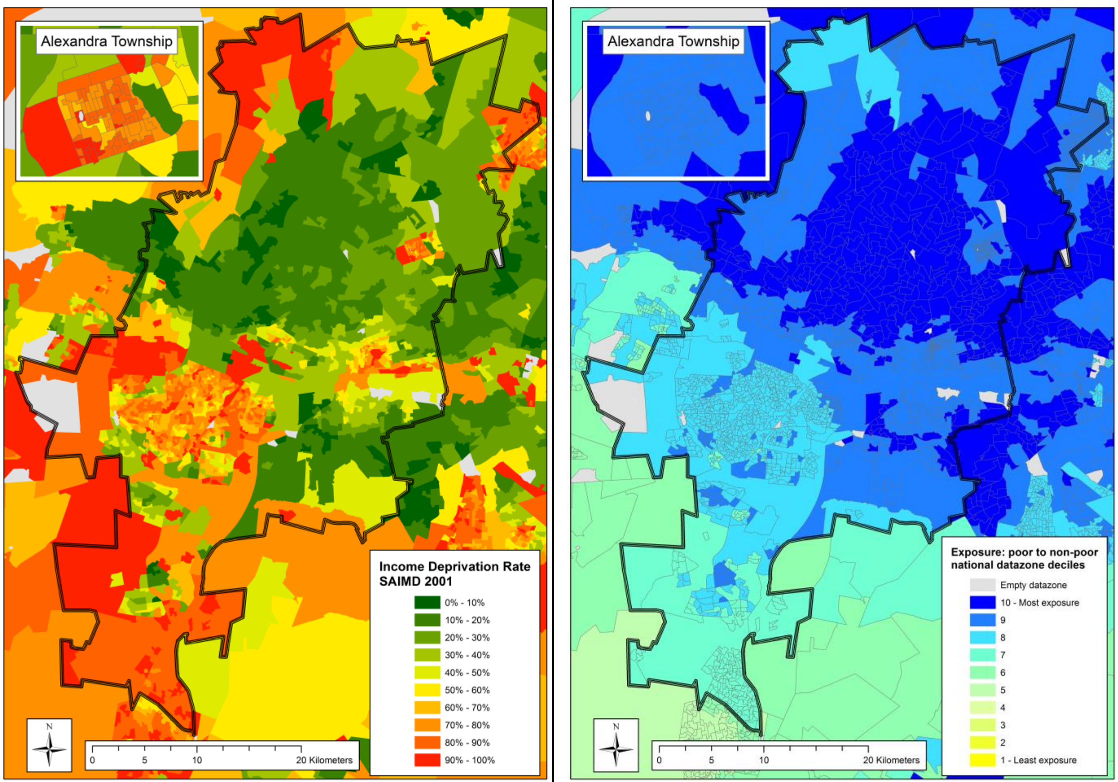


Although the fine-grained detail of datazones cannot be adequately represented on a map of the whole of South Africa, a broad pattern of exposure on this measure is apparent, with the areas of highest exposure typically located in and around the major cities, and the areas of lowest exposure typically located in the predominantly rural areas. In Chapter 3 the analyses showed that it was typically the major urban areas that contained the greatest concentrations of the *least deprived* neighbourhoods, while it was the more rural areas – and particularly the former homelands – that typically contained the greatest concentrations of the *most deprived* neighbourhoods. The initial map of datazone exposure scores shown in Figure 5.1 can therefore be seen to be roughly inverse to the map of datazone deprivation levels shown in Figure 3.6 in Chapter 3. This is as one would expect given the methodological basis of the exposure measure: a poor person living in an affluent (or rather, a less deprived) neighbourhood will be highly exposed to non-poor people (and indeed the visible signs of affluence and therefore the visible signs of inequality) as they go about their daily lives. In contrast, a poor person living in a former homeland may have very little contact with non-poor people (and very little exposure to the visible signs of affluence and therefore inequality) as they go about their daily lives due to the fact that almost everyone in their surrounding geographical area is classed as poor. Whilst the national map of datazone exposure scores shown here in Figure 5.1 provides an indication of this relationship between deprivation and exposure, it is necessary to examine the spatial patterns at a much more detailed level to fully appreciate the way exposure varies between neighbourhoods. I now therefore present data for two selected municipalities: the metropolitan municipality of Johannesburg in Gauteng, and the largely rural municipality of King Sabata Dalindyebo in the former homeland of Transkei in the Eastern Cape.

Figure 5.2 presents the results for the City of Johannesburg and surrounding areas. The left-side panel of Figure 5.2 again shows the datazone level poverty rate, with each colour-coded interval representing absolute rates of poverty.

Figure 5.2: Datazone level poverty rates and Exposure deciles for the City of Johannesburg Metropolitan Municipality in Gauteng

Panel A: deprivation rate	**Panel B: exposure score**



The right-side panel of Figure 5.2 shows the datazone scores on the $aLDPxy_i$* measure using national deciles. It is important to note, therefore, that the two maps in Figure 5.2 present data in slightly different ways: in the exposure map on the right of Figure 5.2 there are equal numbers of datazones across the country in each of the colour-coded intervals, but this is not the case in the poverty rate map on the left.

As is apparent from Figure 5.2, the poor individuals with the highest levels of exposure to socio-economic inequality tend to be those that live in – or close to – the relatively less poor neighbourhoods. As discussed above, this is as we would expect because these poor people are likely to interact more frequently with non-poor people on a daily basis and so experience higher exposure to the opposite end of the socio-economic distribution as they go about their daily lives. So, within Johannesburg, a poor person living in the largely affluent suburb of Sandton would experience very high exposure to inequality due to the very low poverty rates within Sandton and many of the surrounding areas. In panel B of Figure 5.2 we see that neighbourhoods across Sandton are colour-coded dark blue, meaning that these neighbourhoods fall within the highest exposure decile nationally. However, a poor resident of the highly deprived Alexandra township – which borders Sandton to the east and is shown as an expanded inset on the top left of each map – would also be highly exposed to inequality due to the close proximity of the township to the affluence of Sandton. The neighbourhoods of Alexandra can be seen to be mainly colour-coded mid-blue, meaning that they are in the second-highest decile of exposure nationally. Neighbourhoods in Soweto are mainly colour-coded light blue indicating they are mainly within the third-highest exposure decile nationally. A poor resident of Soweto is therefore likely to be relatively highly exposed to inequality when compared to poor residents over the rest of the country, but not as highly exposed as residents of Sandton or indeed Alexandra. This finding is partly due to the greater geographical distance between Soweto and Sandton than was observed between Alexandra and Sandton, as my methodology assumes that distance plays an important role in determining the likelihood of a person visiting each neighbourhood within the defined spatial bounds (see Chapter 4 for the detailed explanation of spatial bounds and spatial weights). Furthermore, Soweto consists of a more substantive spatial cluster of datazones than Alexandra which means that Soweto residents are assumed to carry out a greater proportion of their routine daily activities within the general confines of Soweto than is assumed to be the case for residents of Alexandra.

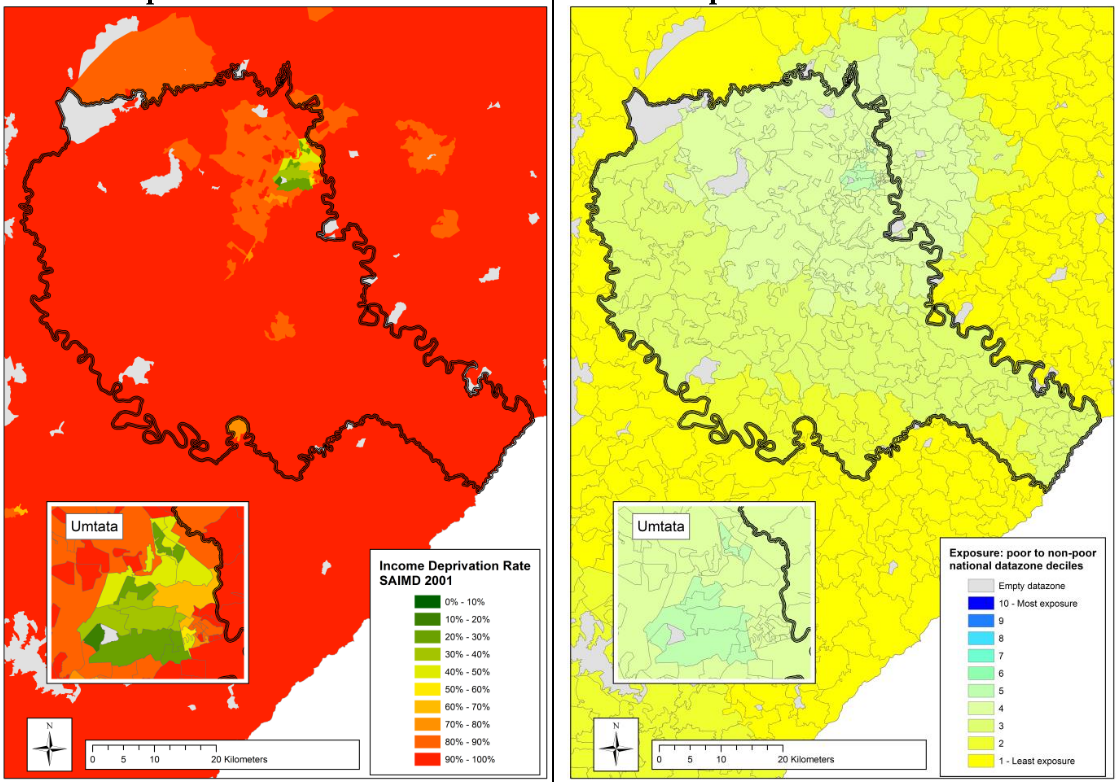
It is also evident that there appears to be more variation between neighbourhoods in Johannesburg in terms of their deprivation rates than in terms of their exposure scores. This is a deliberate feature of the methodology for constructing the exposure scores because, whilst deprivation rate relates only to the composition of the individual datazone neighbourhood, the exposure score takes into account the deprivation level in every datazone in the defined vicinity of the specified datazone and so reflects a weighted average. The exposure methodology essentially acknowledges that individuals may experience inequality in a range of different locations and not solely within the home neighbourhood (see Chapter 4 for a fuller discussion of the strengths of this 'local' approach to constructing segregation measures).

The main broad conclusion from Figure 5.2 is that in a municipality such as Johannesburg, which contains a large concentration of affluent neighbourhoods, the datazone exposure scores across the entire municipality (and indeed possibly in parts of the neighbouring municipalities too) are likely to be high compared to the rest of the country, but the exact level of exposure in any given datazone depends on the complex spatial configuration of datazone deprivation levels across the wider geographical area.

In contrast to the findings for datazones across Johannesburg, Figure 5.3 shows equivalent poverty rates and exposure decile information for datazones in and around the King Sabata Dalindyebo Local Municipality in the Eastern Cape. Whereas Johannesburg is characterised by some of the highest levels of exposure on the $aLDPxy_i*$ measure, the situation in and around King Sabata Dalindyebo municipality is very different. The majority of the datazones in this predominantly rural municipality are in the lower national deciles on this measure, indicating that poor individuals living in these datazones are much less exposed to socio-economic inequality then poor individuals in Johannesburg. It is evident that poor individuals living in and around the town of Umtata do have slightly higher rates of exposure than the rest of the municipality, and this is as one might expect as Umtata contains the only notable concentration of ‘non-poor’ population in that whole geographical area (as demonstrated by the left-side panel of Figure 5.3). The main broad conclusion from Figure 5.3 is therefore that poor people living in areas of extremely and almost uniformly high deprivation are deemed to have far less exposure to inequality than most other poor residents across South Africa, and that even those residents living within less deprived towns (such as Umtata) have lower exposure than residents of places such as Johannesburg because they are essentially surrounded by a sea of extreme deprivation.

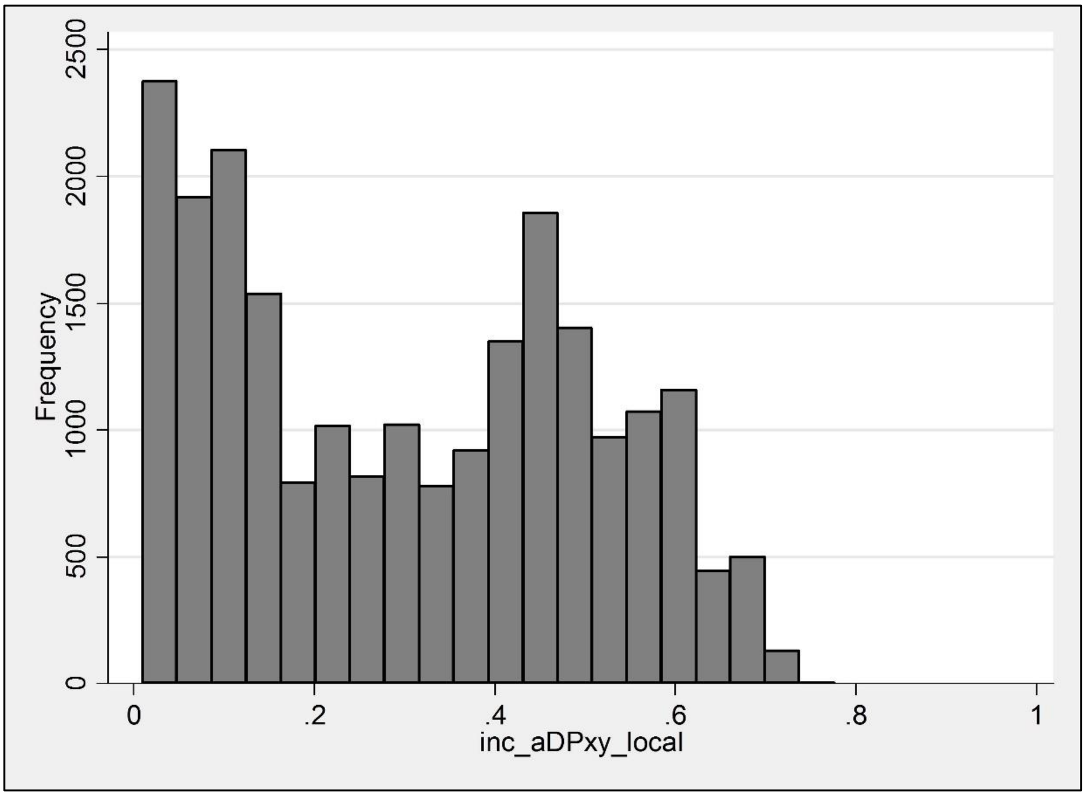
Figure 5.3: Datazone level poverty rates and Exposure deciles for the King Sabata Dalindyebo Local Municipality in the Eastern Cape

Panel A: deprivation rate | **Panel B: exposure score**



Whilst the data on exposure presented in Figure 5.1, Figure 5.2 and Figure 5.3 reflect the *national decile* distribution of $aLDPxy_i*$, it is also possible to examine the actual underlying $aLDPxy_i*$ scores that formed the basis of the decile ranking. Figure 5.4 presents a histogram of $aLDPxy_i*$ scores at datazone level for the whole of South Africa. The horizontal x-axis represents the $aLDPxy_i*$ score, with a potential range of 0 (meaning no exposure) to 1 (meaning complete exposure). The verticle y-axis represents the number of datazones that fall within each of the column bands.

Figure 5.4: Histogram of Datazone *aLDPxy$_i$ scores for all Datazones in South Africa**



The distribution of *aLDPxy$_i$** values shown in Figure 5.4 is bi-modal, with a concentration of datazones exhibiting relatively low levels of exposure and another concentration of datazones exhibiting relatively high levels of exposure. To help unpick this further, Figure 5.5 presents equivalent histograms for two discrete subsets of the datazones: the left-side panel of Figure 5.5 shows the distribution of *aLDPxy$_i$** values for datazones in the six metropolitan municipalities as per the definition at the time of the 2001 Census (namely, Cape Town, Johannesburg, Tshwane, Ekurhuleni, eThekwini and Nelson Mandela Metropolitan Municipalities), while the right-side panel of Figure 5.5 shows the distribution of *aLDPxy$_i$** values for datazones in the 229 non-metropolitan municipalities as per the definition at the time of the 2001 Census.

Figure 5.5: Histogram of Datazone *aLDPxy$_i$ scores, separately by metropolitan and non-metropolitan areas**

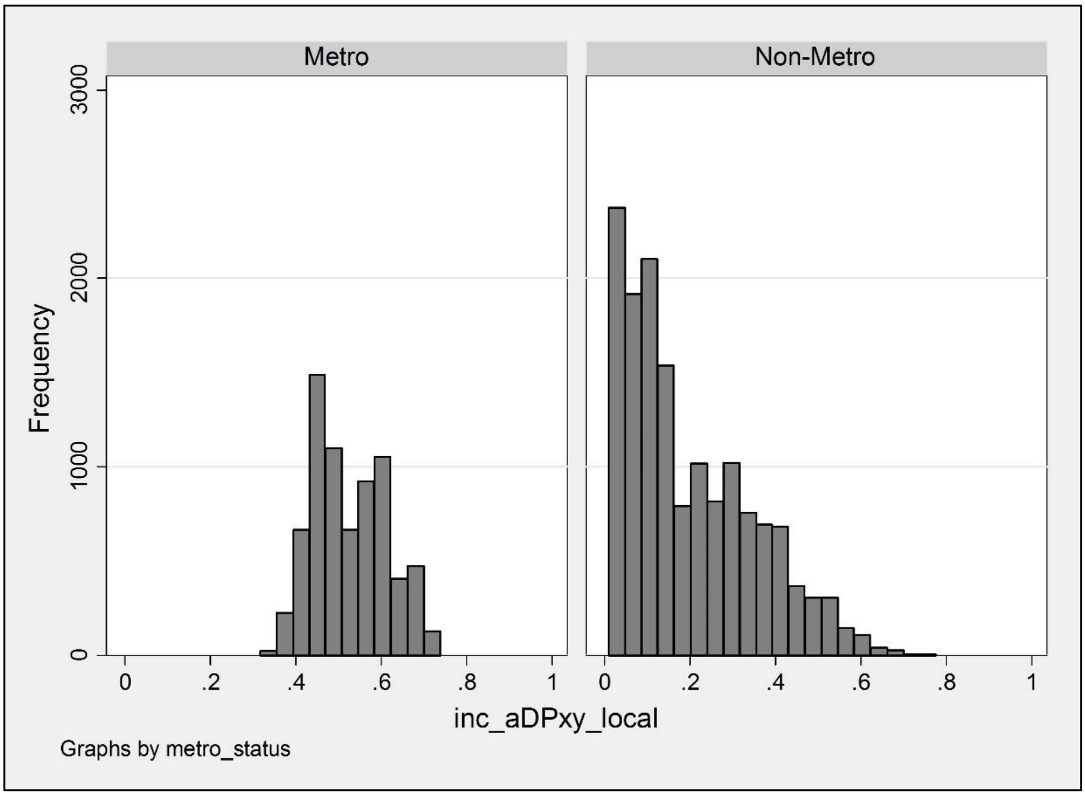


Figure 5.5 clearly shows the contrasting distributions of values on the *aLDPxy$_i$** measure between datazones in metropolitan and non-metropolitan areas. Almost all (95.5%) of the metropolitan datazones exhibit *aLDPxy$_i$** scores above 0.4 on the x-axis, compared to just 12.2% of the non-metropolitan datazones. Whilst there is therefore some overlap in exposure scores between metropolitan and non-metropolitan datazones, this overlap is quite small and the two distributions are considerably different.

I contend that the differential distribution of exposure scores between metropolitan and non-metropolitan areas is an important finding and I structure later analysis within this chapter in the context of this differential. However, before moving on to those particular analyses, I first examine the locations of the highest exposure datazones before supplementing this with a brief analysis of the locations of the lowest exposure datazones.

Returning to consider the national decile distribution of $aLDPxy_i$* shown in Figure 5.1, Figure 5.2 and Figure 5.3, Table 5.1 summarises the geographical locations of the 10% of datazones nationally that have the highest scores on the $aLDPxy_i$* measure.

Table 5.1: Location of Datazones in the 10% highest exposure decile nationally on the $aLDPxy_i$* measure

Municipality	**Number**	**Percentage**
City of Cape Town	986	44.5%
City of Tshwane Metro	502	22.7%
City of Johannesburg Metro	290	13.1%
Ekurhuleni Metro	206	9.3%
Others (23 municipalities)	232	10.5%
Total in the 10% highest exposure decile nationally	2,216	100.0%

These highest exposure datazones are located primarily within the four metropolitan municipalities of City of Cape Town, City of Tshwane, City of Johannesburg and Ekurhuleni. These four metropolitan municipalities account for 89.5% of the highest exposure decile datazones on this measure, with the remaining 10.5% of datazones being spread more thinly over 23 other municipalities (including eThekwini and Nelson Mandela Bay metropolitan municipalities).

The major metropolitan municipalities do, of course, contain higher total numbers of datazones within their boundaries than non-metropolitan municipalities, and so Table 5.2 is presented to show the ten municipalities across South Africa that have the highest proportions of datazones falling within the highest exposure decile on the $aLDPxy_i$* measure.

Table 5.2: The ten municipalities with the highest proportions of datazones in the 10% highest exposure decile nationally on the $aLDPxy_i$* measure

Municipality	Number of datazones in the municipality	Number of datazones in highest exposure decile	Percentage of municipality datazones in the highest exposure decile
Gamagara (Northern Cape)	9	7	77.8%
Stellenbosch (Western Cape)	60	44	73.3%
City of Cape Town (Western Cape)	1,388	986	71.0%
Saldanha Bay (Western Cape)	34	23	67.6%
City of Tshwane Gauteng/North West)	951	502	52.8%
Mossel Bay Western Cape)	37	10	27.0%
City of Johannesburg (Gauteng)	1,599	290	18.1%
George (Western Cape)	67	12	17.9%
Ekurhuleni (Gauteng)	1,188	206	17.3%
Nokeng tsa Taemane (Gauteng)	21	3	14.3%

City of Cape Town, City of Tshwane, City of Johannesburg and Ekurhuleni all rank within the top ten municipalities on this measure and so are reflected in Table 5.2, but there are also smaller municipalities such as Gamagara and Stellenbosch with very high proportions of datazones within the highest exposure decile nationally.

Having established that the highest exposure datazones (i.e. those in the top decile) are not evenly distributed across South Africa but are, in fact, concentrated in certain geographical areas, I will now proceed to examine the wider range of datazone level exposure scores for different spatial groupings. I commence by looking at differential distributions of datazone exposure scores within and between provinces, before focusing on the equivalent distributions within a selection of local municipalities.

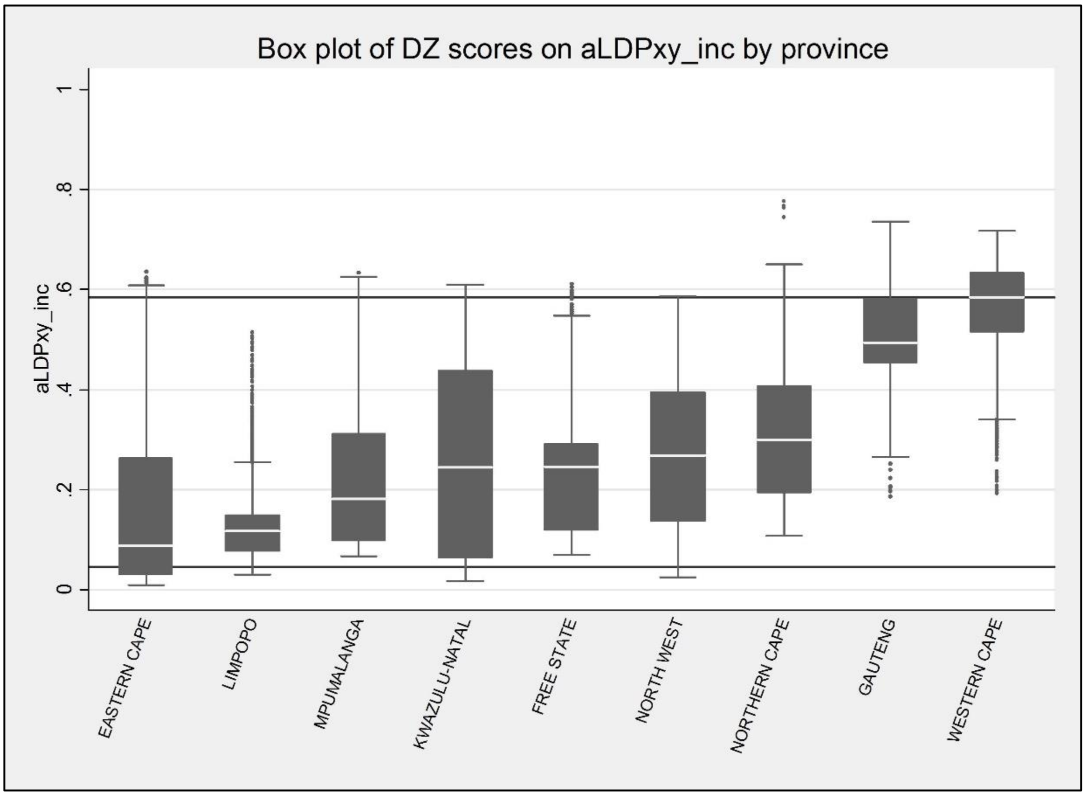
Box plot charts provide a means to assess the distribution of $aLDPxy_i*$ indicator values between different subsets of the datazones. In Figure 5.6 a box plot is used to show the spread of datazone values on the $aLDPxy_i*$ measure for each of the nine provinces in South Africa. The vertical y-axis shows the actual $aLDPxy_i*$ score for the datazones. Each province is represented by a grey rectangular box which represents the Interquartile Range of datazone $aLDPxy_i*$ scores, plus upper and lower whiskers that represent the upper and lower ranges of the datazone scores, with small dots used to represent statistical outliers. The median datazone score for each province is represented by a small white horizontal line running through the grey Interquartile Range box. The upper horizontal reference line[30] represents the cut point for the highest exposure national decile of datazones, with all datazones scoring *above* this line being colour-coded dark blue in Figure 5.1, Figure 5.2 and Figure 5.3, while the lower horizontal reference line[31] represents the cut point for the lowest exposure national decile of datazones, with all datazones scoring *below* this line being colour-coded bright yellow. The provinces are ordered from left to right according to the median datazone $aLDPxy_i*$ score for the province.

It is evident from Figure 5.6 that the Eastern Cape Province has the lowest median exposure value and the Western Cape Province has the highest median exposure value on this measure. The Interquartile Range for the Eastern Cape overlaps with the

[30] y= 0.585
[31] y= 0.045

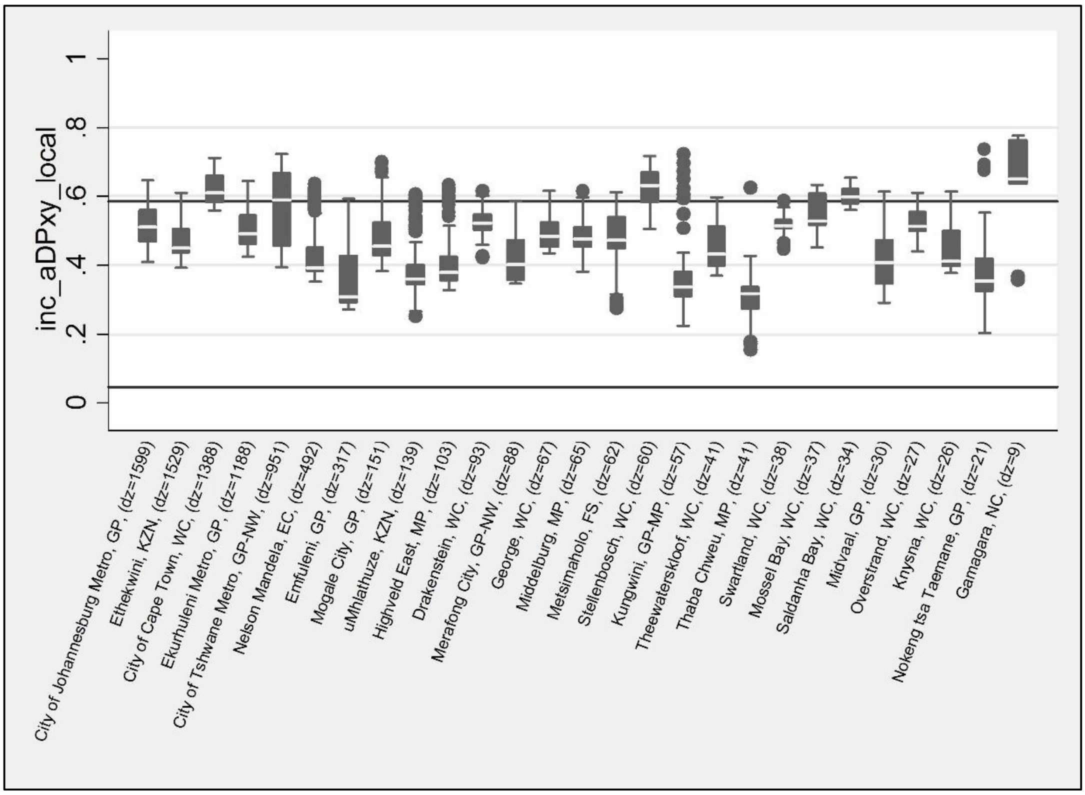
lower horizontal reference line, meaning that more than a quarter of the datazones in the Eastern Cape have exposure values that place them within the 10% of datazones nationally with the lowest exposure. However, even within the Eastern Cape, we see that there are some datazones that have exposure scores that lie above the upper horizontal reference line, indicating that these datazones have exposure scores within the 10% of datazones nationally with the *highest* level of exposure. Overall, however, it is clear that the Eastern Cape as a whole is characterised predominately by areas of low exposure on the $aLDPxy_i*$ measure. This is also the case for the other provinces towards the left-hand side of the chart, such as Limpopo. On the right-hand side of the chart it is evident that Gauteng and the Western Cape exhibit rather different distributions than the other seven provinces, with median values and Interquartile Ranges that sit notably higher up the y-axis than the other provinces. However, even in these two provinces there are datazones where exposure is substantially lower than the majority of the province. Indeed, an important overall finding from these analyses is that although there are differences *between* provinces in the distribution of datazones on the $aLDPxy_i*$ measure, there are also wide variations *within* provinces.

Figure 5.6: Boxplot of Datazone $aLDPxy_i$* scores by province



The boxplot approach to analysing distributional characteristics of subsets of datazones can also be applied to focus on datazones in selected municipalities. The analyses of Table 5.1 above showed that the 10% highest exposure decile of datazones on the $aLDPxy_i$* measure were spread across 27 municipalities (with the majority being located in the four metropolitan municipalities of Cape Town, Johannesburg, Tshwane and Ekurhuleni). In Figure 5.7, the spread of datazone values on $aLDPxy_i$* is shown for each of the 27 municipalities separately. In this chart, the municipalities are ordered from left to right along the x-axis according to the total number (in descending order) of datazones within the municipality so that more populous metropolitan municipalities can be easily identified. The total number of datazones per municipality is shown after the municipality name in the labels along the x-axis.

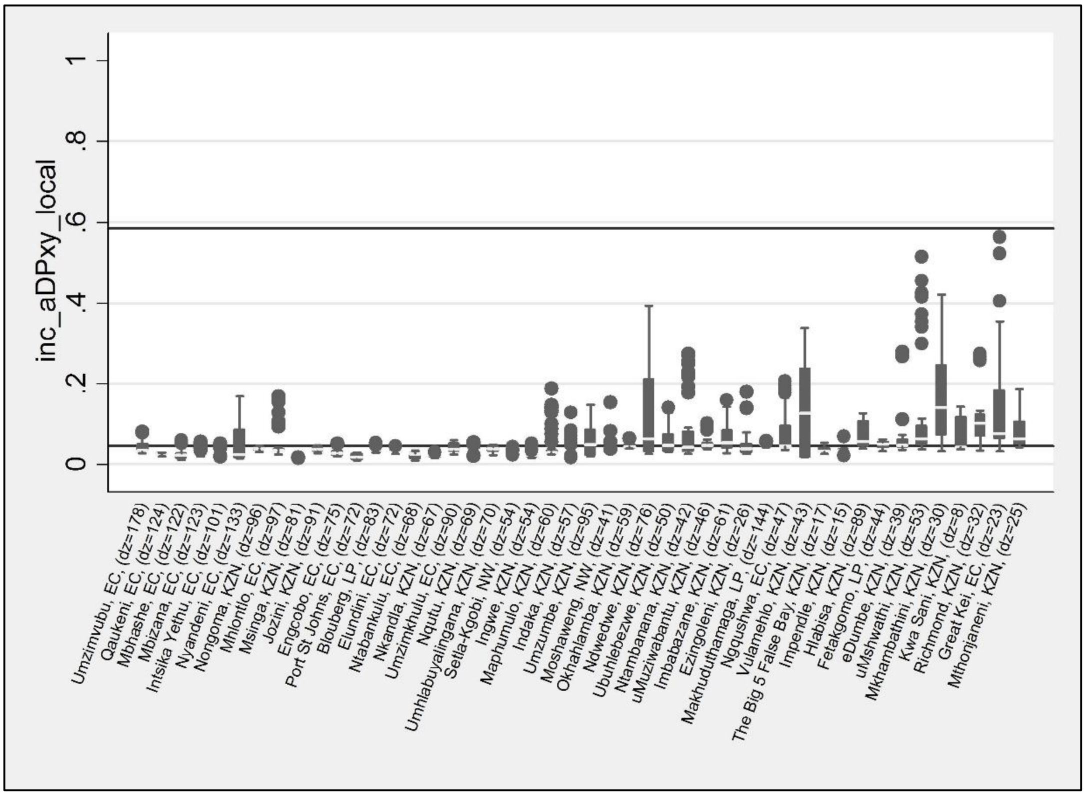
Figure 5.7: Boxplot of Datazone *aLDPxy*$_i$* scores by municipality, for those municipalities containing at least one Datazone in the highest exposure national decile



As is noted above, the criterion for including a municipality in Figure 5.7 was that at least one datazone within the municipality must lie within the *highest exposure* national decile, and this is clearly reflected in the chart, where all municipalities contain at least one area above the upper horizontal reference line. In most cases, the median datazone value for the municipalities shown in Figure 5.7 lies below the upper horizontal reference line, but there are five municipalities (City of Cape Town, City of Tshwane Metro, Gamagara, Saldanha Bay and Stellenbosch) where the median lies above the line, indicating as in Table 5.2 that half or more of the datazones in each of these five municipalities are in the highest exposure decile nationally. In contrast, in other municipalities, only a small minority of the datazones (sometimes only a single datazone) falls in the highest exposure decile. In general, Figure 5.7 shows that there is heterogeneity both within municipalities and between municipalities.

Figure 5.8 focuses on the other end of the *aLDPxy$_i$** spectrum, and includes only those municipalities where one or more datazones falls in the *lowest exposure* national decile. In this chart the municipalities are sorted in ascending order from left to right according to the median datazone score on this measure.

Figure 5.8: Boxplot of Datazone *aLDPxy$_i$ scores by municipality, for those municipalities containing at least one Datazone in the lowest exposure national decile**



Forty six municipalities are represented in Figure 5.8. There are 30 municipalities with median scores below the lower horizontal reference line, indicating that in each of these 30 municipalities over half of the datazones are in the lowest exposure national decile. Furthermore, in six municipalities (Msinga, Nkandla, Ntabankulu, Port St Johns, Qaukeni and Setla-Kgobi) every constituent datazone falls within the lowest exposure national decile. Almost all of these 30 municipalities are located in whole or in part in former homeland areas.

To summarise, the main finding from the analyses here is that poor members of the South African population who live within or close to substantial geographical areas of affluence (or, more accurately, areas of low deprivation) have the highest levels of exposure to inequality because these poor people routinely come into contact with non-poor people and routinely recognise the stark visual signs of affluence and inequality as they go about their daily lives. Those poor individuals who live in areas where almost everyone else is also poor have much lower levels of exposure. The results reveal notable variations in exposure levels both between and within provinces and municipalities, with exposure typically lowest in the mainly rural municipalities within the former homelands, and typically highest in the more urbanised municipalities and particularly the major metropolitan areas.

The analyses of exposure presented so far have all been based upon the measures constructed from the Income and Material Deprivation Domain of the SAIMD 2001. However, as noted in the introduction to this Chapter, I have also generated equivalent measures using the other three suitable domains of the SAIMD 2001: Employment, Education, and Living Environment. Like the Income and Material Deprivation Domain, these other three domains represent the proportion of population deprived on the particular dimension of deprivation and, as such, they are well suited as input datasets for residential segregation measures.

The methodology described above for developing a measure of exposure to inequality based upon the dimension of Income and Material Deprivation was replicated to produce equivalent exposure measures for: Employment; Education; and Living

Environment. For ease of reference, these three new exposure measures are called: *Expos_Emp*, *Expos_Edu* and *Expos_Liv*.

In light of the notable differential between metropolitan and non-metropolitan areas in the exposure measures, I focus explicitly on the metropolitan areas during the following analysis of *Expos_Emp*, *Expos_Edu* and *Expos_Liv* and their comparison to *Expos_Inc*. This permits a more detailed examination of intra-metropolitan exposure levels.

5.3 Results: Spatial patterns of exposure across metropolitan areas

I begin the intra-metropolitan analysis of the four separate deprivation domain-based exposure measures by showing how the datazone exposure scores relate to the respective datazone *deprivation rates* across the metropolitan municipalities. This highlights how the distributions on the exposure score measure and the deprivation rate measure differ systematically due to the former being a weighted average of data derived from the latter (see Chapter 4 for a detailed methodological discussion of the construction process). These analyses also reveal certain differences in exposure scores between the municipalities which I proceed to explore further in this chapter. I then test the extent to which the datazone scores on the four exposure measures correlate and, based upon the results of these correlations, I proceed to develop a composite exposure score drawing upon all four domain-specific exposure measures. In the final piece of empirical analysis in this chapter I utilise my composite exposure index to examine variations between datazones within and between the metropolitan municipalities and look in detail at the sub-municipality pattern of exposure within the City of Cape Town.

The relationships between deprivation rates and exposure scores for metropolitan datazones[32] are shown in Figure 5.9, Figure 5.10, Figure 5.11 and Figure 5.12. In these scatterplots the datazone values on the respective deprivation domain score (i.e. percent of population deprived) are shown along the x-axes and the respective exposure score along the y-axes. The points are colour-coded according to the relevant municipality to highlight differentials in datazone scores both between and within the metropolitan municipalities.

Although the x-axis and the y-axis on each of the four scatterplots share the same *theoretical* minimum (0) and maximum (1), in practice, we see a narrower spread of values on the exposure y-axis then on the deprivation rate x-axis, and we see a greater degree of clustering within municipalities in terms of scores on the exposure measure. This is, of course, to be expected because the y-axis exposure score is a form of weighted average of all datazone deprivation scores within the respective bespoke spatial bounds, while the x-axis deprivation score relates only to the particular datazone in question and therefore is independent of the other datazone deprivation scores (although, in reality, we know that a range of historical and contemporary social, economic and political factors tend to result in clustering of similarly deprived neighbourhoods). The scatterplots indicate that a person's exposure to inequality is affected by the municipality in which they live, because each municipality has its own unique neighbourhood deprivation profile which feeds into the derivation of exposure scores.

[32] In this part of the analysis I am including the two municipalities which were categorised as metropolitan municipalities subsequent to the 2001 census (Mangaung in the Free State and Buffalo City in the Eastern Cape) as they contribute to a more nuanced discussion.

Figure 5.9: Datazone deprivation rate against exposure score - INC

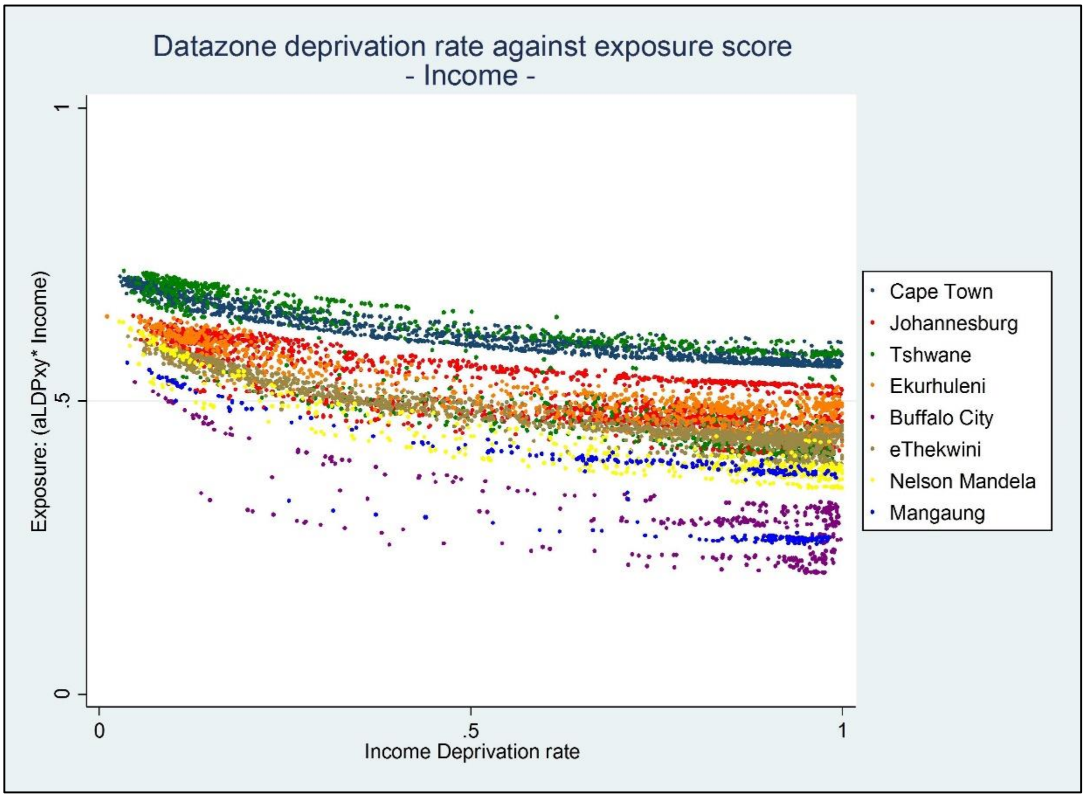


Figure 5.10: Datazone deprivation rate against exposure score - EMP

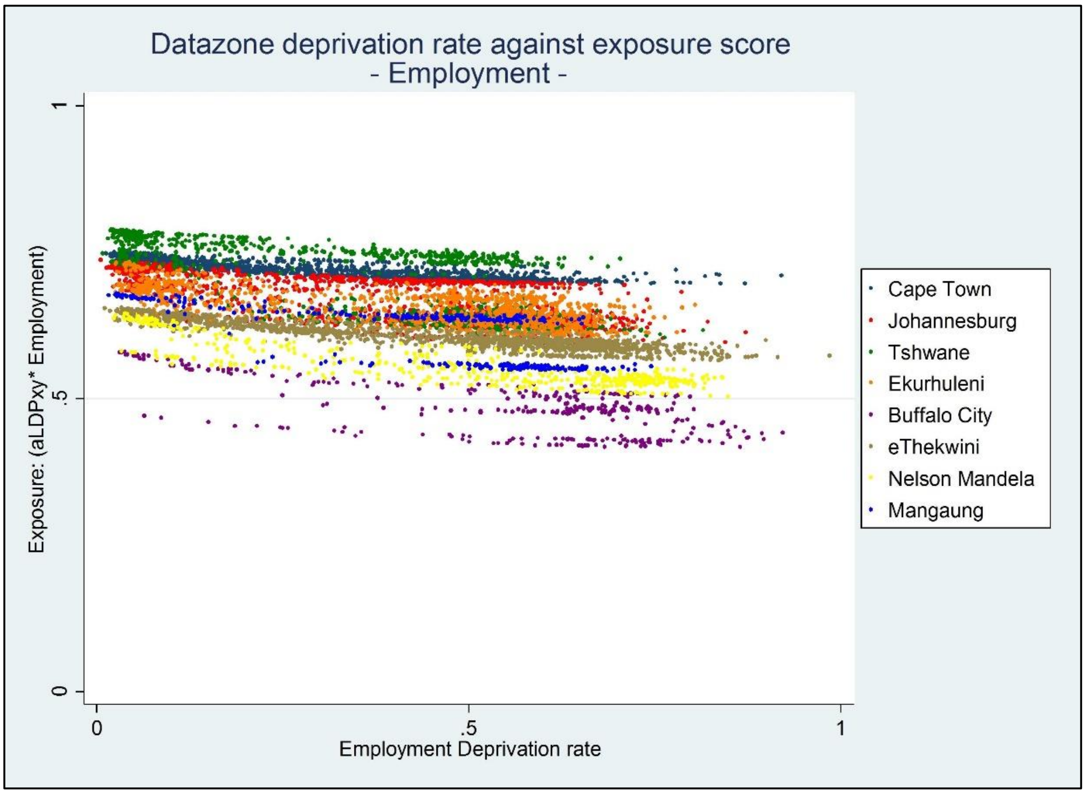


Figure 5.11: Datazone deprivation rate against exposure score - EDU

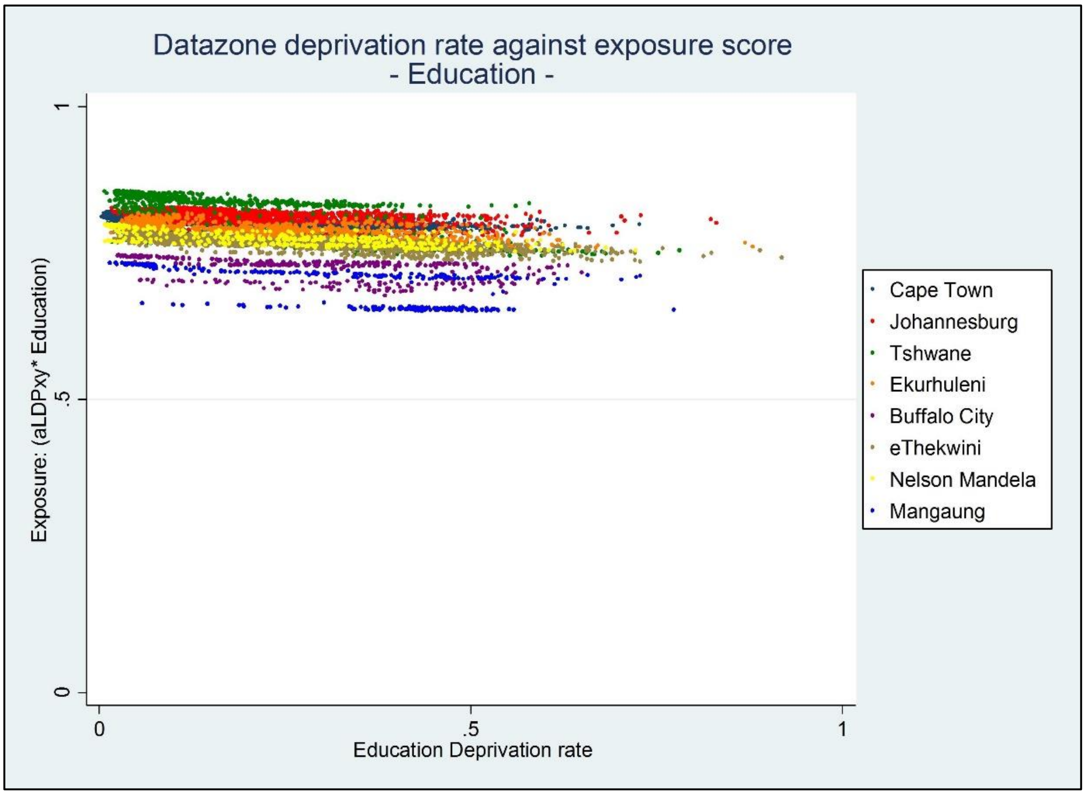
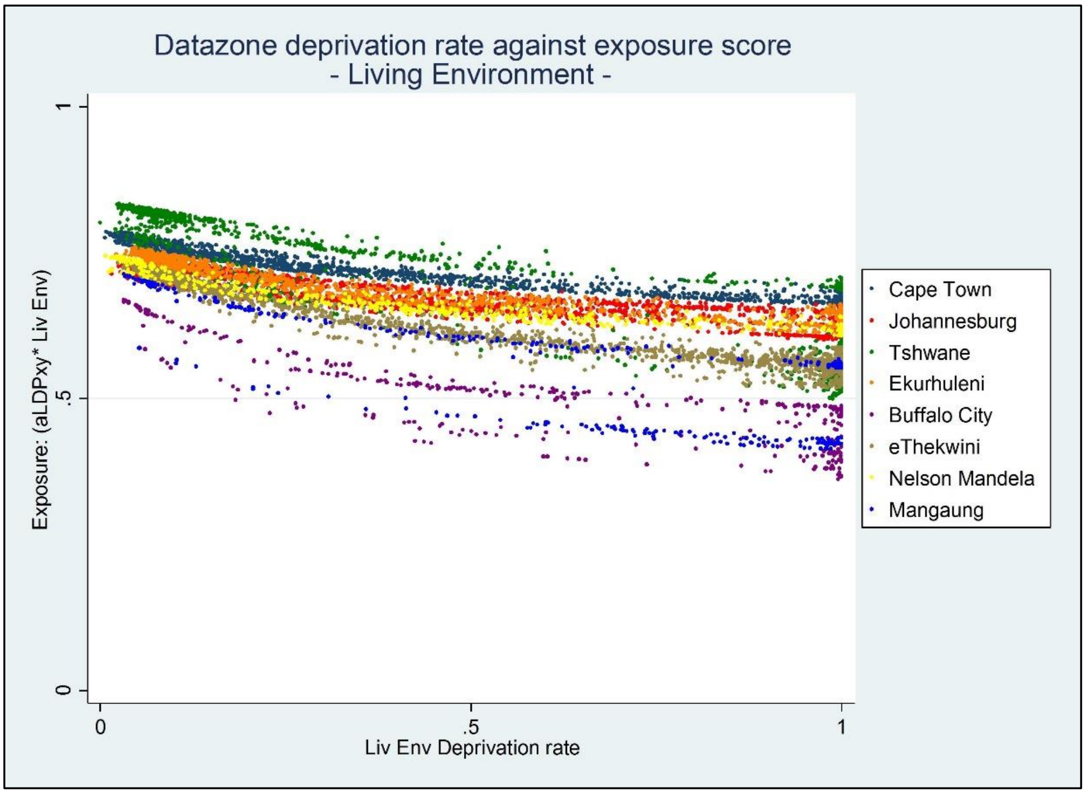


Figure 5.12: Datazone deprivation rate against exposure score - LIV



The green dots (representing Tshwane datazones) and the navy blue dots (representing Cape Town datazones) can be seen to cluster towards the top of the charts, indicating that exposure to inequality is typically high across these two municipalities. An interesting pattern can be observed within Mangaung, whereby there appear to be two notably different clusters of datazones on the exposure measure (denoted by two groups of the lighter-blue dots). This pattern is due to the spatial configuration of deprivation within Mangaung, where there is a concentration of low deprivation datazones in the north-western segment of the Bloemfontein urban area, which can be contrasted with notably higher deprivation rates in the south-eastern segment of Bloemfontein urban area, as well as the township of Botshabelo and much of the surrounding rural area.

Whilst the scatterplots provide a valuable visual depiction of the relationships between deprivation rates and exposure scores, a clearer picture of the within- and between-municipality distributions of datazone exposure scores can be provided through box plots. Figure 5.13 shows the distribution of datazone level exposure scores per metropolitan municipality on each of the four dimensions assessed. The y-axis represents the respective exposure score (ranging from a hypothetical minimum value of zero to a hypothetical maximum value of one) and the municipalities are presented along the x-axis, ordered from left to right in descending order according to the median datazone exposure score within the municipality. These box and whisker plots again show the interquartile range as well as the full extent of the values per municipality.

In terms of exposure to inequality on the income dimension, Cape Town can be seen to rank highest according to the median datazone exposure score, followed by Tshwane and Johannesburg. Although the median datazone values on this measure are similar for Cape Town and Tshwane, these two municipalities exhibit notably different distributions, with Cape Town seen to have a relatively tight spread compared to Tshwane's broader spread. This indicates much more variation in exposure levels on this measure between the constituent datazones in Tshwane than between the constituent datazones in Cape Town. All of Cape Town's datazones have exposure scores on this measure of above 0.55 whereas all of Buffalo City's datazones have exposure scores on this measure of below 0.55.

In terms of exposure to inequality based upon the employment deprivation dimension, we again see Tshwane and Cape Town ranked highest and Buffalo City ranked lowest on the median datazone value. Tshwane exhibits a broader spread of datazone values than Cape Town indicating greater variation within Tshwane than within Cape Town.

Cape Town is ranked third highest on the median datazone score according to the exposure to education inequality, with Tshwane and Johannesburg ranked above it. Note here that the y-axis for the education-based measure has a smaller range than for the other three dimensions, indicating that there is less overall variation in education levels between metropolitan datazones on this measure than on the other three measures.

Finally, in terms of exposure to inequality based upon the living environment dimension, we see Tshwane and Cape Town again ranked highest on the median datazone score and Buffalo City ranked lowest. Once more, Tshwane can be seen to exhibit a broader range of datazone exposure scores than Cape Town, indicating greater internal variation within Tshwane than Cape Town.

The results presented in Figure 5.13 below highlight a number of commonalities between the distributions on the four datazone-level exposure scores (income, employment, education and living environment). Looking across all 7,800 metropolitan datazones it is possible to assess the correlations between these four measures, and this is presented in Table 5.3.

Figure 5.13: Exposure scores across metropolitan municipalities

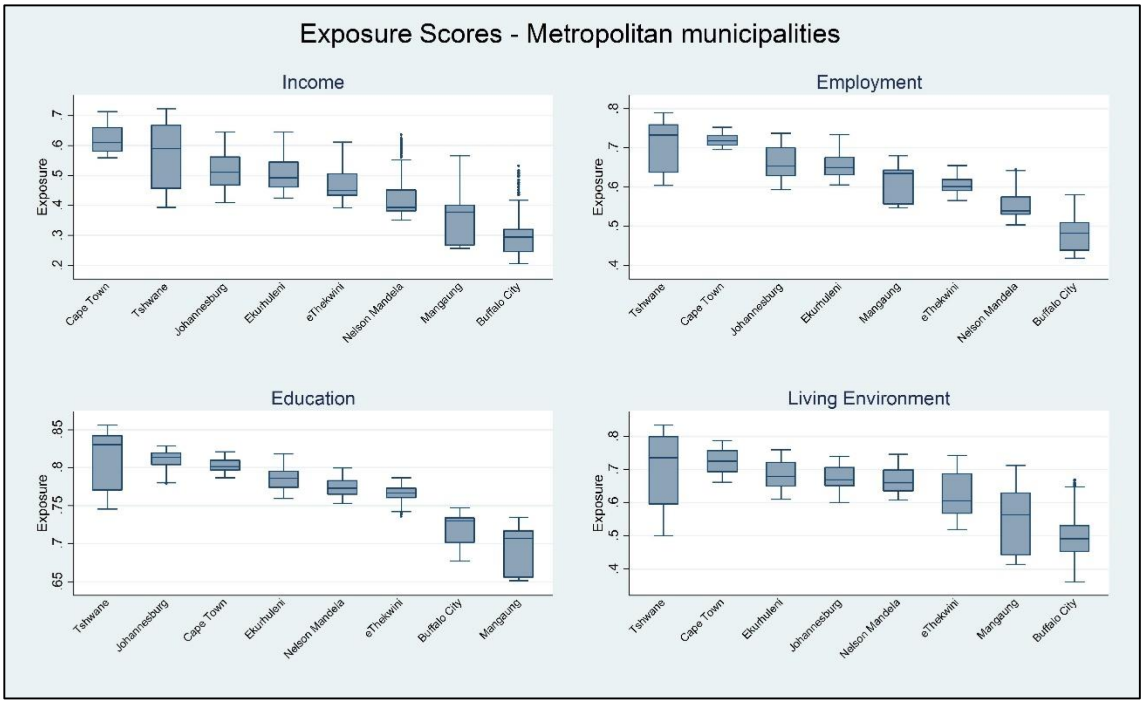


Table 5.3: Spearman rank correlation coefficients between the four dimension-specific exposure measures, all metropolitan Datazones

All Metros	*Expos_Inc*	*Expos_Emp*	*Expos_Edu*	*Expos_Liv*
Expos_Inc	1			
Expos_Emp	0.9171	1		
Expos_Edu	0.8104	0.8281	1	
Expos_Liv	0.8947	0.7821	0.7225	1

We see here from Table 5.3 that strong correlations exist between exposure score on all four dimensions. The strongest correlation is between the income exposure score and the employment exposure score (0.9171) whilst the weakest correlation is between the education exposure score and the living environment exposure score (0.7225). Equivalent correlations can also be undertaken focusing only on the datazones within the municipality of Cape Town, and these coefficients are presented in Table 5.4. Here we see that each of the coefficients is higher in the Cape Town-specific table than in the table above that included all metropolitan datazones across the eight metropolitan municipalities.

Table 5.4: Spearman rank correlation coefficients between the four dimension-specific exposure measures, Cape Town datazones only

Just Cape Town	*Expos_Inc*	*Expos_Emp*	*Expos_Edu*	*Expos_Liv*
Expos_Inc	1			
Expos_Emp	0.9344	1		
Expos_Edu	0.8752	0.8861	1	
Expos_Liv	0.9339	0.8592	0.8116	1

Conceptually, it might be preferable to have one single exposure measure for the purpose of mapping and analysing spatial variations across South Africa because, in reality, each of the four separate measures is essentially constructed as a proxy for an underlying measure of the ‘lived experience of inequality’. The analyses presented above suggest that combining the four separate exposure indices into a single composite measure does appear to be an empirically valid action to take for the purpose of mapping and spatial analysis[33].

[33] Whilst thematic mapping and spatial analysis of exposure levels is a suitable application of a composite exposure index, other analytical purposes require the separate domain-specific exposure scores. Indeed, in the modelling of statistical associations between lived experience of inequality and attitudes to inequality

To operationalise a composite measure I combined the four separate exposure measures using weights derived through maximum likelihood factor analysis. Each of the four separate datazone level exposure scores was first ranked within the group of metropolitan datazones, and then each of the rank distributions was normalised prior to incorporation in the factor analysis routine. I employed maximum likelihood factor analysis and the resulting factor loadings are shown in Table 5.5.

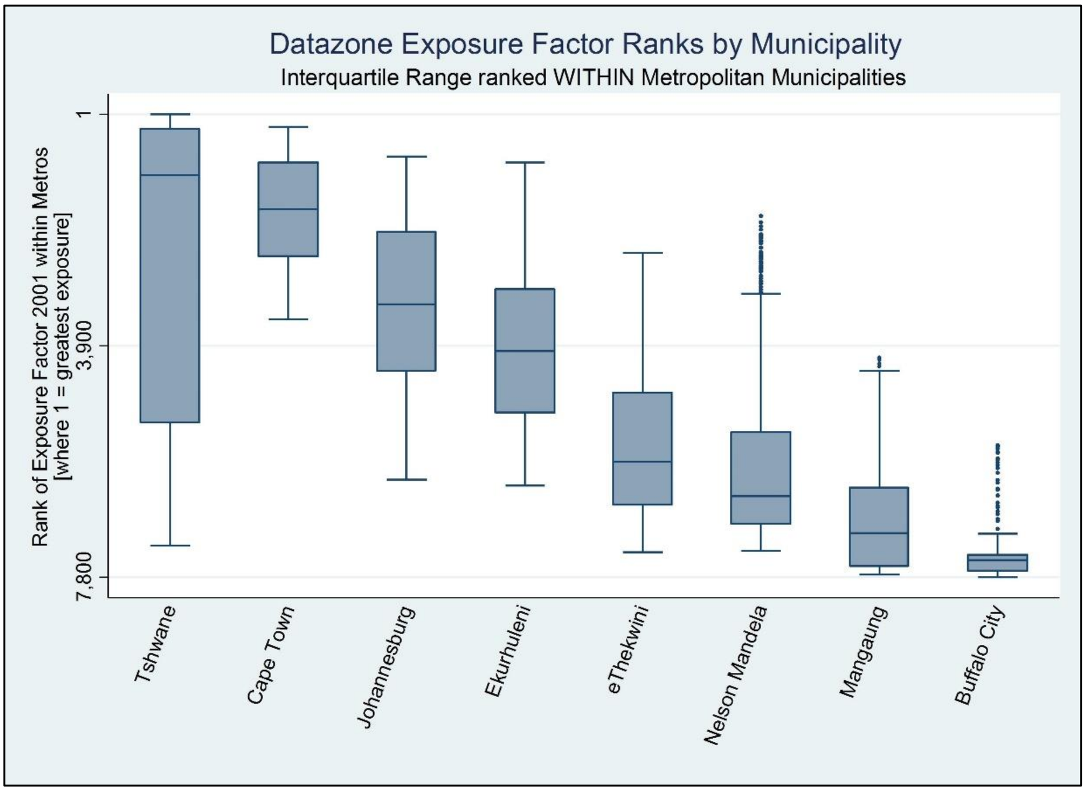
Table 5.5: Factor loadings from construction of Expos_Fac

Variable	**Factor1**	**Uniqueness**	**Standardised factor loading**
inc	0.990	0.021	0.270
emp	0.930	0.136	0.254
edu	0.836	0.301	0.228
liv	0.911	0.171	0.248

The factor analysis generated a single factor (with an Eigenvalue of 3.372) and the new composite exposure measure was called '*ExposFac*'. The 7,800 metropolitan datazones were ranked from highest score on my new *ExposFac* measure. The overarching results by municipality are presented in Figure 5.14, where rank 1 on the y-axis relates to the datazone with the highest overall composite exposure score on the *ExposFac* measure. The municipalities are again ranked from left to right according to the median datazone score on this measure.

in Chapter 7 of this thesis I opt to use a single domain-specific exposure measure rather than the composite measure.

Figure 5.14: Datazone Exposure Factor ranks by Municipality (ranked across metros only)



As one might expect, the broad patterns identified in the four separate exposure distributions presented in Figure 5.13 above carry through to the new composite *ExposFac* measure. Primarily, we see that Tshwane and Cape Town exhibit the highest median datazone values and Buffalo City exhibits the lowest median datazone value. The spread of values with Tshwane is much broader than the spread in Cape Town, indicating that there is more homogeneity in datazone exposure scores with Cape Town than within Tshwane. All of Cape Town's datazones are in the 'top half' of the metropolitan rank distribution, whereas certain datazones in Tshwane exhibit some of the lowest levels of exposure to inequality across all metropolitan areas.

A more comprehensive picture can be seen from Table 5.6 which shows the number of datazones per metropolitan municipality that fall within each within-metro decile on the *ExposFac* distribution. The highest exposure decile of datazones are concentrated in Cape Town (328 datazones) and Tshwane (448 datazones), plus a very small number in Johannesburg (just four datazones). At the other end of the spectrum, almost all the datazones in Buffalo City are within the lowest exposure decile of datazones.

Table 5.6: Number of Datazones per metropolitan municipality that fall within each within-metro decile on the composite Exposure Factor score

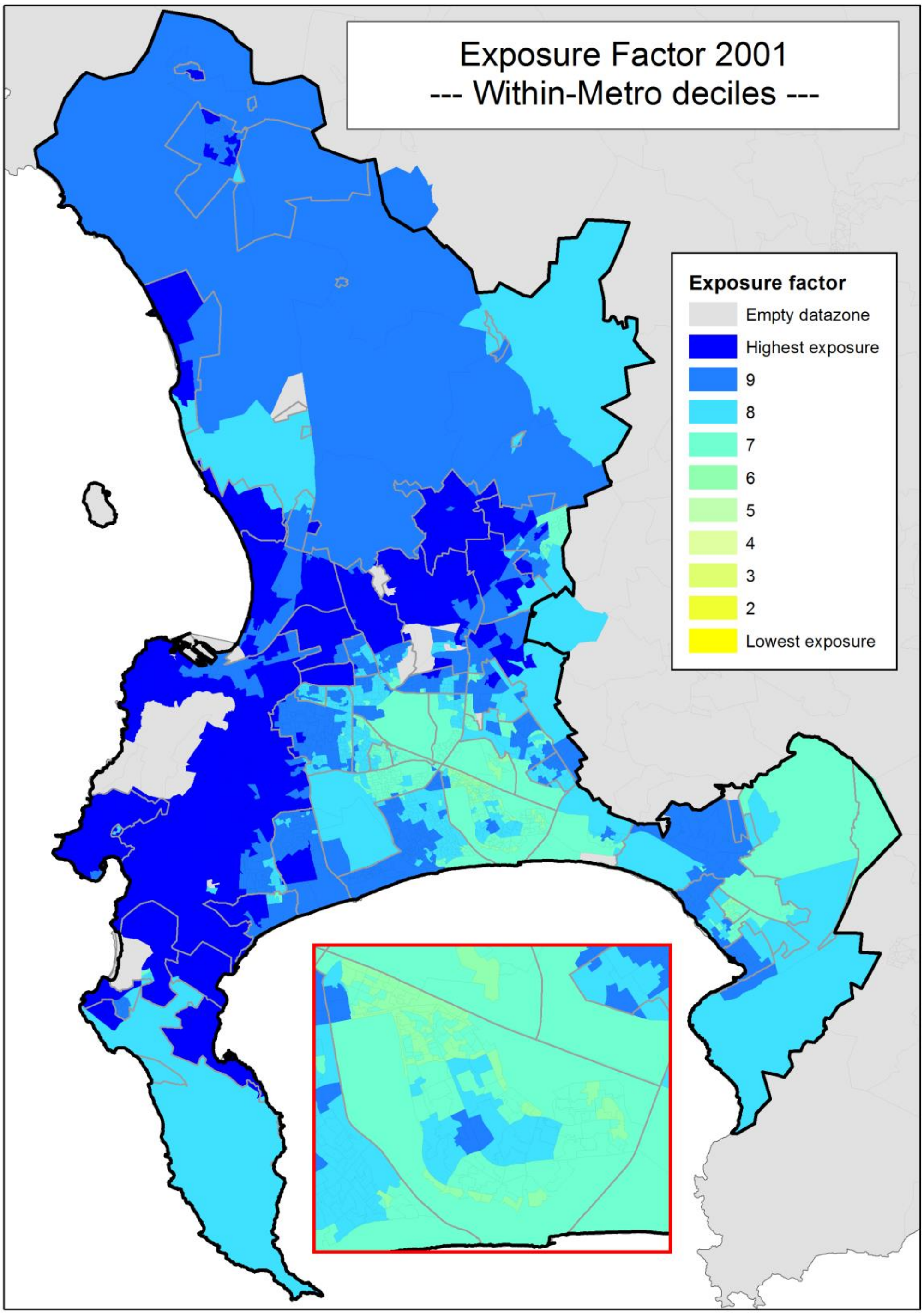
Exposure factor decile	BC	CT	Ek	Jb	Mg	NM	Ts	eT	Total
10 highest exposure	0	328	0	4	0	0	448	0	780
9	0	352	65	269	0	0	94	0	780
8	0	342	155	250	0	19	13	1	780
7	0	292	99	255	0	33	35	66	780
6	0	74	237	260	0	13	41	155	780
5	0	0	254	293	25	22	27	159	780
4	0	0	253	154	18	51	85	219	780
3	16	0	124	114	31	76	66	353	780
2	12	0	1	0	80	180	106	401	780
1 lowest exposure	311	0	0	0	160	98	36	175	780
Total	**339**	**1,388**	**1,188**	**1,599**	**314**	**492**	**951**	**1,529**	**7,800**

I now turn to look at the spatial distribution of exposure to inequality within the City of Cape Town metropolitan municipality. Cape Town is chosen here for specific consideration because it contains a large proportion of the highest-exposure datazones but none of the lowest-exposure datazones (unlike, for instance, Tshwane which contains a broader spread which includes a concentration of very high exposure datazones). Figure 5.15 shows the spatial configuration within Cape Town of the within-metro decile distribution provided by Table 5.6 above. The 328 datazones within Cape Town that fall within the highest exposure decile across the eight metropolitan municipalities are colour coded dark-blue on the map. The light-grey boundaries show the Census Main Places within Cape Town and are provided to aide spatial interpretation of the distribution[34].

Within Cape Town, the highest levels of exposure on the *ExposFac* measure are observed within the more affluent northern and southern suburbs, whilst lower levels of exposure are observed in the more deprived areas south-east of the city centre. The datazones within the large township of Khayelitsha are typically within deciles 6, 7 and 8 of the within-metro decile distribution (with decile 10 representing the highest exposure decile of metropolitan datazones), although we see a geographical pocket of datazones within the central section of Khayelitsha where exposure levels are in the decile 9. As was observed in the deprivation rate analyses within Chapter 3, this central part of Khayelitsha is characterised by notably lower rates of deprivation than the surrounding parts of the township. As such, poor residents of the central (high exposure) part of Khayelitsha are likely to have greater interaction with non-poor people as they go about their daily lives. Despite this particular spatial variation, the map does give the visual impression that the pattern of exposure across Khayelitsha is broadly similar to the patterns across other areas along the N2 highway such as Nyanga.

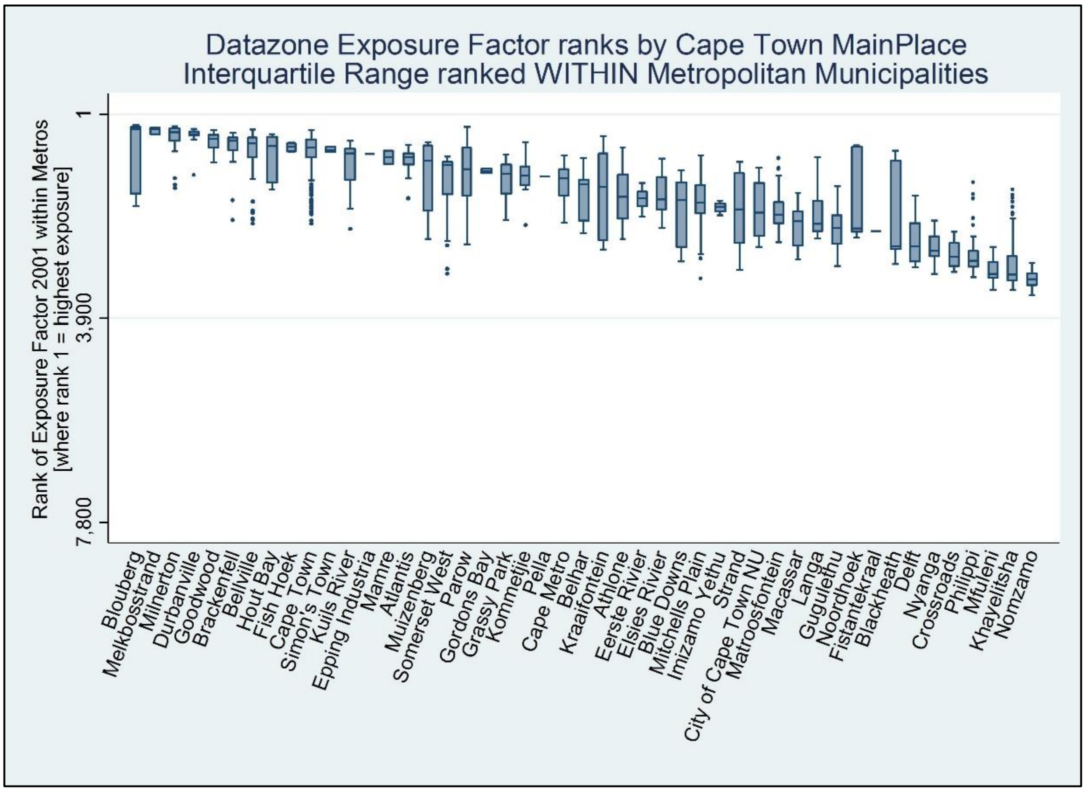
[34] Note: I have used the 2011 Census Main Place boundaries to aide spatial interpretation of the map

Figure 5.15: Within-metro deciles of the *ExposFac* composite exposure measure



To unpick these Cape Town-specific patterns of exposure further, I compare the datazone distributions of *ExposFac* scores between the various 2011 Census Main Places. It is important to note again here (as was noted in the equivalent discussion of the SAIMD ranks within Cape Town above), that the 2001 datazones do ***not*** nest perfectly within the 2011 Main Places. However, overalying the 2011 Main Places onto the 2001 datazones does provide a valuable *general* sense of the spatial patterning within Cape Town

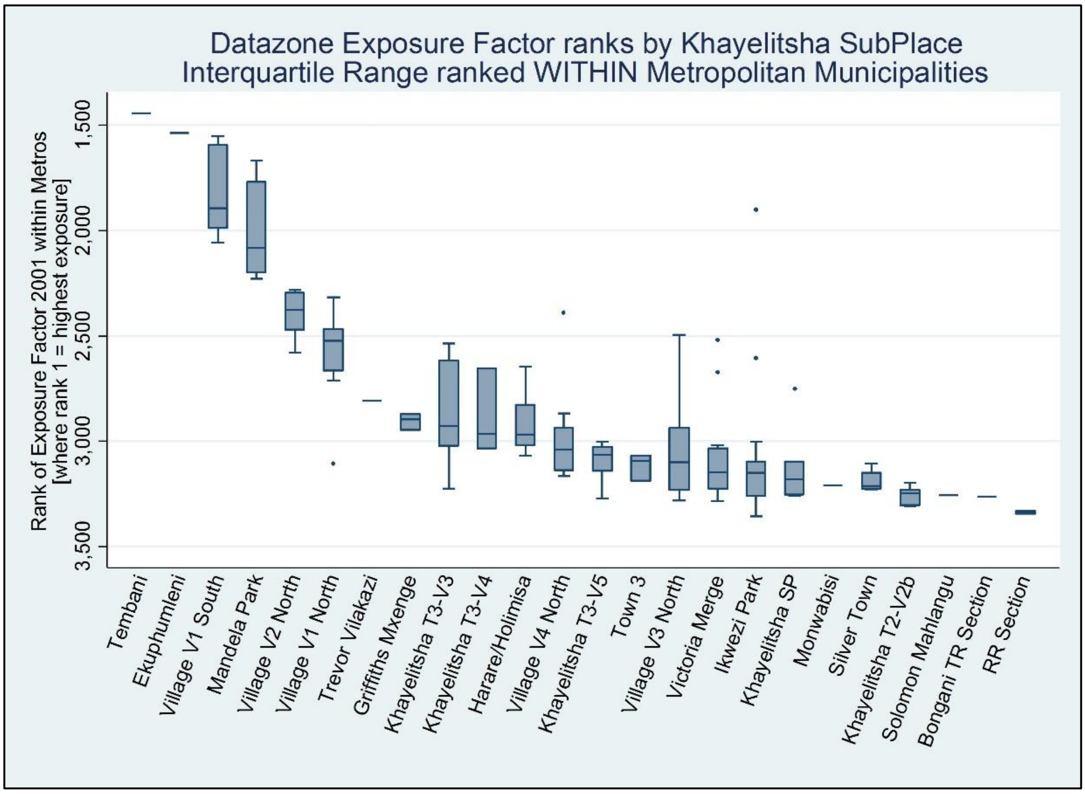
Figure 5.16: Datazone Exposure Factor ranks by Cape Town MainPlace (ranked across metros only)



The Khayelitsha Main Place area can be seen to be positioned towards the right-hand side of Figure 5.16, with the second-lowest median datazone *ExposFac* rank of all Cape Town Main Places (with Nomzano Main Place having a slightly lower median rank). However, although the median rank value places Khayelitsha at the lower end of the Cape Town distribution, there are still some datazones within Khayelitsha that rank relatively highly on this measure compared to the rest of Cape Town. In addition, as noted above, all Cape Town datazones, including all those within Khayelitsha, rank within the 'top half' of the within-metro distribution when all 7,800 metropolitan datazones are considered together. So, although levels of exposure to inequality are typically lower within Khayelitsha than elsewhere in Cape Town, then are still typically higher than most other metropolitan areas of South Africa.

To focus down even further, Figure 5.17 shows equivalent data for datazones just within Khayelitsha, this time grouped according to 2011 Census Sub Place. Again it is important to note that 2001 datazones do ***not*** nest perfectly within 2011 Sub Places (and, indeed, the 2011 Sub Places are less coterminous than the 2011 Main Places with the 2001 datazones), and so these distributions should be regarded as providing a general indicative picture rather than a definitive account.

Figure 5.17: Datazone Exposure Factor ranks by Khayelitsha SubPlace (ranked across metros only)



In some cases, single datazones are associated with single Sub Places, meaning it is impossible to calculate an interquartile range. However, despite these limitations, some valuable information can be extracted from the chart. For instance, it is evident that the datazones within the 'Village V1 South' Sub Place exhibit notably higher within-metro ranks on the *ExposFac* measure than datazones within the 'Silver Town' Sub Place.

In summary, there is a broad inverse relationship at datazone level between the poverty rate and the level of exposure to inequality experienced by the poor. A poor person living in a largely affluent area will be highly exposed to the obvious signs of inequality on a daily basis, whereas a poor person living in a largely poor area will have much less exposure to inequality. However, this relationship is not a perfect inverse correlation because people's exposure to inequality occurs more widely than in their own particular home datazone neighbourhood. We therefore see much higher levels of exposure to inequality within poor neighbourhoods that are situated close to affluent neighbourhoods, than within poor neighbourhood that are surrounded by other poor neighbourhoods.

The datazones exhibiting the highest levels of exposure to inequality are typically found within urban areas and, more specifically, within the major metropolitan municipalities. There is a relatively high degree of positive correlation between the four different dimension-specific measures of exposure to inequality presented here, which offers an empirical justification in support of the conceptual motivation for combining the four measures into a single composite exposure measure using factor analysis. The analyses of the resultant '*ExposFac*' measure across the metropolitan datazones highlights the differential patterns of exposure both between and within the eight metropolitan municipalities, with Cape Town and Tshwane exhibiting the highest levels of exposure, but with far greater heterogeneity within Tshwane than within Cape Town.

When analyses are undertaken *within* Cape Town, we see clear differentials in exposure levels across the city, enabling the identification of generalised patterns across Census Main Places and Census Sub Places.

5.4 Conclusion

In this chapter I have presented a range of new analyses of the spatial distribution of exposure to socio-economic inequality. Levels of exposure amongst poor people were found to be highest in the urban areas and particularly the major metropolitan areas, while being considerably lower in the more rural areas and especially the former homeland areas. Cape Town was seen to contain a large proportion of the highest-exposure datazones across the country. When considering only the metropolitan municipalities, the neighbourhoods with the highest levels of exposure amongst the poor population were typically those that had the lowest poverty rates. However, important differentials were observed within metropolitan municipalities between neighbourhoods with similar poverty rates due to the spatial structuring of the municipality and the respective datazones' geographical locations within the municipality. Poor neighbourhoods that lie close to affluent neighbourhoods tend to have higher exposure scores than poor neighbourhoods lying far from affluent neighbourhoods.

The measures of exposure to inequality presented here provide new and valuable contributions to the evidence base concerning inequality in South Africa. These measures reveal the extent to which people's experiences of inequality are geographically contoured according to where they live and carry out their routine daily activities. Notable differences in levels of exposure to inequality have been observed between neighbourhoods that have fairly similar levels of poverty/deprivation and located within the same municipality, with the differences in exposure due to the spatial configuration of the broader neighbourhood geography. In other words, geography – and particularly place of residence – plays an important role in determining people's exposure to inequality as they go about their daily lives.

The purpose of 'drilling down' from national level, to the combined metropolitan areas, to each metropolitan area separately, to just the City of Cape Town, and finally to neighbourhoods within the City of Cape Town and indeed neighbourhoods within a single township with Cape Town, is to demonstrate how people's lived experience of inequality varies according to where within South Africa a person lives. I contend that neither the conventional measures of inequality, such as national level Gini coefficients, nor the traditional 'global' or 'spatial' measures of residential segregation adequately capture the fine-grained geographical variations in lived experience of inequality which I argue are provided by my new local exposure measures.

These measures of exposure to inequality – which I argue represent a measure of people's 'lived experience' of inequality - can be used as explanatory variables in analyses of people's attitudes towards inequality and options for redress. In Chapter 7 I use these exposure measures in conjunction with the South African Social Attitudes Survey (SASAS) to explore these potential linkages between experience of inequality and attitudes to inequality/redress. However, before I turn to this, in Chapter 7, I first wish to look in more depth in Chapter 6 at the way in which levels of exposure relate to the underlying levels of deprivation.

Chapter 6: Community level ‘intensity’ of exposure to inequality

6.1 Introduction, aims and objectives

As described in my account of the exposure methodology in Chapter 4, and further demonstrated in the analyses of spatial patterns of exposure in Chapter 5, there is a broad negative correlation between datazone deprivation rates and datazone exposure scores across the country (Pearson correlation coefficient = -0.6939). In other words, amongst the ‘poor’ population, levels of personal exposure to inequality are typically highest for those individuals that live in and/or close to areas of relatively low deprivation. For example, I showed in Figure 5.2 that some of the highest levels of exposure were found for poor residents of the mainly affluent suburbs in and around Sandton in Johannesburg. In contrast, levels of personal exposure to inequality are typically lowest for those ‘poor’ individuals that live in and/or close to areas of relatively high deprivation. For example, I showed in Figure 5.3 that some of the lowest levels of exposure were found for poor residents of the mainly rural former homelands. However, I have also showed through the empirical analysis in Chapter 5 that there are some parts of South Africa that are characterised both by high levels of deprivation *and* high levels of exposure to inequality. It is these areas that I wish to focus on here in Chapter 6. Specifically, the aim here is to address research SubQ4: Are there any neighbourhoods across South Africa with high rates of deprivation and high lived experience of inequality?

Examining the spatial relationships between deprivation and exposure is of relevance both in terms of the analyses I undertake in Chapter 7 and in terms of the broader applications of my exposure measures in other research (not presented here). For instance, as I discussed in Chapter 2, many theories concerning crime causation highlight poverty/deprivation *and* socio-economic inequality as potential determinants of crime (e.g. Becker, 1974; Brantingham and Brantingham, 1981; Sampson, 2006; Sampson and Groves, 1989; Shaw and McKay, 1942; Wikstrom, 2004). In his situational action theory (SAT) of crime causation, Wikstrom regards poverty/deprivation and inequality as some of the "causes of the causes" of crime occurrence (Wikstrom, 2006a, p.62), in that poverty/deprivation and inequality shape individuals' attitudes towards "rule breaking" and this in turn shapes the "moral context" of the community. Wikstrom suggests it is the interaction between an individual's propensity to break moral rules and the community's propensity to reject, permit or facilitate moral rule breaking that is the primary cause of criminal activity. In this sense, poverty/deprivation and inequality are seen as some of the causes of certain social attitudes and it is these social attitudes which are the direct causes of crime[35] (Wikstrom, 2004; Wikstrom, 2006a; Wikstrom, 2006b; Wikström, 2005; Wikström, 2014; Wikström and Treiber, 2009).

Wikstrom's SAT provides a constructive framework for considering how the socio-economic stresses and pressures generated by the experience of both poverty/deprivation and inequality may shape people's attitudes (and therefore subsequently their "action responses"). SAT assumes that a range of socio-spatial

[35] Whilst I do not explicitly consider the relationships between crime, poverty/deprivation and exposure to inequality in this thesis, my exposure measures can be used alongside poverty/deprivation rates to explore issues of crime causation in other research, outside the scope of this thesis. Indeed, as I note in Chapter 7, I have undertaken a separate piece of research looking at this very issue, but have not pursued that line of enquiry in my thesis due to reasons which are elaborated in Chapter 7.

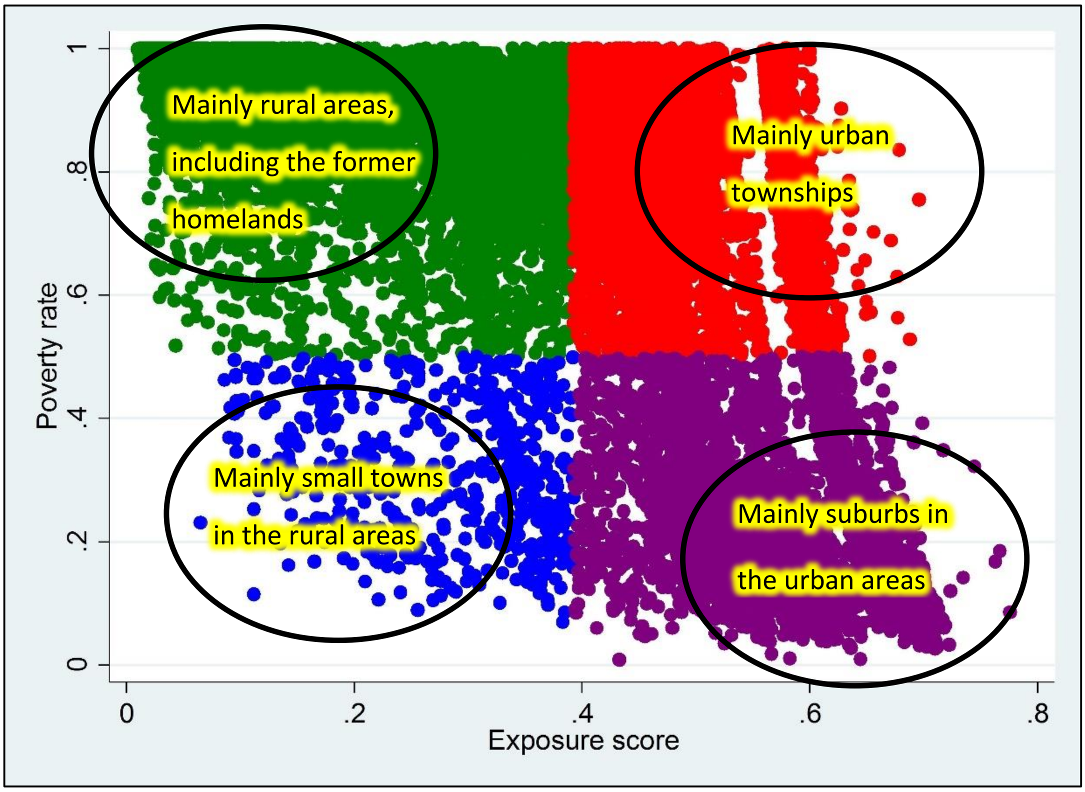
factors, including both poverty/deprivation and inequality, may shape people's attitudes. Furthermore, SAT assumes that people's attitudes are shaped not only by individual experience of these socio-spatial factors, but also by the broader spatial context within which individuals live and carry out the daily activities. In other words, people's attitudes are shaped by the situational contexts in which they have lived during their lives (e.g. neighbourhood characteristics), as well as their own personal experiences. In light of this, one might expect that individuals who are themselves deprived and who live in areas of high deprivation and where they are also highly exposed to inequality may have the most sharply attuned attitudes to inequality and options for redress. By extension, it may be those neighbourhoods characterised by high rates of poverty/deprivation and high levels of exposure to inequality that are most vulnerable to being hotspots of crime and/or social unrest.

In this chapter I therefore look in detail at the spatial distribution of those neighbourhoods that are characterised by high levels of deprivation and exposure. I begin by analysing the exposure measure based on the INC domain and how this relates to the rate of deprivation on the underlying income and material deprivation domain of the SAIMD 2001 at datazone level. I create a new empirical measure, which I term 'intensity', as a measure of community level interaction of exposure to inequality and experience of deprivation. I then proceed to replicate this process for the other three applicable domains of deprivation from the SAIMD and their associated exposure scores. Following the analytical approach adopted in Chapter 5, I first consider all datazones in the country and I then focus down to look at spatial variations between and within the metropolitan municipalities.

6.2 The relationship between deprivation rate and exposure score at datazone level

Before I discuss the methodological construction of my ‘intensity’ measure, it is first instructive to review the extent to which deprivation rates and exposure scores relate at datazone level. Figure 6.1 shows a scatterplot covering all datazones in South Africa, and displays the deprivation rates on the y-axis and the exposure score on the x-axis. To illustrate the different combinations of deprivation level and exposure level that are observed across South Africa, the datazones have been colour-coded according to a four-category scheme. The y-axis differentiates between those datazones with an income deprivation (poverty) rate above and below 0.5 (i.e. those datazones coloured green and red in the chart are areas where half or more of their population is deprived on the income and material deprivation domain). On the x-axis the datazones are split using a cut-off placed at the midpoint of the range of exposure to inequality. Although I appreciate this is a somewhat crude approach to take, it serves a valuable purpose as it allows me to show a general pattern which helps to set the context for the new intensity measure later in the chapter. The crude four-category groupings are not designed to be of equal size and Table 6.1 shows the numbers of datazones that fall within each category. To accompany Figure 6.1 and Table 6.1, I also present Figure 6.2, Figure 6.3, Figure 6.4 and Figure 6.5, which show the spatial patterning of this four-category classification across South Africa, Johannesburg, Cape Town and King Sabata Dalindyebo, respectively.

Figure 6.1: Scatterplot of the four-group categorisation



Note: x-axis split using a cut-off placed at the midpoint of the range of exposure.

Note: y-axis split at poverty score = 0.5 (i.e. 50% of population deprived).

Table 6.1: Proportional split between the four-category classification in Figure 6.1

category	Freq.	Percent	Cum.
Green	12,735	57.46	57.46
Red	5,150	23.24	80.69
Blue	540	2.44	83.13
Purple	3,739	16.87	100.00
Total	22,164	100.00	

Figure 6.2: Four-way categorisation of datazones: all South Africa

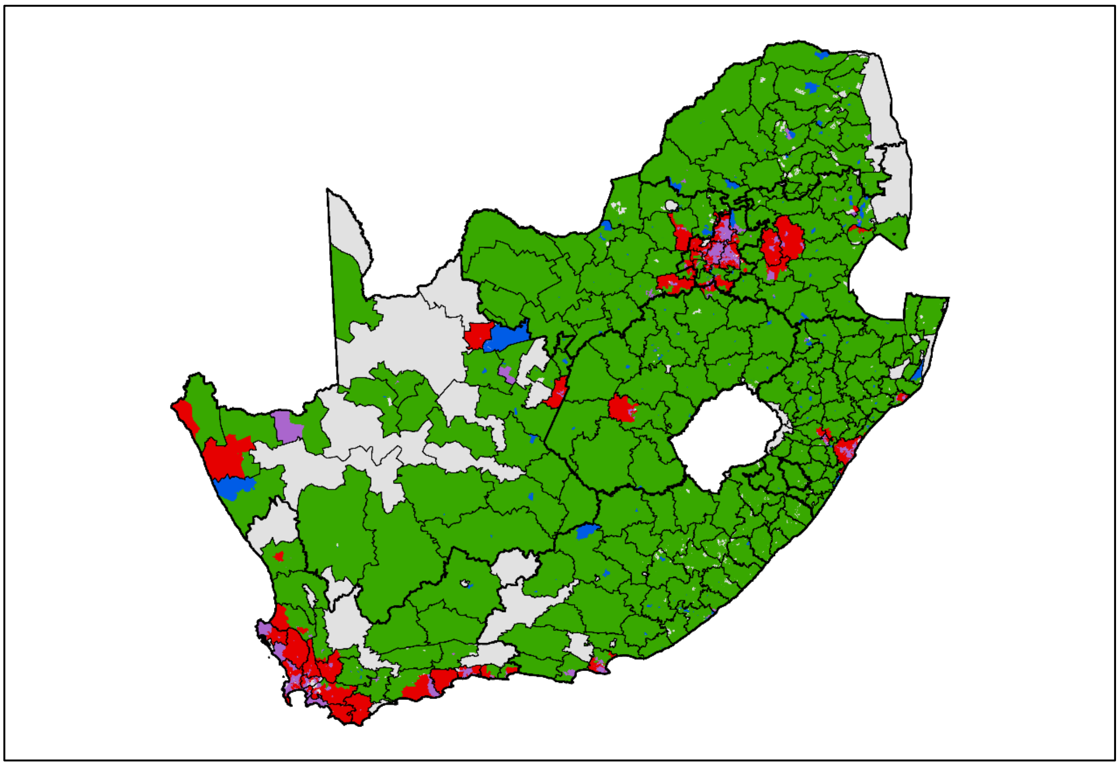


Figure 6.3: Four-way categorisation of datazones: focus on Johannesburg and surrounding areas

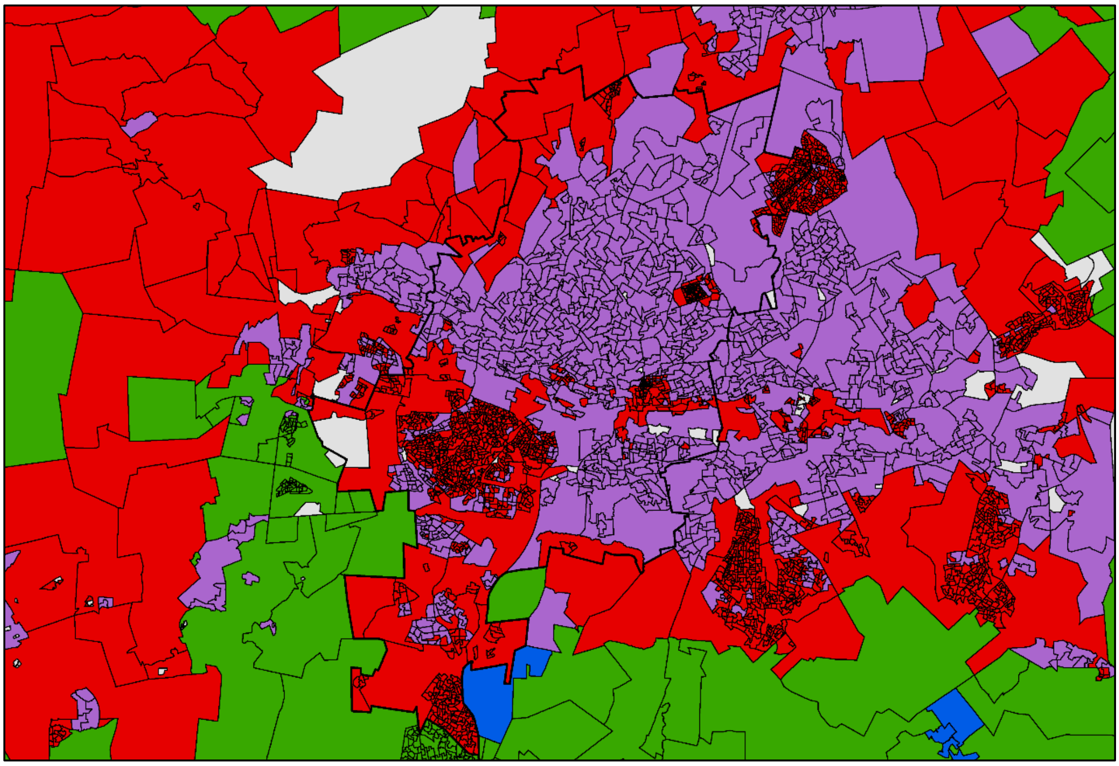


Figure 6.4: Four-way categorisation of datazones: focus on Cape Town and surrounding areas

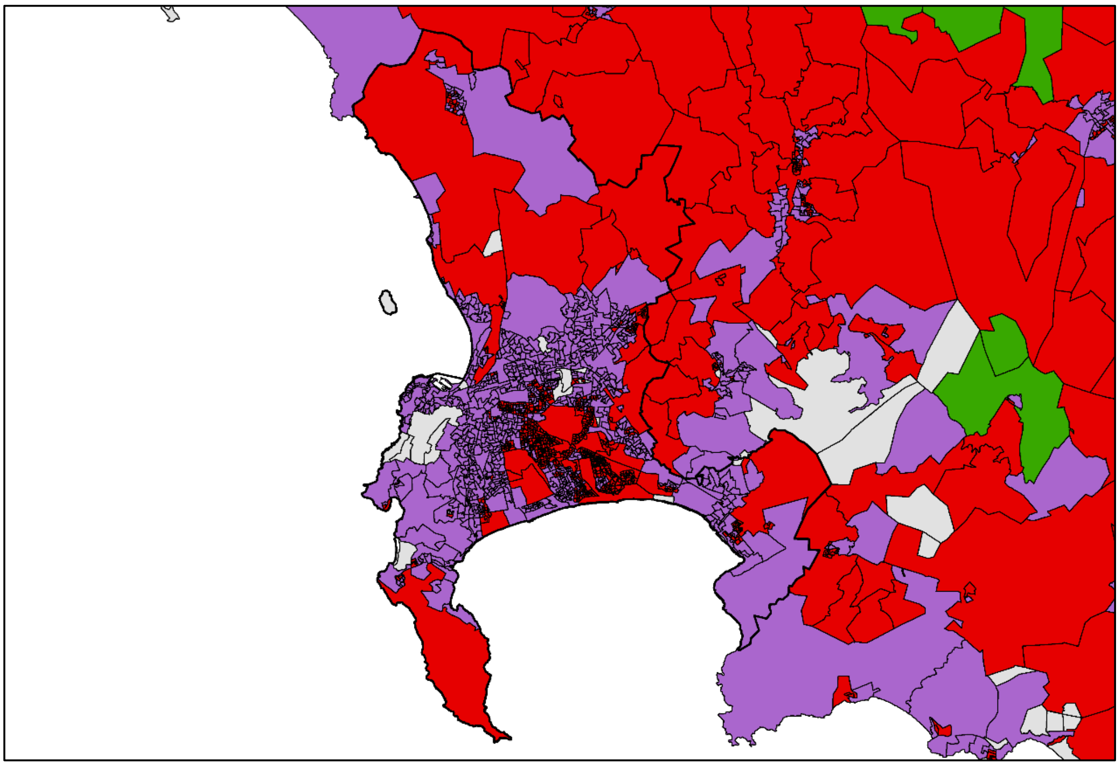
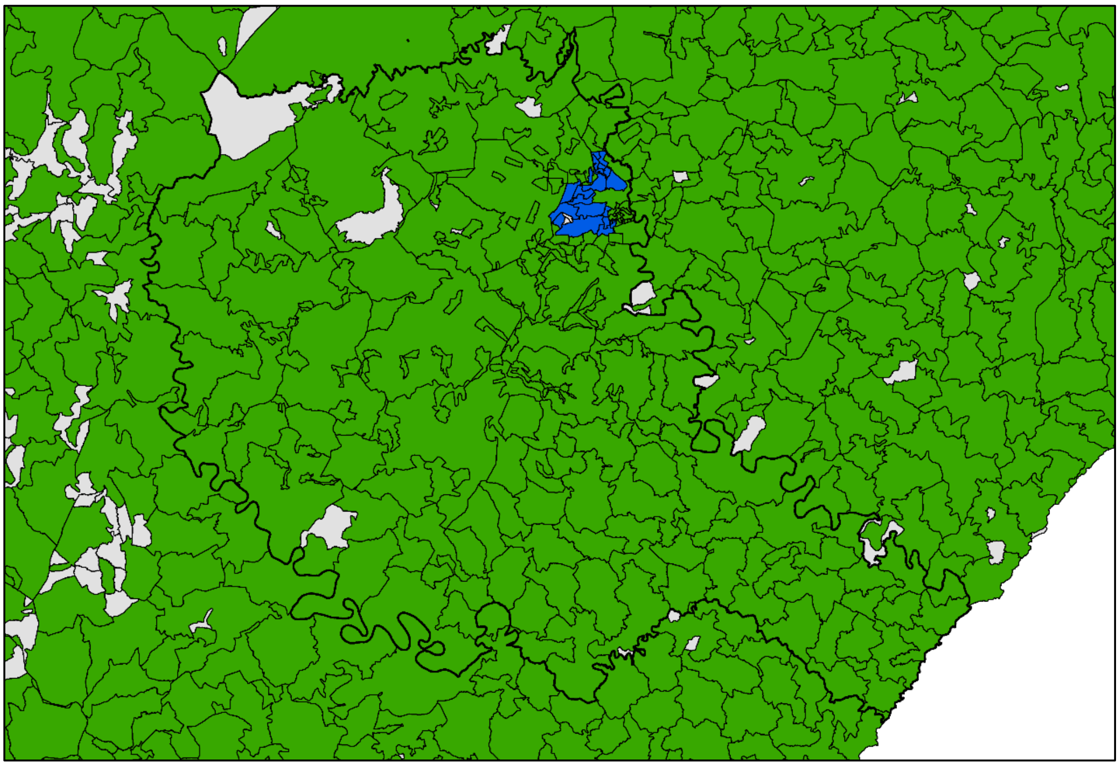


Figure 6.5: Four-way categorisation of datazones: focus on King Sabata Dalindyebo municipality and surrounding areas



It is evident from Table 6.1 that the majority of datazones in the country are colour-coded green in Figure 6.2. These are areas where, broadly, the deprivation rate is high but the exposure score is low. We see from the associated maps that the majority of these green areas are located in the rural areas, including the former homelands. Table 6.1 shows that the blue colour-coded areas on the chart represent only a small fraction of the overall datazones. These are areas where, broadly, the deprivation rate is low and the exposure score is also low. We see from the associated maps that these areas are typically found in the smaller towns scattered across the country, such as the town of Umtata in King Sabata Dalindyebo municipality. Table 6.1 shows that there are a more sizeable number of datazones colour-coded purple on the chart, and these areas are where, broadly, the deprivation rate is low and the exposure score is high. We see from the associated maps that these areas are typically found in the suburbs of the larger urban areas, particularly the metropolitan municipalities, such as the southern suburbs of Cape Town or the suburb of Sandton in Johannesburg. Finally, those datazones colour-coded red are where, broadly, the deprivation rate is high and the exposure score is also high. The associated maps reveal that these areas are typically found in the urban townships, particularly in and around the major metropolitan municipalities.

I suggest here that the *combined* effects of high rates of poverty/deprivation and high exposure to inequality may make the red colour-coded datazones towards the top-right of the chart most vulnerable to crime and social unrest. I suggest that where either or both of these two factors is lower, the vulnerability to crime and unrest is also lower. Similarly, in the analyses of associations between exposure and attitudes in Chapter 7, I hypothesise that people living in areas characterised by both high deprivation and high exposure to inequality may have the strongest aversion to inequality as their attitudes are shaped both by their own personal experiences of inequality and by the experiences of others in their local neighbourhood.

To explore this issue in more detail, I move away from the relatively crude four-category classification presented above and proceed to introduce a new ‘intensity’ measure which is constructed with the aim of highlighting more clearly those areas with the highest combined levels of poverty/deprivation and exposure to inequality.

6.3 Building the new community ‘intensity’ measure of exposure to inequality

The intensity measure is constructed as the mathematical product of the datazone deprivation rate and the datazone exposure score. In other words:

$$Intensity\_xy = \text{deprivation\_rate} * aLDPxy_i$$

As both the deprivation rate and the exposure score range between 0 and 1, the resulting intensity score also ranges between 0 and 1, with a higher score indicating greater intensity of community exposure to inequality[36].

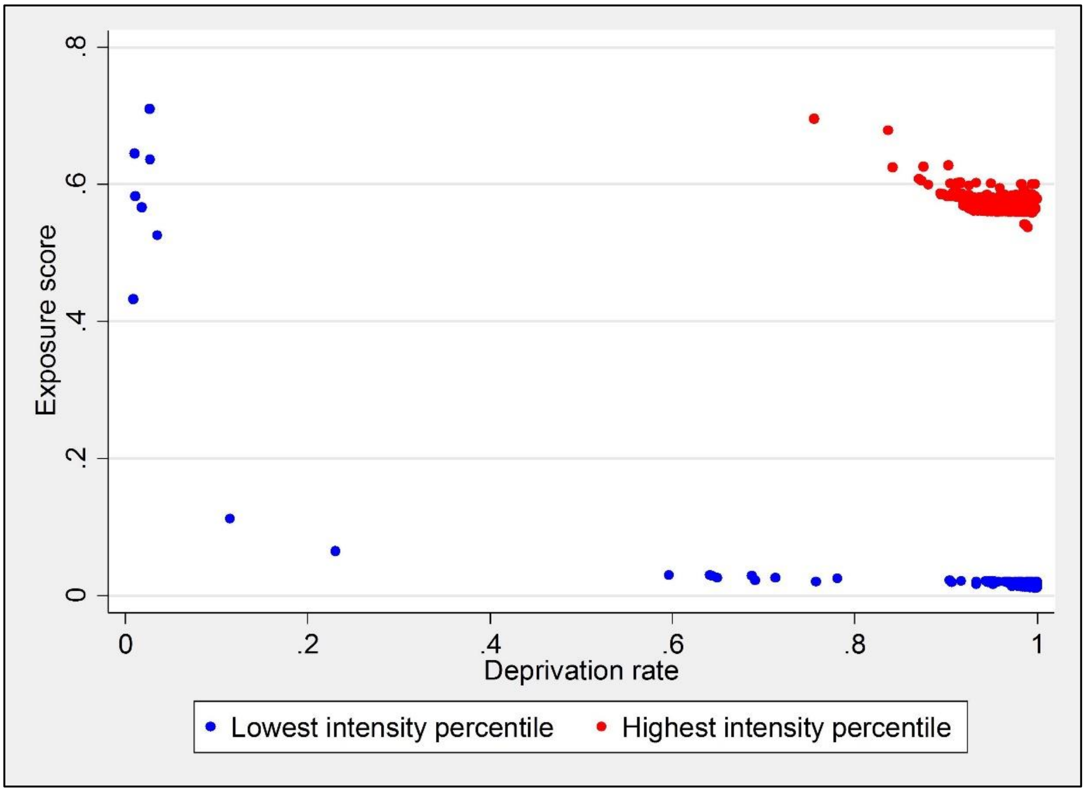
The *intensity_xy* score cannot be interpreted in any absolute sense (i.e. it does not relate to a proportion of population (as the poverty rate does) and it does not relate to a probability (as the exposure measure does)), but its interpretation is nevertheless

[36] Note: it is in practice not technically possible for a Datazone to score a value of 1 because in order to do so would require the Datazone poverty rate to be 1 (i.e. a rate of 100%) and the Datazone exposure score to be 1 (i.e. everyone in that Datazone and all surrounding Datazones to be non-poor), which is clearly contradictory and therefore impossible to achieve.

relatively straightforward: the higher the *intensity_xy* score, the higher the combined effects of high poverty *and* high exposure to inequality.

To demonstrate how the effects of poverty rate and exposure score interact to generate the *intensity_xy* scores, Figure 6.6 presents a scatter plot of the relationship between these two variables for two groups of Datazones.

Figure 6.6: Scatter plot showing the relationship between poverty rate and exposure to inequality for the highest and lowest *intensity_xy* percentiles of Datazones in South Africa



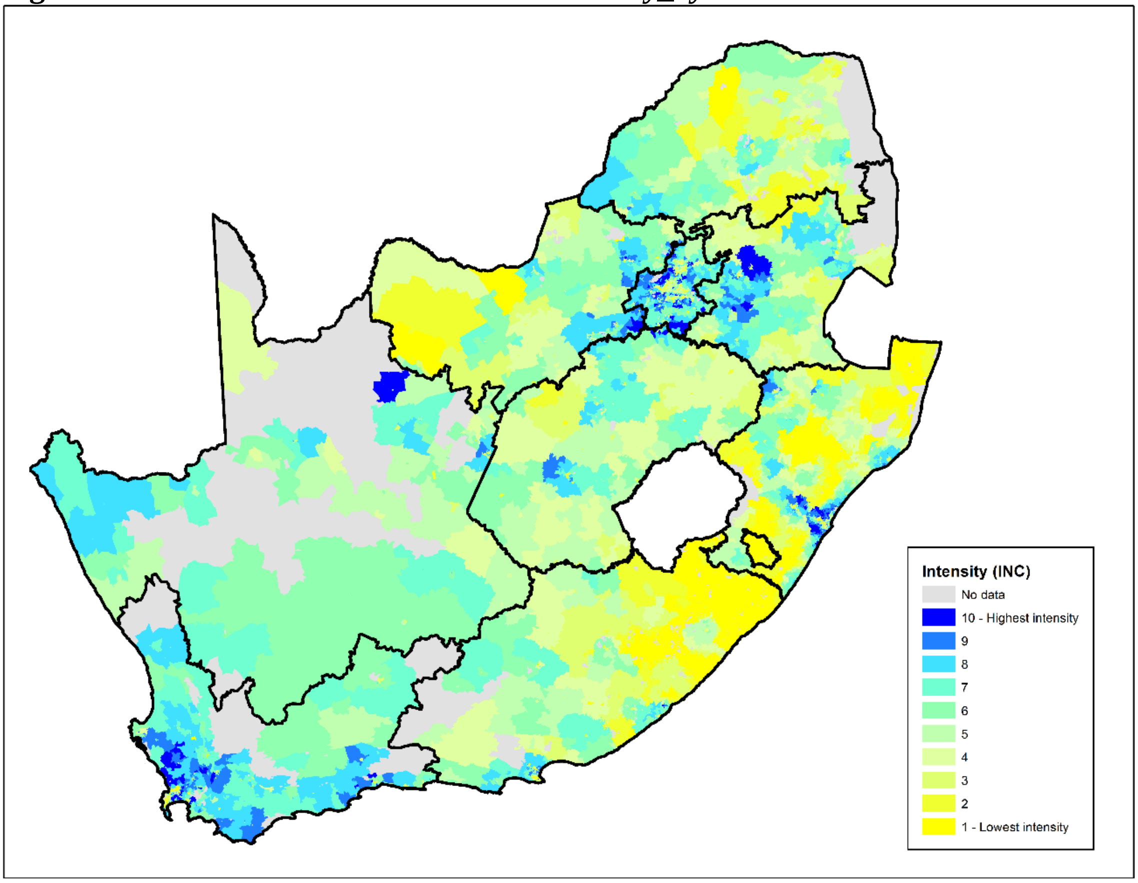
The y-axis shows the datazone exposure score while the x-axis shows the datazone poverty rate. The red dots depict those Datazones which are the 1% highest *intensity_xy* datazones across the country, whilst the blue dots depict the 1% lowest *intensity_xy* datazones across the country. From Figure 6.6 it is apparent that those datazones that score the highest on the intensity measure exhibit both relatively high rates of poverty and relatively high levels of exposure. It is also evident that those datazones that score the lowest on the *intensity_xy* measure exhibit either relatively low levels of poverty, or relatively low levels of exposure, or relatively low levels of both. An important feature of the intensity measure is that it permits large cancellation effects between the two constituent components. This is a deliberate methodological feature as my argument is that areas of high intensity should be characterised by high levels of both deprivation *and* exposure to inequality.

6.4 Results: Intensity scores for all datazones across the country

The analyses of intensity of exposure to inequality are structured into two sub-sections. I first consider the spatial distribution of intensity scores across the whole country using the intensity measure constructed from the income and material deprivation domain (i.e. INC). I then proceed to focus on the metropolitan areas only and expand the analyses to look across intensity measures constructed from all four domains (i.e. INC, EMP, EDU and LIV).

Figure 6.7 shows the *intensity_xy* scores for all datazones in South Africa. The datazones have been grouped into national deciles according to their scores on the *intensity_xy* distribution. Those areas colour-coded bright yellow are the 10% of datazones with the *lowest* levels of intensity on this measure, whilst those areas colour-coded bright blue are the 10% of datazones with the *highest* level of intensity on this measure. The other eight colour-coded categories represent the other eight national deciles of datazones when ranked on this measure.

Figure 6.7: National datazone deciles of *intensity_xy*

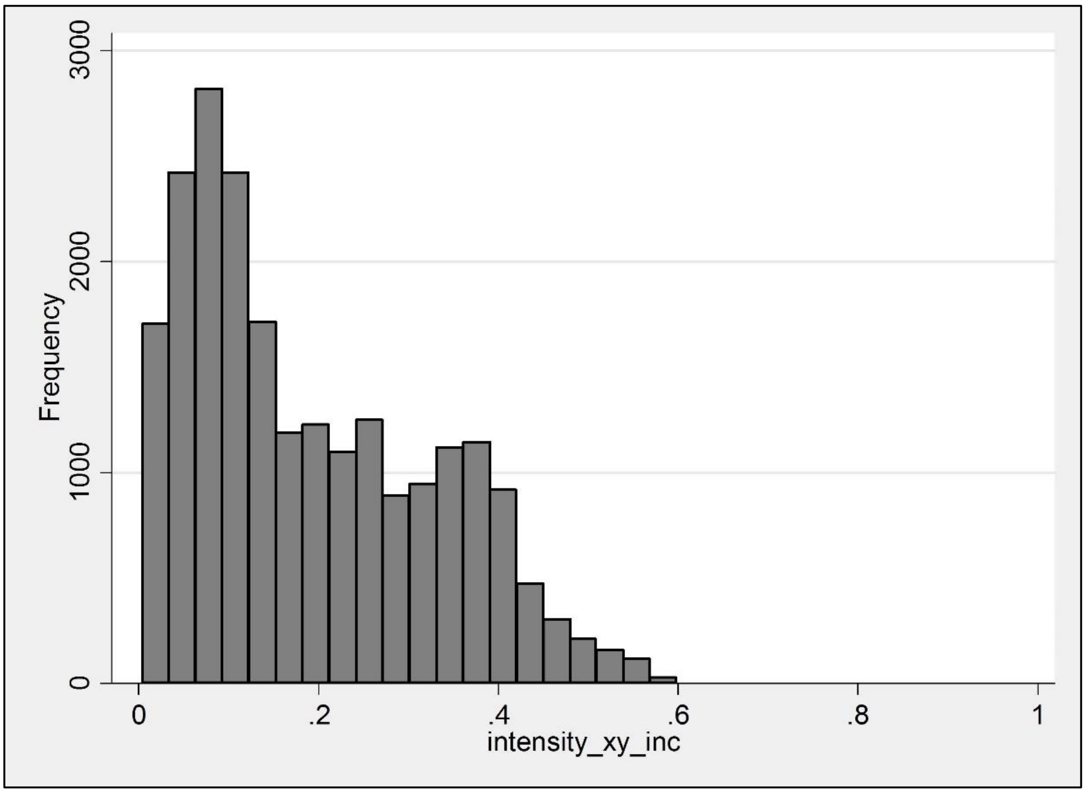


As noted above, whilst national level maps can be problematic for showing the level of detail provided by datazones, it is evident from Figure 6.7 that the highest intensity datazones appear to be located in and around the major urban areas. I will now explore this in more detail using a range of empirical analyses.

Figure 6.8 shows the distribution of datazone level *intensity_xy* scores for all datazones in South Africa. There is a clear positive-skew to the data. The highest intensity score is a value of 0.5987 which relates to a datazone covering the 'Site 5' informal settlement that borders the township of Dunoon in Cape Town. This particular datazone has a deprivation score of 0.9984 (i.e. almost 100% of the population is deprived) and an exposure score of 0.5997. In contrast, the lowest intensity datazone is located in the Bergrivier municipality in the Western Cape where the deprivation score is 0.0088 (i.e. less than 1% of the population is deprived) and the exposure score is

0.4327. In this particular datazone, while the exposure level is relatively high, there are hardly any poor people and so the intensity measure is very low (0.0038). The fourth-lowest intensity datazone, on the other hand, is located in the Ntabankulu municipality in the Eastern Cape, with an intensity score of 0.0096. In this particular datazone the deprivation score is extremely high (0.9991) but the exposure score is very low (0.0097).

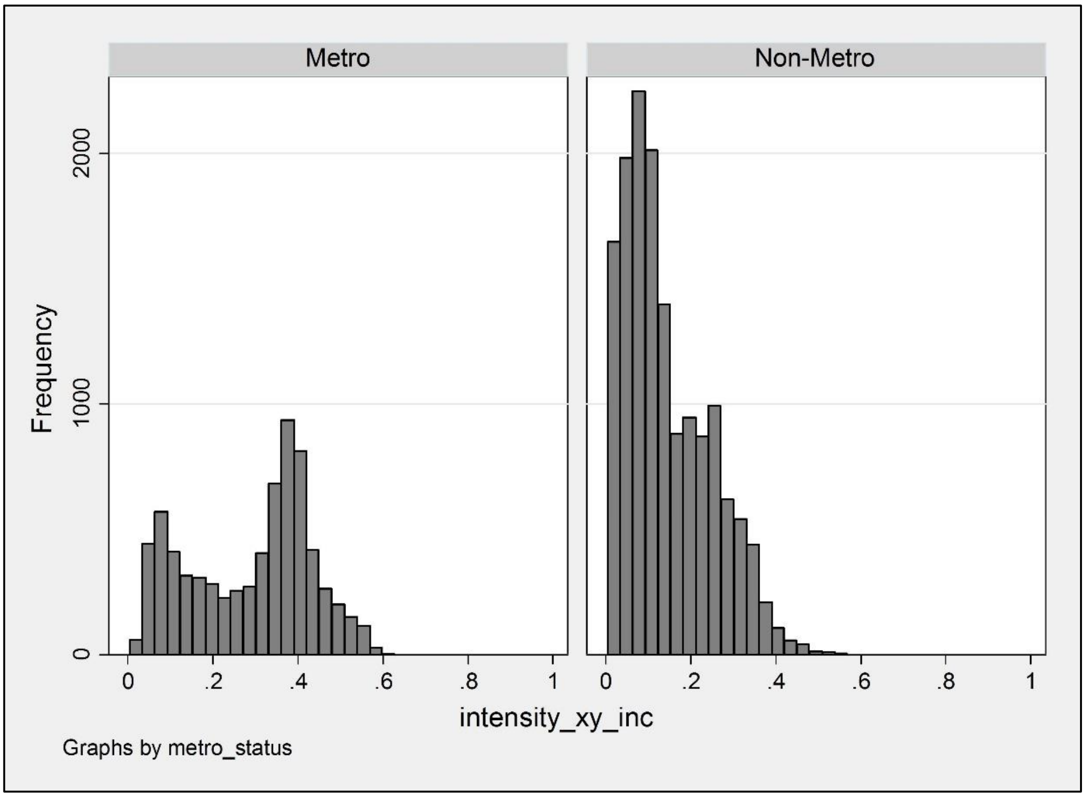
Figure 6.8: Histogram of Datazone level '*intensity_xy*' scores for all Datazones in South Africa



The spread of datazone level intensity scores shown in Figure 6.8 also indicates a slight bi-modal distribution, with a smaller secondary peak of values just below 0.4 on the intensity scale. To unpick this bi-modal element further, I performed some sub-group analyses which revealed that there was again a notable differentiation in distributions between metropolitan and non-metropolitan datazones.

Figure 6.9 shows the datazone intensity scores disaggregated by metropolitan-non-metropolitan status, again showing intensity scores along the x-axis.

Figure 6.9: Histogram of Datazone level '*intensity_xy*' scores for all Datazones in South Africa, separately by metropolitan/non-metropolitan status

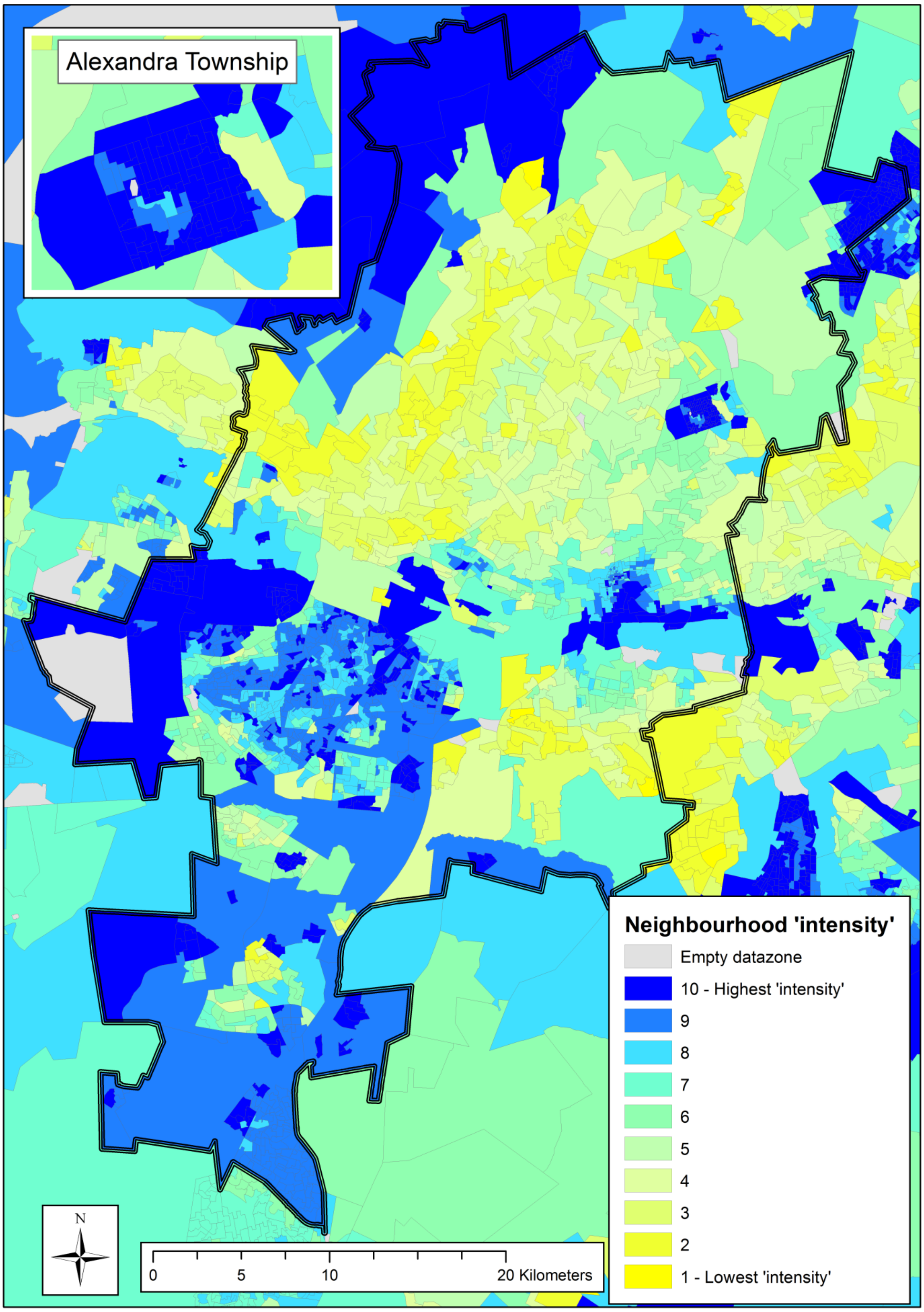


For the non-metropolitan datazones the distribution of *intensity_xy* values lies primarily towards the lower intensity end of the scale, which relates closely to the relatively low levels of exposure in the majority of non-metropolitan areas (compared to the metropolitan areas). The metropolitan datazones exhibit a bimodal distribution on the *intensity_xy* measure, with one concentration of areas scoring relatively low on this measure and another concentration scoring relatively high. From Figure 5.5 above we know that all the metropolitan datazones scored relatively highly on the $aLDPxy_i$* exposure measure, so the metropolitan datazones that score relatively low values on the *intensity_xy* measure will be those with the relatively low deprivation rates.

The power of the *intensity_xy* measure is best demonstrated cartographically using a case study approach. Figure 6.10 shows a map of the City of Johannesburg area with the datazones colour-coded according to the national deciles of the *intensity_xy* distribution. The areas coloured bright yellow are the 10% of datazones across South Africa with the lowest levels of *intensity_xy*, while the areas coloured bright blue are the 10% of datazones across South Africa with the highest levels of *intensity_xy*. The bright blue areas are therefore those that are characterised by relatively high levels of deprivation *and* relatively high levels of exposure to inequality.

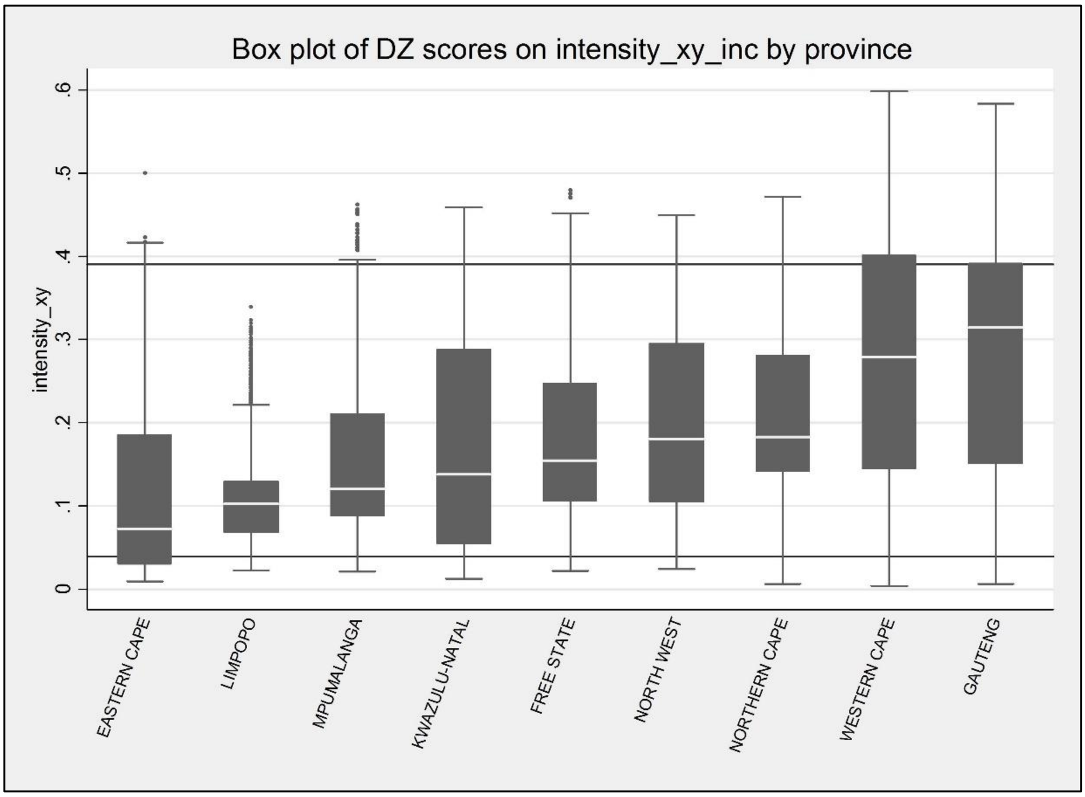
Figure 6.10 reveals that areas in and around Sandton score relatively low on the *intensity_xy* measure, despite scoring relatively highly on the $aLDPxy_i$* exposure measure, reflecting the fact that although poor people living in Sandton will be acutely aware of socio-economic inequality, there are very few poor people in Sandton and therefore the intensity of community sense of injustice is likely to be relatively low. However, in the proximately located township of Alexandra, the level of *intensity_xy* is extremely high, reflecting the high degree of deprivation in Alexandra and the high degree of exposure to inequality due to being surrounded by relatively less deprived (and in many cases very affluent suburbs).

Figure 6.10: Map of Datazone level '*intensity_xy*' score, national deciles



Each of the nine provinces contains some datazones in the lowest national decile of *intensity_xy*, whilst all provinces except Limpopo contain some datazones in the highest *intensity_xy* national decile. Figure 6.11 shows the datazone distributions on *intensity_xy* for each of the nine provinces separately, with the upper horizontal reference line representing the cut point for the highest intensity decile (i.e. all datazones scoring above this line would be coloured bright blue on the map) and the lower horizontal reference line representing the cut point for the lowest *intensity_xy* national decile (i.e. all datazones scoring below this line would be coloured bright yellow on the map).

Figure 6.11: Box plot of Datazone level *intensity_xy* scores by province



It is evident from Figure 6.11 that the provinces of Western Cape and Gauteng exhibit notably different distributions to the other seven provinces, with far greater proportions of datazones in Western Cape and Gauteng scoring towards the upper end of the *intensity_xy* distribution.

Of the 2216 datazones that constitute the highest *intensity_xy* national decile, 1981 (i.e. 89.4% of that decile of datazones) are located in the five metropolitan municipalities of Cape Town, Johannesburg, Tshwane, Ekurhuleni and eThekwini. The remaining 235 datazones from that highest intensity national decile (10.6% of that decile of datazones) are spread across a further 32 municipalities.

Table 6.2: Location of datazones in the 10% highest *intensity_xy* decile nationally

Municipality	Number	Percentage
City of Cape Town	487	22.0%
Ekurhuleni Metro	447	20.2%
City of Johannesburg Metro	409	18.5%
Ethekwini	389	17.6%
City of Tshwane Metro	249	11.2%
Others (32 municipalities)	235	10.6%
Total in the 10% highest exposure decile nationally	2,216	100.0%

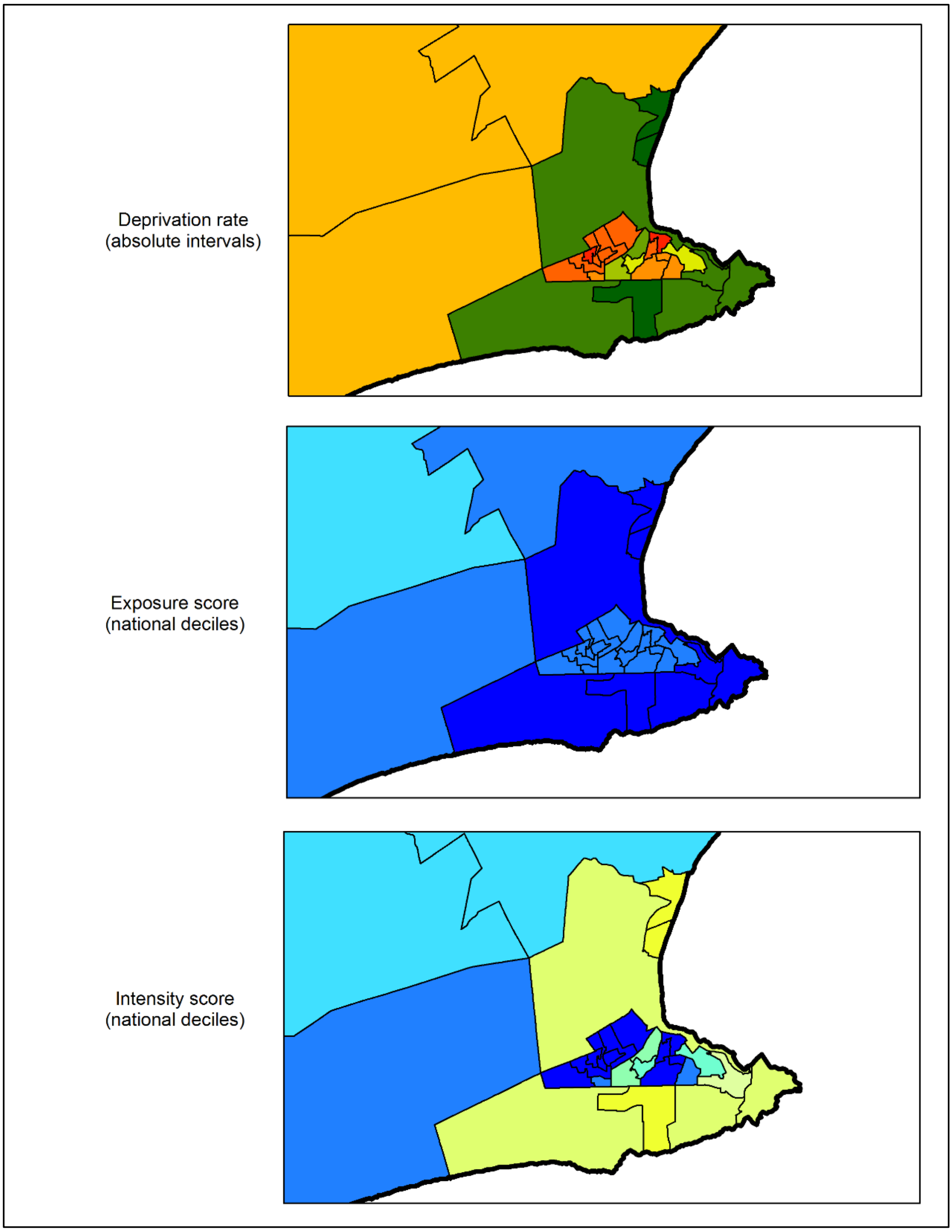
In terms of the *proportion* of datazones in each municipality that fall within the highest *intensity_xy* decile nationally, Ekurhuleni ranks top, with 37.6% of the datazones in that municipality being within the highest *intensity_xy* decile nationally. Table 6.3 shows the ten municipalities with the greatest proportions of datazones in this highest intensity decile, and it is evident that all five of the municipalities listed in Table 6.2 are also listed in Table 6.3. From this we can conclude that the highest intensity datazones are primarily, although not entirely, located in metropolitan areas.

Table 6.3: The ten municipalities with the highest proportions of datazones in the 10% highest *intensity_xy* decile nationally

Municipality	Number of datazones in the municipality	Number of datazones in the 10% highest intensity decile nationally	Percentage of municipality datazones in the 10% highest intensity decile nationally
Ekurhuleni (Gauteng)	1,188	447	37.6%
Mossel Bay (Western Cape)	37	13	35.1%
City of Cape Town (Western Cape)	1,388	487	35.1%
Metsimaholo (Free State)	62	21	33.9%
Middelburg (Mpumalanga)	65	21	32.3%
George (Western Cape)	67	19	28.4%
City of Tshwane (Gauteng)	951	249	26.2%
Overstrand (Western Cape)	27	7	25.9%
City of Johannesburg (Gauteng)	1,599	409	25.6%
eThekwini (KwaZulu-Natal)	1,529	389	25.4%

Before moving on to examine the metropolitan areas in more detail, and to expand the breadth of the intensity measures to cover those based on the EMP, EDU and LIV domains, I will first give a little more detail on the results for Mossel Bay. This is a non-metropolitan municipality located on the southern coastline approximately midway between the cities of Cape Town and Port Elizabeth. Although this municipality is fairly small in terms of population size, consisting of just 37 datazones in total, over a third of these datazones are within the highest intensity national decile. To illustrate the socio-spatial patterning that generated these high intensity scores I have placed three maps together in Figure 6.12. In panel A of Figure 6.12 I present the deprivation rate using the same absolute interval colour coding approach as in Chapter 3, where the dark green areas are datazones with deprivation rates less than 10%, compared to the bright red areas which are datazones where between 90% and 100% of the population are deprived. In Panel B of Figure 6.12 I present the exposure scores using the same national decile colour coding as in Chapter 5, where the bright yellow areas are datazones in the lowest exposure national decile and the dark blue areas are datazones in the highest exposure national decile. Finally in Panel C of Figure 6.12 I present the intensity scores using the same colour coding approach as earlier in Chapter 6, where the bright yellow areas are those datazones in the lowest intensity national decile and the dark blue areas are those datazones in the highest intensity national decile. The maps focus specifically on the main town of Mossel Bay in order to display sufficient detail. The remainder of the municipality (not shown on the maps) includes some additional low-deprivation datazones along the coastal stretch to the north of the area shown, plus large swathes of sparsely populated rural inland areas. The area shown on the map therefore captures the vast majority of the municipality's constituent datazones. It is evident from Figure 6.12 that the town of Mossel Bay consists of a concentration of very deprived datazones bordering and indeed interspersed with a series of considerably less deprived datazones. It is this stark contrast in deprivation levels between proximately located datazones in a relatively confined geographical space, including the presence of a number of highly deprived datazones, which leads to the relatively high levels of intensity of exposure revealed in Table 6.3.

Figure 6.12: Mossel Bay 'intensity' case study



In summary, therefore, with regards to the lived experience of the poor population in South Africa, the $aLDPxy_i$* exposure measure captures the experience of inequality from the perspective of individual day-to-day experiences, whilst the *intensity_xy* measure captures a community effect relating to the interaction between high neighbourhood poverty rates and high exposure to inequality. Exposure to inequality, neighbourhood poverty rate and the interaction between the two (i.e. *intensity_xy*) may each potentially influence the attitudes of poor members of South African society, as I will explore in Chapter 7. Before moving on to those analyses, however, I first wish to look in more detail at patterns of intensity within the metropolitan areas. As part of this, I assess the degree of commonality or difference between intensity measures calculated based on the four domain-specific measures of deprivation and exposure.

6.5 Results: Intensity scores for metropolitan datazones

Equivalent intensity measures have been constructed for each of the other three appropriate dimensions of deprivation measured in the SAIMD (namely, employment, education and living environment). Although these new measure have been constructed for all datazones across South Africa, the focus here in this section is solely on the metropolitan areas and so the decile-based analyses (including the maps) are based on deciles within the metropolitan datazones only ('within-metros'). To further illustrate the methodology, Figure 6.13, Figure 6.14 and Figure 6.15 focus on the township of Khayelitsha in Cape Town and show the within-metros decile distribution for the four sets of deprivation rates, exposure scores and intensity scores, respectively. As noted above, each intensity score is the mathematical product of the respective deprivation rate (on a scale of 0 to 1) and exposure score (also on a scale of 0 to 1). As such, those datazones that have relatively high deprivation rates and relatively high exposure scores will have relatively high intensity scores. However, in cases where either or both of the deprivation rate or exposure score is relatively low, then the intensity score will also be relatively low.

The shift from national deciles used above to within-metro deciles used in the following analyses has the effect of increasing the level of differentiation between metropolitan datazones. This shift does not in any way affect the calculation of the underlying intensity scores, only the analytical application of them. The ranking scheme therefore ranges from rank 1 which is the highest intensity datazone within the metropolitan areas, to rank 7,800 which is the lowest intensity datazone within the metropolitan areas[37].

The patterns of intensity within Khayelitsha are similar but not identical across the four domain-specific measures. The exposure scores are typically in the mid- to upper deciles of the 'within-metros' distribution, with slightly higher levels in the central portion of the township which can be seen to have somewhat lower deprivation rates than the areas around the edges of the township. The higher deprivation rates around the township edges can be seen to result in higher intensity scores in these areas. These detailed maps demonstrate how levels of intensity of community exposure vary quite dramatically between relatively proximate parts of the township. This finding can be extended across the metropolitan areas and indeed across the entire country. It highlights the added value of adopting a neighbourhood-based approach to examine socio-spatial issues: simply using a Khayelitsha average would mask these measurable differences between constituent datazones within the township thereby obscuring the potential statistical associations between deprivation/exposure/intensity and other social phenomena of interest.

[37] Note that here I present the deprivation rates using within-metro deciles, in contrast to the approach adopted earlier in this thesis of showing deprivation rates using absolute intervals.

Figure 6.13: Dimension-specific deprivation levels within Khayelitsha, showing deciles ranked across all metropolitan municipality datazones

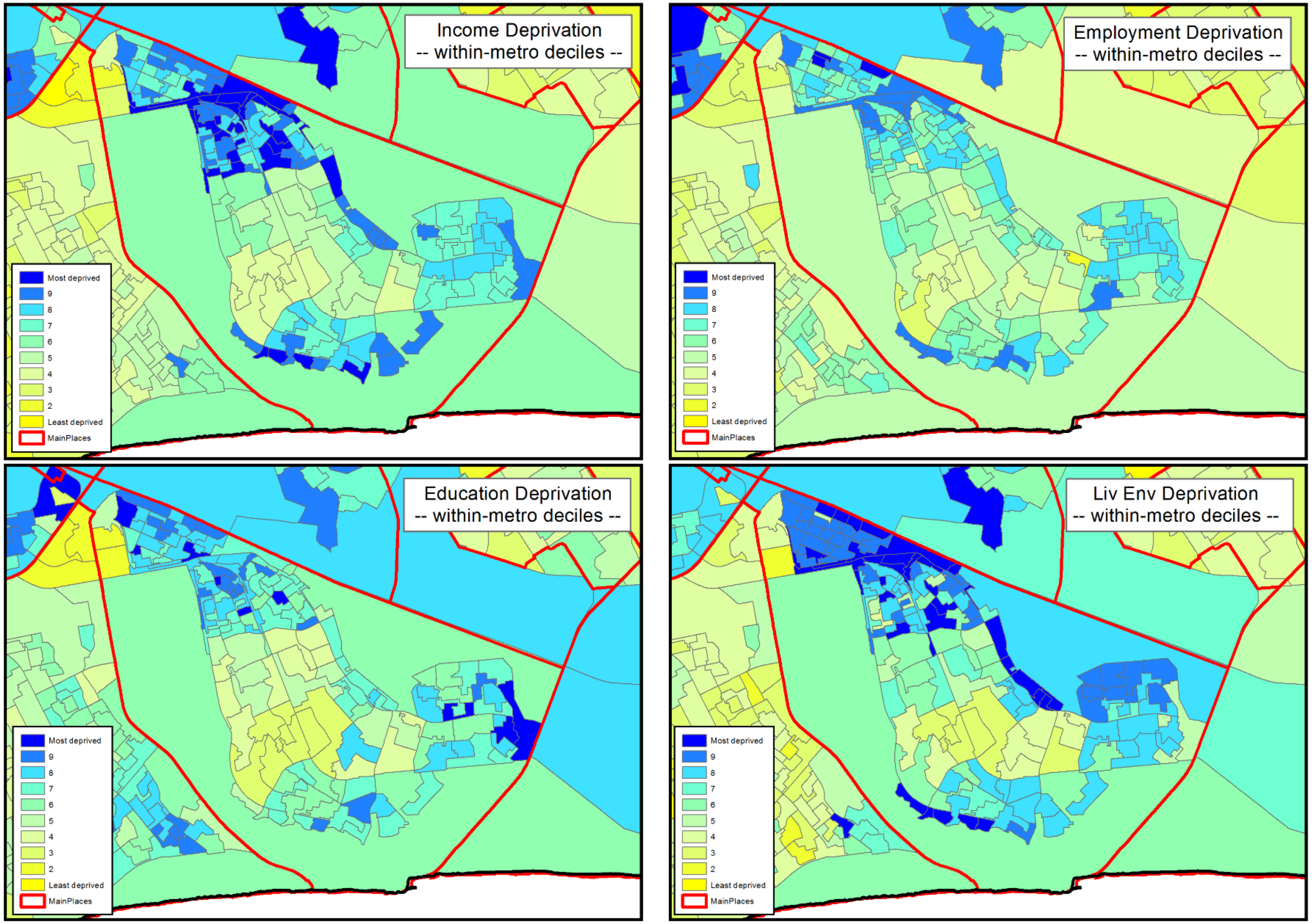


Figure 6.14: Dimension-specific Exposure levels within Khayelitsha, showing deciles ranked across all metropolitan municipality datazones

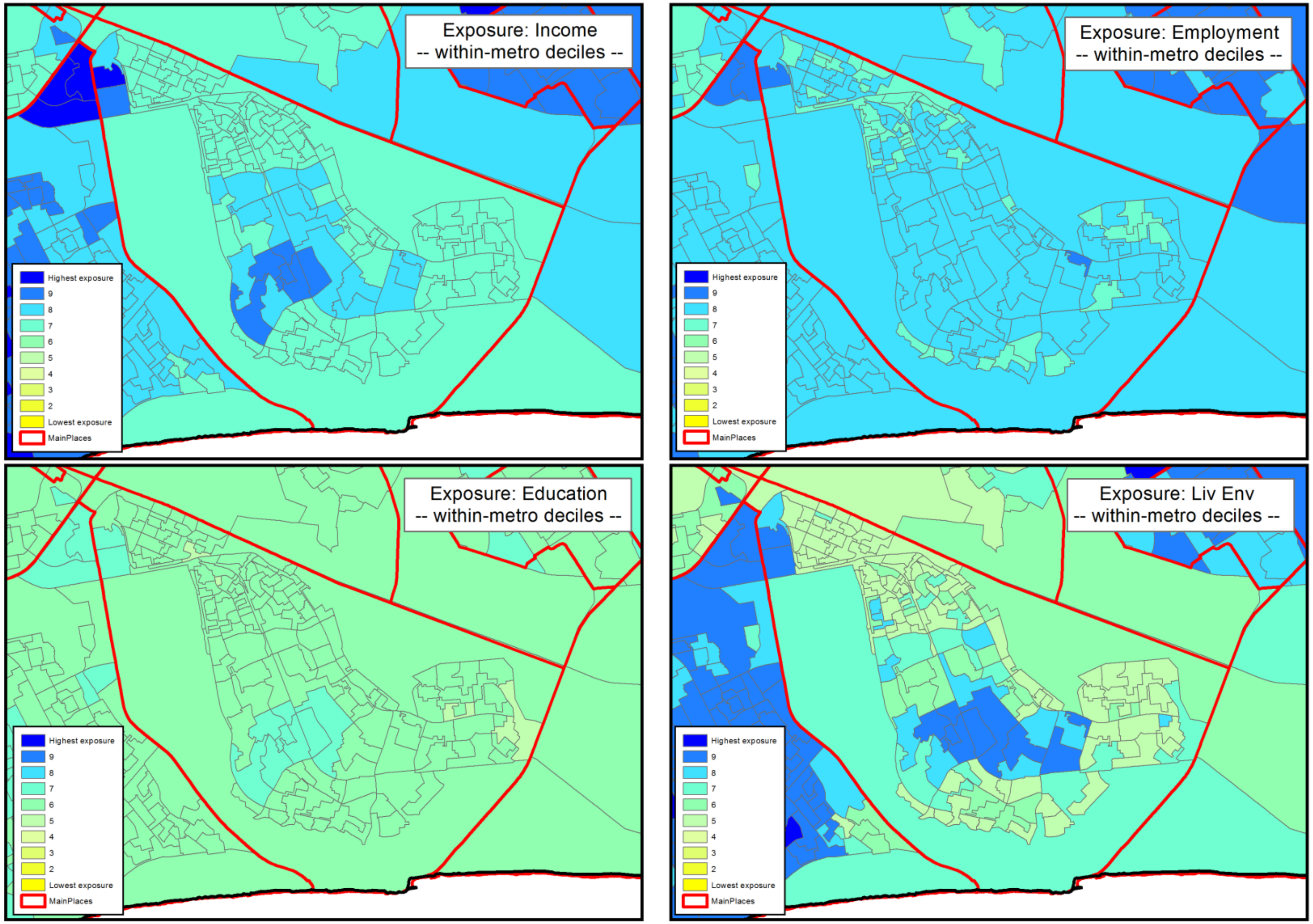
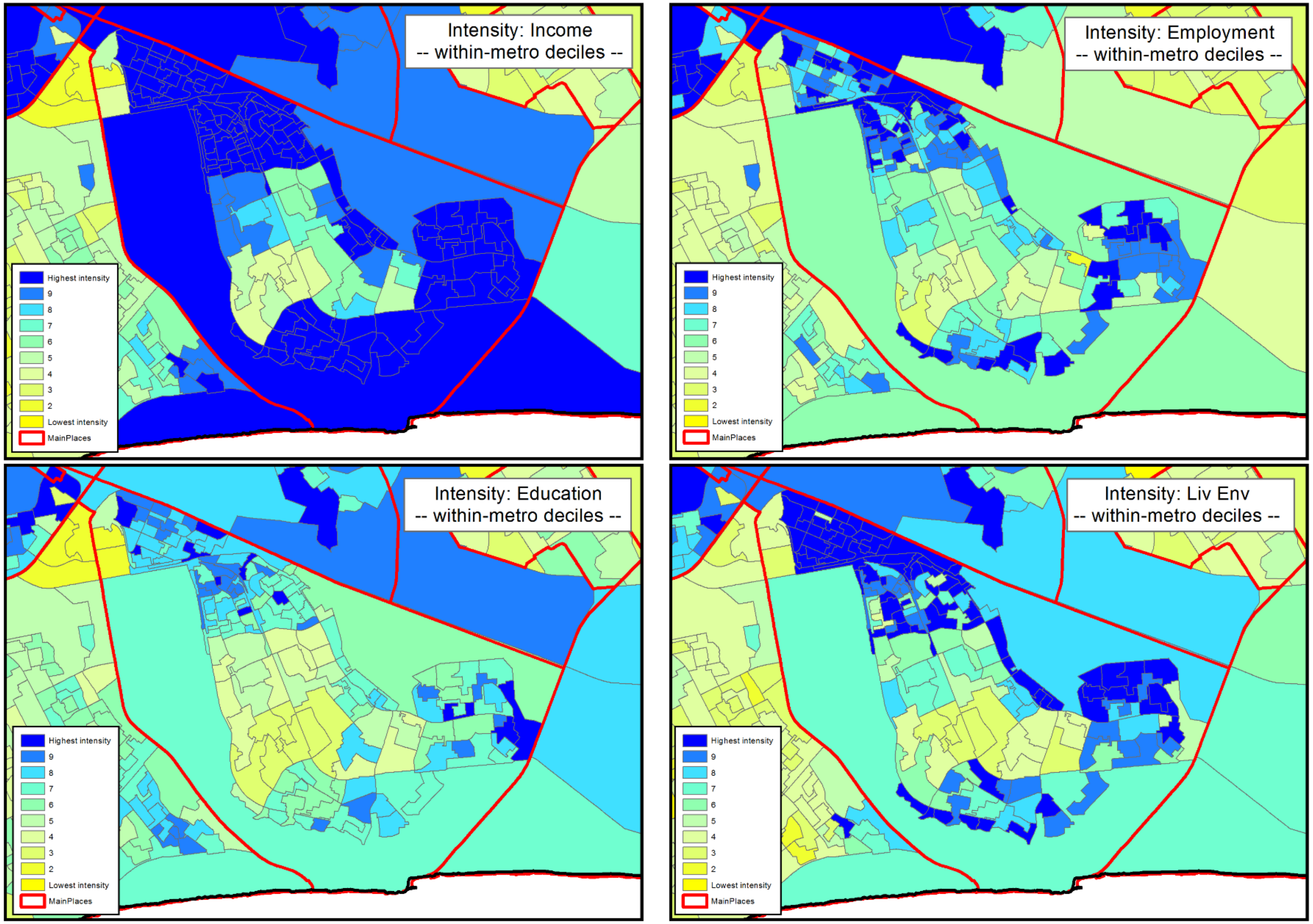


Figure 6.15: Dimension-specific 'Intensity' levels within Khayelitsha, showing deciles ranked across all metropolitan municipality datazones



As per the analyses of exposure presented above, there is a conceptual basis for exploring whether the four separate intensity measures can be combined into a single overall composite intensity measure. The correlation coefficients presented in Table 6.4 and Table 6.5 also confirm that there is an empirical basis for generating a composite measure of intensity. The coefficients are strong between all four dimensions of intensity when assessed across all metropolitan datazones, and are stronger still when assessed only across the Cape Town datazones.

Table 6.4: Spearman rank correlation coefficients between the four dimension-specific exposure measures, all metropolitan datazones

All Metros	*intensity_inc*	*intensity_emp*	*intensity_edu*	*intensity_liv*
intensity_inc	1			
intensity_emp	0.8245	1		
intensity_edu	0.7960	0.7068	1	
intensity_liv	0.8810	0.7529	0.8399	1

Table 6.5: Spearman rank correlation coefficients between the four dimension-specific exposure measures, Cape Town datazones only

Just Cape Town	*intensity_inc*	*intensity_emp*	*intensity_edu*	*intensity_liv*
intensity_inc	1			
intensity_emp	0.9329	1		
intensity_edu	0.8666	0.8122	1	
intensity_liv	0.9320	0.8780	0.8402	1

The same methodology was adopted to produce a new composite Intensity Factor variable as was described in the construction of the Exposure Factor composite variable above. In summary, the 7,800 metropolitan datazones were ranked independently on each of the four dimension-specific intensity measures (income, employment, education, and living environment), and the four sets of ranks transformed to normal distributions. These four normally distributed rank variables were then entered into a maximum likelihood factor analysis. The factor analysis generated a single factor (with an Eigenvalue of 3.072), and the generated factor weights were used to combine the four intensity dimensions to produce the new composite Intensity Factor measure. Table 6.6 shows the factor loadings reported.

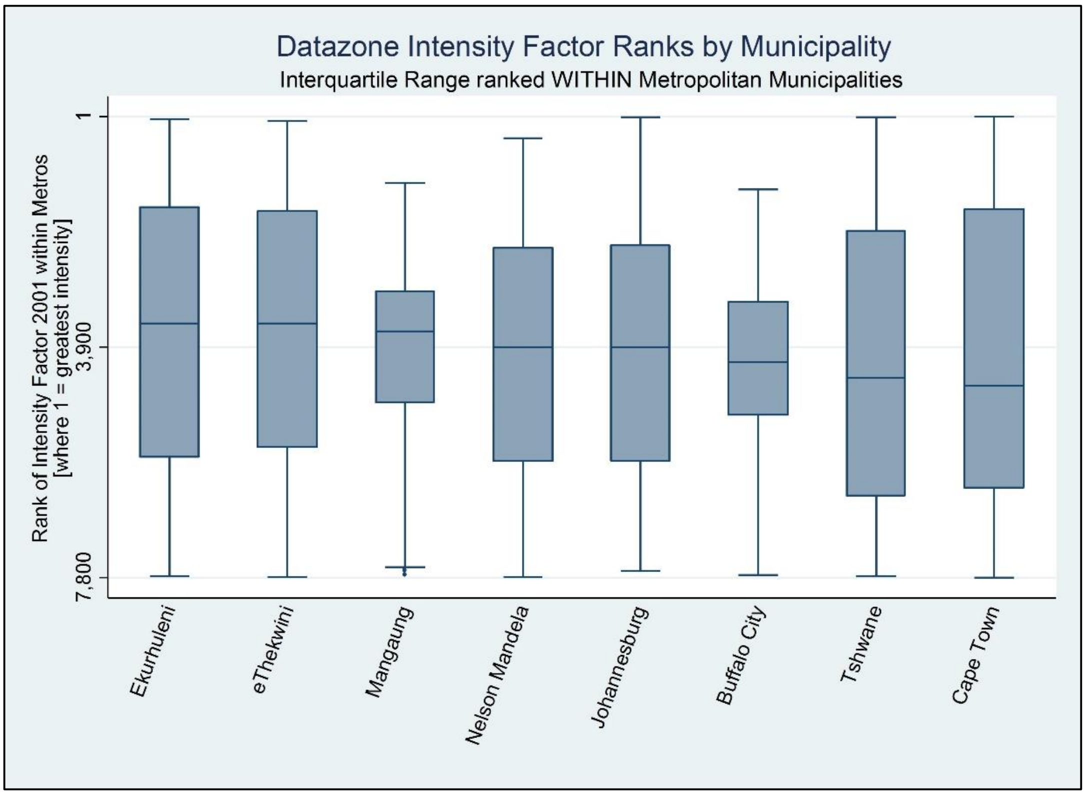
Table 6.6: Factor loadings from construction of *Intensity_Fac*

Variable	Factor1	Uniqueness	Standardised factor loading
inc	0.949	0.099	0.271
emp	0.806	0.351	0.230
edu	0.819	0.329	0.234
liv	0.922	0.149	0.264

Figure 6.16 shows the datazone rank distribution across metropolitan municipalities on the new Intensity Factor measure. As per the analyses above, rank 1 relates to the highest intensity metropolitan datazone, and rank 7,800 relates to the lowest intensity metropolitan datazone.

A remarkable degree of commonality exists on this Intensity Factor measure between the datazone distributions across the eight metropolitan municipalities. The median datazone ranks are fairly similar across the eight municipalities, ranging from a high median rank of 3,498 in both eThekwini (absolute value of 3,498) and Ekurhuleni (average value of 3,498, where the two datazones at the centre of the municipality distribution are ranked 3,496 and 3,500), down to a low of 4,554.5 in Cape Town. Each municipality contains some datazones that rank at the lowest intensity end of the distribution, and all municipalities except Mangaung and Buffalo City contain some datazones that rank at the very highest intensity end of the distribution.

Figure 6.16: Datazone Intensity Factor ranks by municipality (ranked across metros only)



The similarity observed in the metropolitan distributions shown in Figure 6.16 is a function of the calculation of the intensity measure. In broad terms, as the level of deprivation increases the level of exposure decreases. This has an equalising effect on the levels of intensity when the datazone distributions are compared between the eight municipalities. It is important to state here that the primary purpose of the exposure and intensity measure is *not* to compare cities against one another in terms of their median ranks, but rather to show detailed spatial patterns of inequality at neighbourhood level *within* the municipalities.

To illustrate this further, Table 6.7 presents summary statistics for the distributions of data seen in Figure 6.16, showing counts of datazones per municipality that fall within each decile of datazones when ranked across the eight municipalities combined. We see that Cape Town has the largest number of datazones in the highest intensity decile, but also the largest number in the lowest intensity decile. Overall, we see a relatively even spread of datazones across the deciles in most municipalities, with the exception of the relative lack of datazones from Mangaung or Buffalo City in the highest intensity deciles.

Table 6.7: Datazone level Intensity Factor, by within-metro decile

Intensity factor decile	**BC**	**CT**	**Ek**	**Jb**	**Mg**	**NM**	**Ts**	**eT**	**Total**
10 highest intensity	0	255	168	130	0	11	93	123	780
9	3	91	132	144	7	66	90	247	780
8	28	90	117	159	27	53	115	191	780
7	52	87	118	180	62	58	83	140	780
6	63	83	124	185	98	58	44	125	780
5	82	122	104	164	35	64	88	121	780
4	47	169	92	151	33	47	82	159	780
3	19	139	104	193	8	31	91	195	780
2	20	88	137	205	16	31	143	140	780
1	25	264	92	88	28	73	122	88	780
Total	**339**	**1,388**	**1,188**	**1,599**	**314**	**492**	**951**	**1,529**	**7,800**

I turn now to focus specifically on the spatial pattern of intensity, measured through the new composite Intensity Factor variable, within the municipality of Cape Town. Figure 6.17 shows cartographically the spatial distribution of the data for Cape Town that was presented in Table 6.7 above. The dark blue colour-coded datazones are those that are within the 10% highest intensity decile across the eight metropolitan municipalities, whilst the bright yellow colour-coded datazones are those that are within the 10% lowest intensity decile across the eight metropolitan municipalities.

With regards to the relatively affluent northern and southern suburbs, we see that despite scoring highly on the exposure measure (see Figure 5.15), these areas score relatively lowly on the intensity measure. So, in these areas, deprived individuals are exposed to inequality to a high degree as they go about their daily lives but, because there are relatively few deprived people living within those neighbourhoods, there is a relatively low neighbourhood-level intensity of exposure.

The greatest concentration of high intensity datazones within Cape Town can be seen in the area running south-east along the N2 highway away from the city centre. However, there are also notable geographical pockets of high intensity areas in certain outer-lying townships, such as Imizamo Yethu (close to Hout Bay on the western slopes of Table Mountain); Fistantekraal (to the north-east of Durbanville); Boekombos (in Kraaifontein); Witsand (in Atlantis); Nonzamo (to the south of Somerset West); Vrygrond (Muizenberg); and Masiphumelele (to the south of Table Mountain).

Figure 6.17: Datazone level within-metro deciles on the Intensity Factor measure

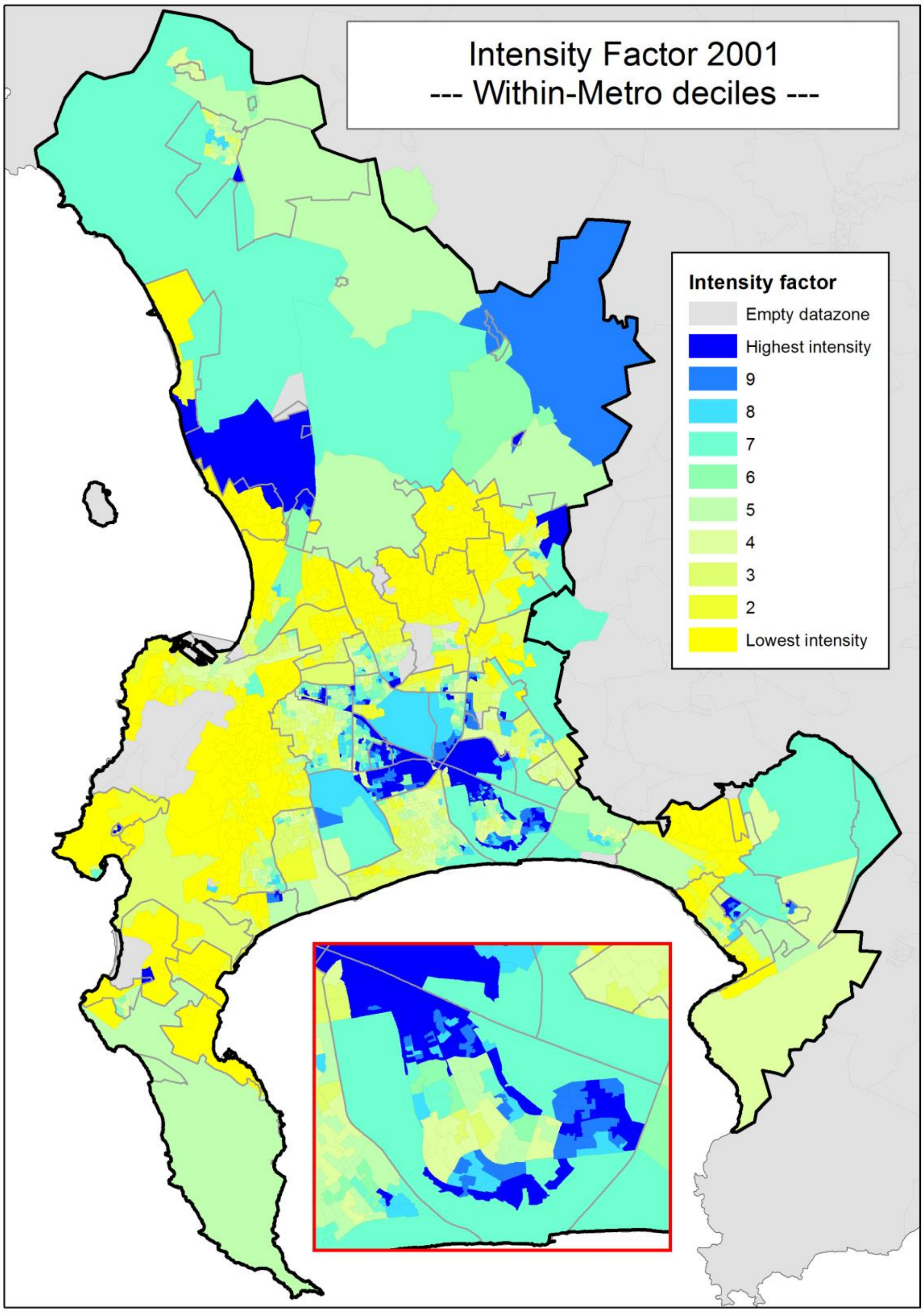
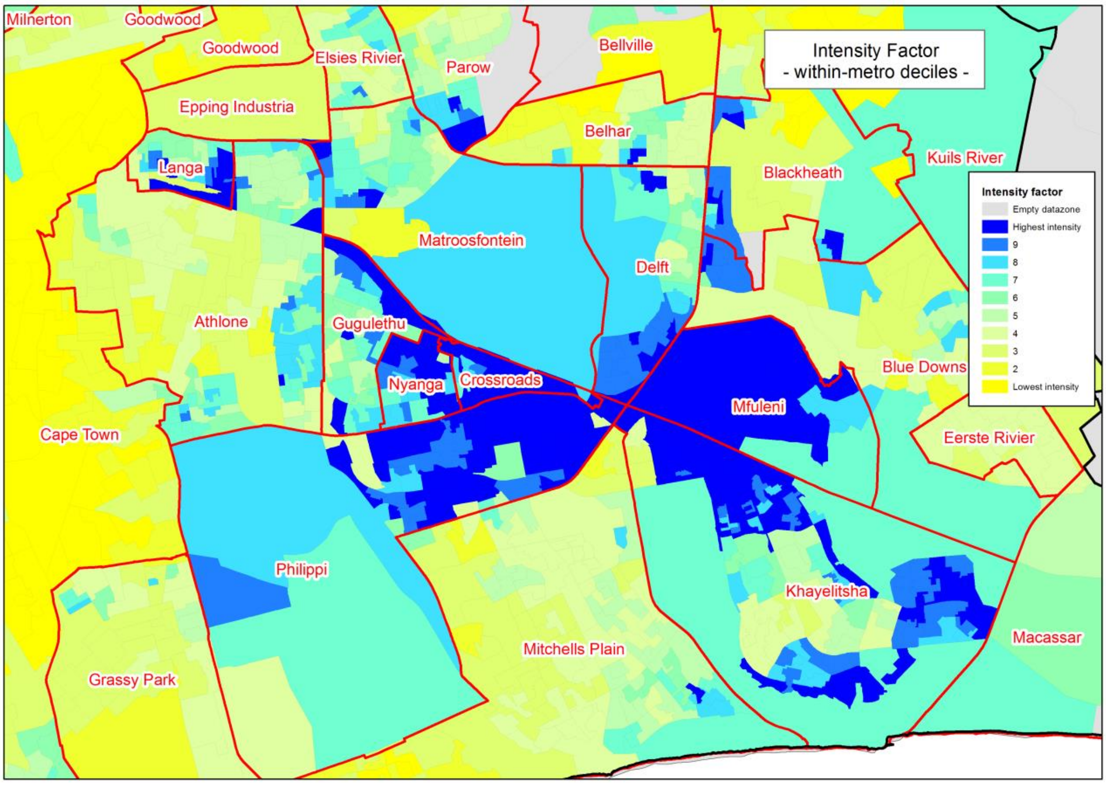


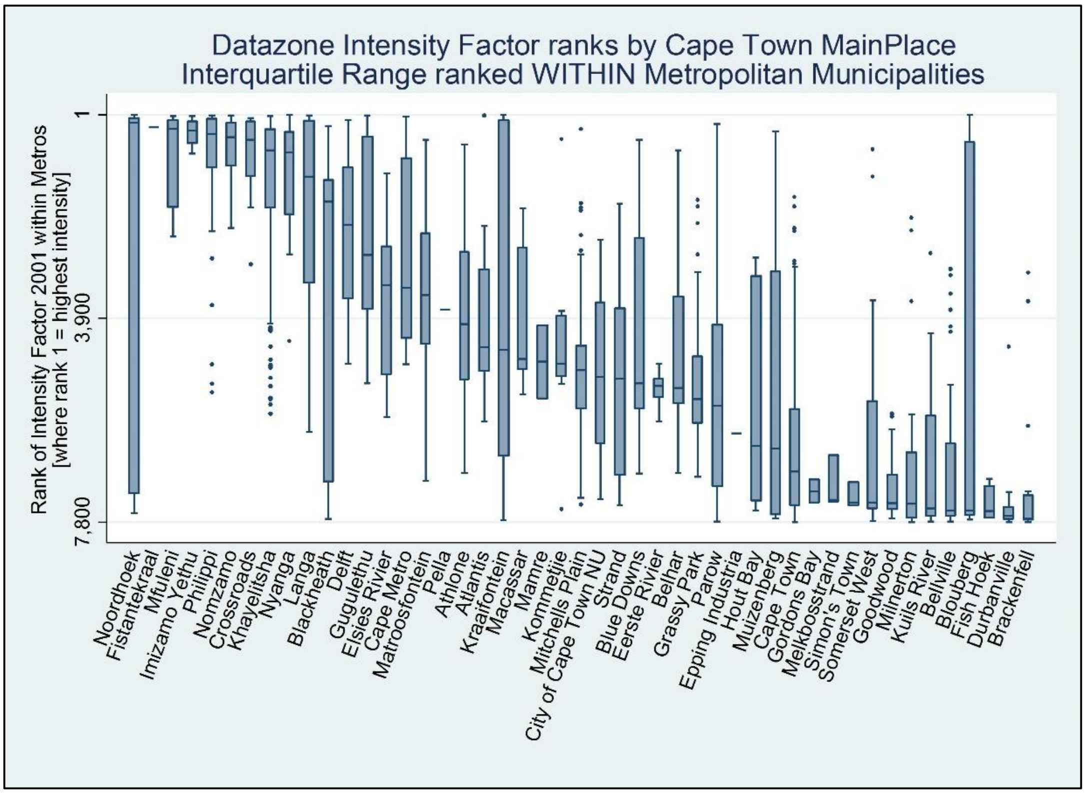
Figure 6.18 provides a close-up for the areas running along the N2 corridor. High intensity neighbourhoods can be seen in Langa, Gugulethu, Nyanga, Crossroads, Philippi, Delft, Mfuleni, and Khayelitsha. In this particular map, the 2011 Census Main Place boundaries are shown in red, and the Main Places are labelled. Within Khayelitsha, the areas of highest intensity can be seen to be located in the northern part of the township close to the N2 highway, and also in the southern and eastern part of the township. The more central part of the township can be seen to exhibit somewhat lower levels of intensity than the northern, southern and eastern parts.

Figure 6.18: Datazone level Intensity Factor within-metro deciles with focus on the Cape Flats



Graphically, the distribution of datazone level intensity ranks across Cape Town Main Places is provided in Figure 6.19, where again the Main Places are ranked from highest median datazone rank on the left-hand side, through to lowest median datazone rank on the right-hand side. As was noted in the equivalent discussions of deprivation rates and exposure scores in the sections above, some Main Places contain only a few datazones (sometimes only a single datazone), whereas others (such as Khayelitsha) contain many datazones. For this reason, the comparison between Main Places can only give an indicative sense of the variation in intensity.

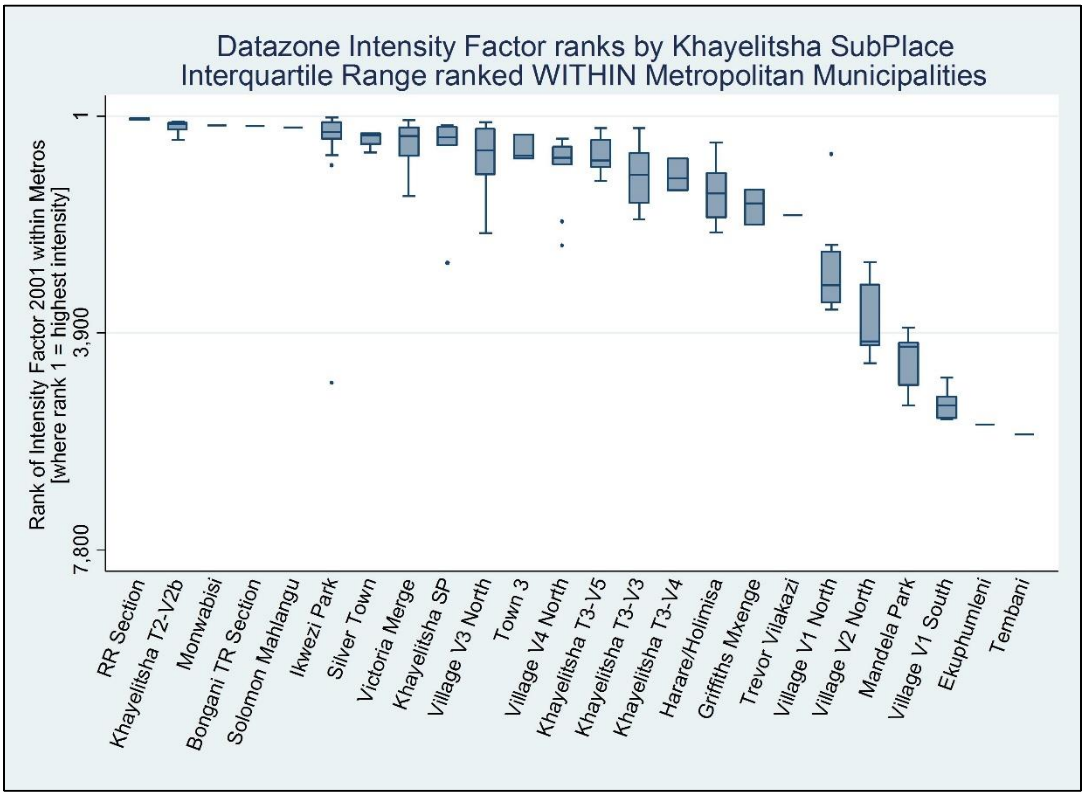
Figure 6.19: Datazone Intensity Factor ranks by Cape Town MainPlace (ranked across metros only)



The Main Place with the highest median datazone intensity score in Cape Town is Noordhoek which is located on the peninsular, just to the south of Table Mountain. The range of datazone intensity scores in Noordhoek can be seen to be particularly wide. This is due to the presence within the Noordhoek Main Place of both affluent neighbourhoods, such as Sunnydale, and also very deprived neighbourhoods, specifically, parts of the Masiphumelele township. Levels of exposure to inequality amongst the poor population are relatively high in both the affluent and deprived parts of the Main Place, resulting in levels of intensity being high in places such as Masiphumelele but low in places such as Sunnydale. In contrast, at the lower intensity side of the chart, the Main Place of Fish Hoek can be seen to cover only datazones that rank at lower intensity end of the within-metro distribution. Indeed, all of the datazones within the Fish Hoek Main Place are in the lowest intensity decile of metropolitan datazones. This particular finding is due to Fish Hoek being a relatively affluent area with no major pockets of deprivation.

Finally, Figure 6.20 focuses down further to explore indicative patterns of intensity between the constituent Sub Places within Khayelitsha. Again, as noted above, datazones do not nest perfectly within Sub Places (or indeed Main Places), and so the data presented here should only be used to give a general sense of the variation between Sub Places. It is possible to conclude, however, based upon a combination of the thematic maps and the boxplots, that, for example, residents of the neighbourhood of Silver Town in Khayelitsha experience a higher intensity of exposure to inequality than residents of, say, Mandela Park.

Figure 6.20: Datazone Intensity Factor ranks by Cape Town Main Place (ranked across metros only)



6.6 Conclusion

The aim in this chapter was to address research SubQ4: Are there any neighbourhoods across South Africa with high rates of deprivation and high lived experience of inequality? I was able to show, using the initial four-category classification of 'high deprivation – low exposure', ''low deprivation – low exposure', 'low deprivation – high exposure' and 'high deprivation – high exposure', that there are indeed some datazones that are characterised by both high levels of deprivation and high levels of exposure to inequality. The measures of community 'intensity' of exposure to socio-economic inequality that I proceeded to develop and present provide an indicator of those neighbourhoods that are subject to the twin stressors of high poverty rates and high exposure to socio-economic inequality. These neighbourhoods may be the locations where the attitudes of the poor are likely to be most clearly sharpened towards inequality and other social issues concerning social justice. Furthermore, it could be hypothesised that these areas may potentially be most vulnerable to social unrest or violent crime.

The analyses of the new intensity measure (based on the national distribution of intensity scores computed using the INC domain as the basis), revealed that the highest intensity datazones were overwhelmingly (but not entirely) located within the major metropolitan municipalities. These high intensity areas were often seen to be adjoining or in otherwise close proximity to very low intensity neighbourhoods, indicating the importance of adopting a detailed neighbourhood level approach.

Relatively high correlations were observed between the four dimension-specific measures of intensity developed here which (as for the exposure measures discussed above) offers an empirical justification in support of the conceptual motivation for combining the measures into a single composite indicator using factor analysis. The derived Intensity Factor measure was analysed over the metropolitan datazones and revealed a broad spread of datazone level scores within each of the eight metropolitan municipalities. Much less variation was seen *between* the metropolitan municipalities on the intensity measure than on the exposure measure, which is as one would expect given the methodological features of each measure.

Cape Town was seen to contain a disproportionately high number of datazones that fall within both the highest ten percent and the lowest ten percent of metropolitan datazones on the Intensity Factor measure. Those Cape Town datazones that fall within the highest ten percent category are largely located along the N2 corridor in areas such as Langa, Nyanga, Gugulethu, Crossroads, Philippi and Khayelitsha, as well as certain outer-lying townships such as Masiphumelele. In this respect, Cape Town is clearly a city of contrasts, with considerable differences in levels of community intensity of exposure between its constituent neighbourhoods.

Having analysed spatial patterns of deprivation in Chapter 3, exposure in Chapter 5 and intensity in Chapter 6, I will now turn in Chapter 7 to explore whether any or all of these three empirical measures is statistically associated with attitudinal outcomes concerning inequality and options for redistributive redress.

Chapter 7: The relationship between exposure to inequality and attitudes to inequality

7.1 Introduction

In this chapter I present an example of how my new spatial measures of the lived experience of inequality (developed in Chapter 4 and analysed in Chapters 5 and 6) may be used to better understand other social phenomena. As I argue in Chapters 1 and 2, there is a strong *conceptual* basis for contending that the lived experience of inequality may influence people's social attitudes and indeed their action responses, such as their propensity to engage in acts of social unrest or acts of crime (Wikstrom, 2006a). However, the lack of suitable empirical measures of the lived experience of inequality at a detailed geographical level has so far prevented researchers from fully testing these potential relationships.

In this chapter, I seek to establish whether people's lived experience of inequality is associated with their attitudes to inequality and options for redress. Gaining a better understanding of the public's views about societal inequalities, the role of government in securing equality of opportunity and outcome, and appropriate social policies is essential given the country's context of historic social divisions and persisting spatial disparities. An appreciation of the nature, dynamics and complexities of people's views towards redistributive justice will help policymakers navigate the challenging politics typically associated with designing and implementing sustainable developmental policies (Graham, 2002). In addition, monitoring people's support for different policy options is a crucial element in ensuring that policy-making is suitably evidence-based, accountable and adaptive. The specific research question I address in this chapter is therefore Sub Q5: To what extent are people's attitudes to inequality and options for redress associated with their lived experience of inequality?

In separate work, I have also explored potential associations between exposure and rates of violent crime[38]. However, any analyses of South Africa crime rates is subject to the (very valid) criticisms of crime under-reporting and under-recording, which I acknowledge in my other work. In this thesis I only wish to present one example of the application of my exposure measures to other social outcomes and I therefore rejected the crime rate option due to the issues of under-reporting and under-recording of crime.

It is important to reiterate again that the analyses presented here are just one example of how my exposure measures could be used and, indeed, I deliberately do not claim to establish causal relationships between my spatial measures of exposure and attitudinal responses concerning inequality and/or redress. As I will explain further below, the analyses here are focused on establishing whether there are any *statistical associations* between exposure and attitudes as this is an important first step in this field of inquiry.

Before I introduce and discuss the data, methods and results from the analyses, I will first briefly acknowledge the early work in the developed world that led to the formulation of various social attitude theories relating to inequality. Such theories include those relating to people's belief in a just world (e.g. Lerner, 1965), inequality

[38] See http://www.idrc.ca/EN/AboutUs/Donor_Partnerships/saic/Pages/107365.aspx for details of a collaborative research project between the Human Sciences Research Council, University of Oxford and University of Rio de Janeiro. This project is part of the Safe and Inclusive Cities programme headed by the Canadian IDRC (with contributions from UK DfID). As part of this project I developed multilevel models to test for associations between rates of violent crime at police precinct level and a range of possible crime generating/facilitating or preventing factors in South Africa. I included deprivation rates and aggregates of my exposure and intensity measures within the set of independent variables. This research will be published as a chapter of a book dedicated to the project findings (due for publication by HSRC Press).

aversion (e.g. Shorrocks, 1980), system justification (e.g. Jost and Banaji, 1994), and also conservatism (e.g. Wilson and Patterson, 1968). Whilst this field of research has continued to expand rapidly over time in developed countries, covering attitudes towards social security, taxation and a raft of redistributive policies (e.g. Hills, 2004; Kluegel and Smith, 1986; Orton and Rowlinson, 2007; Piketty, 1995), there has been far less research undertaken in this field in developing countries (Roberts, 2014). Kalati & Manor (1999; 2005) focused on a group of 'elites' in South Africa in the late 1990s and examined attitudes within this particular social group towards notions of poverty and social solidarity. Roberts (2006) used the nationally representative South African Social Attitudes Survey (SASAS) to examine issues including inequality aversion, relative deprivation and levels of support for a raft of different redistributive social policies, while Noble et al (2008) and Surender et al (2010) also used the SASAS data resource to look explicitly at social attitudes towards social grants.

More recently, and as I will refer to further in the discussion below, Roberts (2014) used a time series of SASAS data to examine patterns and trends in social attitudes on inequality aversion, perceived class tensions, perceived pay differentials, support for government's role in income redistribution and support for various redistributive polices, and included analyses of differential responses across a variety of social, economic, demographic and political sub-groups. Roberts concluded that there is strong evidence that South Africans are acutely aware of the high levels of inequality in their society and they find this concerning. He argued that there is strong and consistent evidence that the majority of the population believe the gap in incomes between the highest earners and lowest earners is too large and there is a consistent preference for a more equal social structure. He demonstrated that the population is acutely aware of tensions between social classes and between racial groups, both of which relate to inequality in a historical as well as contemporary sense. In terms of wage disparities, Roberts's analyses show a general preference for lower paid workers to be paid more than at present, and higher paid workers to be paid less than at present, thus reducing the wage differentials between the lowest and highest paid workers. With regards to the role of the state, Roberts found a strong consensus that the government has a responsibility to reduce the differences in income between people with high incomes and those with low incomes. Class- and race-based gradients of support were observed in respect of government role in redistribution, although even amongst the better-off sub-group there was a majority in support of this. In terms of specific redistributive policies, Roberts found greater a consensus of support across population sub-groups for class-based polices (e.g. support for the unemployed) than for race-based policies (e.g. land reform). Roberts's sub-group analyses showed, for instance, that responses varied according to demographic factors, educational level, employment status, and self-perceived social status. However, when taken as a whole, the results suggest a common identification of and concern with social injustice and a broad-based commitment to government redistribution and class-based social policies.

In this chapter I refer explicitly to the work of Roberts (2014) which formed part of the deliverables from the ESRC/NRF funded research project upon which this thesis draws. Roberts's analyses were designed to set the scene for future analytical work, including multivariate statistical analysis. The aim in this chapter is to build upon his work to develop statistical models to test for associations between my new exposure measures and two selected dependent variables – one relating to inequality aversion and one relating to support for government redistribution – whilst controlling for relevant independent factors, including those identified by Roberts in his sub-group analyses.

Having stated the purpose of this chapter, I will now proceed to introduce the data and methods used and then present the key empirical findings. More specifically, I begin by introducing the South African Social Attitudes Survey (SASAS) which provides information on, amongst other things, people's attitudes towards inequality and redress. I then discuss the model development strategy in the context of the two input datasets (the individual level SASAS dataset and the datazone level exposure measures). I provide an overview of the data preparation methodology and, finally, I present the results and discuss the findings.

The explicit focus of this chapter is on assessing whether there is any evidence of statistical association between my measures of exposure to inequality and the selected attitudinal outcomes concerning inequality and redress. In light of this, in the discussion of the results I place explicit emphasis on the magnitude and direction of these particular associations. Whilst I do comment briefly on the other statistically significant results concerning the control variables, these are presented as background context only and I therefore do not go into any detail on these.

With regard to the choice of exposure measure for the analyses, I utilise only the original INC-based measures so I do not use the EMP, EDU or LIV based measures, nor the composite measures developed in Chapters 5 and 6. I give further details on the reasons for this in the discussion below.

Finally, as discussed above in Chapters 4, 5 and 6, and as I discuss further below, the main interest is on the experiences of the poor population of South Africa and so the results presented in this chapter relate only to the poor population. I have, however, also undertaken equivalent analysis for the non-poor population and these results and commentary are provided in Appendix E.

7.2 South African Social Attitudes Survey

In order to examine societal attitudes to inequality and redress policy, I utilised the South African Social Attitudes Survey (SASAS). This is a nationally representative social survey of people aged 16 and older which has been run annually by the Human Sciences Research Council (HSRC) since 2003. It is similar in design to the British Social Attitudes Survey. The full SASAS questionnaire is translated into seven of the eleven official languages and goes into the field in October and November each year, having first been approved by the HSRC's ethics committee. The target population for the survey is individuals aged 16 and over, excluding those living in special institutions such as hospitals or prisons. As SASAS is part of the International Social Survey Programme (ISSP) community, it carries some common modules included in international social attitudes surveys. Of particular relevance for this thesis is the Social Inequality IV Module, which was included in the seventh SASAS round in late 2009. Appendix E provides some technical notes on SASAS for background context, including geographical stratification. The sampling approach is designed, in part, to avoid geographical bias. However, as with all attitudes surveys, it is important to be mindful of human-induced bias because respondents' personal interpretations of survey questions on attitudes are intrinsically subjective.

The analyses presented here draw directly upon the spatial inequality measures developed and analysed in Chapters 4, 5 and 6 and combine these spatial data with individual data on social attitudes from SASAS.

SASAS contains a number of different questions that relate either directly or indirectly to people's attitudes towards and perceptions of inequality in South Africa. In this chapter I focus on two particular survey questions and develop statistical models to test whether people's responses to these questions are influenced by their lived experience of inequality. The two survey questions I examine as the dependent variables in the statistical models are:

> Q186: "To what extent do you agree or disagree that differences in income in South Africa are too large?"
>
> Q187: "To what extent do you agree or disagree that it is the responsibility of the government to reduce the differences in income between people with high incomes and those with low incomes?"

The first question, which is about inequality aversion, was selected as a dependent variable because it captures, in a broad sense, people's perceptions about societal inequality. The question on government role in income redistribution was selected as a dependent variable because it relates directly to potential responses to tackle inequality.

Roberts (2014) includes summary results tables for both questions and they are replicated here as Table 7.1 and Table 7.2. With reference to Table 7.1, it is evident that when the SASAS survey data are pooled over the ten consecutive years from 2003 to 2012, an average of 85% of the population either agree or strongly agree that income differences in South Africa are too large. Furthermore, whilst there is a degree of variation in the figures year-on-year, the general pattern remains fairly constant over the time period. As Roberts concluded, these data give a clear indication of the South African population's aversion to income inequality.

Table 7.1: Aversion to income inequality in South Africa, 2003-2012

"Income differences in South Africa are too large"	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**All years**
Strongly agree	38	39	36	41	43	40	41	38	41	46	40
Agree	44	42	38	42	44	45	48	50	45	44	44
Neutral	6	7	8	5	7	7	4	5	6	5	6
Disagree	4	5	9	6	3	3	2	4	3	3	4
Strongly disagree	2	4	4	1	1	1	0	0	1	0	1
(Do not know)	5	3	6	5	3	5	4	3	4	1	4
Total	100	100	100	100	100	100	100	100	100	100	100
% strongly agree / agree	83	81	73	84	87	84	90	89	86	91	85
Mean score *	4.12	4.06	3.92	4.18	4.26	4.20	4.28	4.22	4.21	4.33	4.19
Base N	2462	2781	2844	2934	3164	3310	3284	3155	3016	2473	29423

Source: Roberts (2014) using data from SASAS over the 2003-2012 period
Note: * The mean scores are based on a reversed scale, where 1=strongly disagree and 5=strongly agree. 'Do not know' responses were combined with neutral responses.

Table 7.2: Attitudes towards government redistribution of income

"Responsibility of the government to reduce the differences in income between people with high incomes and those with low incomes"	2006	2007	2008	2009	2010	2011	2012	All years
Strongly agree	35	31	28	26	26	23	26	28
Agree	43	39	37	44	45	38	40	41
Neutral	7	12	13	12	11	15	15	12
Disagree	9	11	13	12	13	14	13	12
Strongly disagree	2	4	4	3	4	6	4	4
(Do not know)	3	4	5	2	2	4	2	3
Total	100	100	100	100	100	100	100	100
% strongly agree / agree	78	70	65	70	71	61	66	69
Mean score *	4.00	3.83	3.73	3.78	3.76	3.58	3.70	3.77
Base N	2934	3162	3307	3286	3156	3014	2471	21330

Source: Roberts (2014) using data from SASAS over the 2006-2012 period
Note: * The mean scores are based on a reversed scale, where 1=strongly disagree and 5=strongly agree, such that higher values represent greater support for government-led redistribution. 'Do not know' responses were combined with neutral responses.

With reference to Table 7.2, it is evident that the majority of the population agrees or strongly agrees that it is the responsibility of the government to reduce differences in income between people with high incomes and those on low incomes. Based upon the pooled data over the seven year period between 2006 and 2012, an average of 69% of the population agreed or strongly agreed with this statement. However, as Roberts notes, whereas the levels of inequality aversion seen in Table 7.1 remain fairly constant over the time period, we see from Table 7.2 that there is a notable decline in support for the view that the government has a responsibility to reduce income differences. Despite this decline, the data for 2012 shows that a clear majority still believe the government does have such a responsibility to reduce income differences. Roberts (2014) does not explore these temporal shifts further and, indeed, this is also outside the scope of this thesis.

Although I have replicated Roberts's two tables above to show year-on-year consistency and/or change, as I will explain below, the focus here is explicitly on the 2009 round of SASAS which contained the ISSP module on social inequality. The temporal patterns are shown here only for context.

The SASAS data are nationally representative and, as such, cover people across the entire income distribution. As described in detail in Chapter 4, the final measures of lived experience of inequality developed for this thesis consist of two complementary variants: the exposure of the poor to the non-poor; and the exposure of the non-poor to the poor. The asymmetric nature of the exposure indices allows me to explore associations with attitudinal responses from the perspective of a poor person and separately from the perspective of a non-poor person. Four separate statistical models were therefore developed:

Table 7.3: The four models developed

		Poor/non-poor perspective	
		poor	non-poor
Dependent variable	*ineqavr*	Model A	Model B
	govredr	Model C	Model D

The focus within this Chapter is on the two models that relate to the experiences and attitudes of the poor subgroup. As evidenced above in Chapter 3 and further discussed below, the majority of the total population of South Africa is poor (based upon a range of the most commonly applied poverty thresholds – see e.g. Leibbrandt et al (2010); Noble et al (2009b)). Furthermore, the role that a person's experience of inequality may play in determining their action responses (including, for example, participation in protest action or criminal activity) needs to be considered in the context of the person's experience of poverty and deprivation. This is not to underestimate the importance of understanding the experiences and attitudes of the non-poor, especially as these non-poor people will typically include those in political power, but rather an acknowledgement of the additional stresses that inequality may present when combined with poverty. The focus of the interpretation in this chapter is therefore on the poor models (A & C). A full set of results is also presented for the non-poor models (B & D) in Appendix E.

7.3 Data preparation methodology

In order to test whether people's individual attitudes towards inequality and redress are associated with their lived experience of inequality it was necessary to perform three main data preparation steps. The first step involved attaching datazone level measures of exposure to inequality (and datazone level deprivation rates) to the individual survey respondent-level SASAS dataset. The second step involved reviewing the SASAS questionnaire and retaining only those survey questions (plus the merged datazone level exposure and deprivation measures) that were regarded as being potentially important control variables. The third step involved splitting the retained SASAS dataset into the poor/non-poor subset to enable appropriately specified models to be run for both groups.

7.3.1 Methodology for merging area-level data into SASAS

Attaching area-level variables to individual-level survey records requires details of the geographical home location of each survey respondent. As discussed in Appendix H, the sampling methodology underpinning the SASAS data collection utilised the Enumeration Area (EA) geography in the stratification process and so each survey respondent has an EA code. In order to preserve respondent confidentiality, the EA codes are excluded from the standard SASAS datasets that are made available to researchers. However, for the purpose of the ESRC-funded research project upon which this thesis draws, special dispensation was granted by the HSRC to enable the respondents' home EA codes to be included in the base survey dataset. The EA code represents the key data linkage variable through which the area-level variables were attached to the individual-level survey records.

As discussed in Chapter 3, the datazone geography was generated by combining EAs in such a way as to maximise adherence to a number of rules, including population size thresholds and population homogeneity measures (Avenell et al., 2009; Openshaw and Baxter, 1977). As such, EAs nest perfectly within datazones and so it was possible to generate an EA-to-datazone lookup table. This EA-to-datazone lookup table was matched into the individual-level survey response records using the EA code as the common link variable.

Once the datazone code was successfully attached to the survey records, it was then possible to merge in the datazone level exposure and deprivation measures using the datazone code as the common link variable. Upon completion of this matching process the EA code was deleted from the matched dataset to preserve respondents' confidentiality.

The matched dataset therefore consisted of all the individual-level survey variables plus the home datazone code and datazone-level exposure to inequality and poverty measures.

I acknowledge here an important limitation of this modelling exercise, in that the datazone level exposure and deprivation scores that are attached to each individual respondent relate to the respondents' home addresses as at the time of the 2009 SASAS data collection fieldwork. As such, I am unable to take account of people's exposure to inequality (or their neighbourhood deprivation rates) in places they may have lived prior to the SASAS survey date. This limitation is discussed further in Chapter 8 and I offer some recommendations for how this may be improved upon in future research.

7.3.2 Selecting the control variables

The 2009 SASAS questionnaire contained a large number of survey questions. Only a minority of these questions related to issues around inequality (as the survey also contained various other modules on topics such as experience of crime and fear of crime). In order to make the modelling process manageable, I (together with Ben Roberts and Hope Magidimisha from the HSRC – please see Appendix A for further details on joint work in this regard) undertook a review of the entire SASAS questionnaire and identified a subset of questions deemed to be particularly important for this purpose. The objective here was to ensure that the models controlled for as many measureable explanatory factors as possible in order to maximise confidence in any observed associations between the exposure measures and the response variables. The control variables were selected to reflect known correlates of the dependent variables (see reference to Roberts (2014) above). Table 7.4 lists the SASAS variables that were selected for inclusion in the base modelling dataset.

Table 7.4: Variables selected for base modelling file

	Variable name	Var. type	Variable description
ID / wt	uniqueid	-----	Individual survey respondent unique identifier code
	benchwgt	-----	Composite survey weight
Dependent variables	ineqavr	Ordinal	Q186. To what extent do you agree or disagree that differences in income in South Africa are too large?
	govredr	Ordinal	Q187. To what extent do you agree or disagree that it is the responsibility of the government to reduce the differences in income between people with high incomes and those with low incomes?
Demographics	age	Numerical	Q238. Age of respondent in completed years.
	agesq	Numerical	Derived 'Aged squared' indicator, based on Q238
	race	Categorical	Respondent's population group (taken from 'respondent selection procedure' questions).
	marstat	Categorical	Q239. What is your current marital status?
	hhper	Numerical	Number of persons in this household (taken from 'respondent selection procedure' questions)
Assets	assetindex	Numerical	Derived 'Asset Index' indicator, based on 25 separate items (Q267-Q300): e.g. Q281. Does your household have a washing machine (in working order)?
Social status	edu	Categorical	Q242. What is the highest level of education that you have ever completed?
	empl	Categorical	Q246. What is your current employment status?

Self-perceived social status and trajectories	spoor	Ordinal	Q151. Would you say that you and your family are… e.g. 'wealthy/very comfortable'
	topbott00	Numerical	Derived indictor (transformed to 0-100 scale) based on Q198. In our society there are groups which tend to be towards the top and groups which tend to be towards the bottom. Where would you put yourself on a scale of 1 to 10, where 10 is the top and 1 the bottom?
	ssocmobc	Ordinal	Q2. In the last 5 years, has life improved, stayed the same or gotten worse for people like you?
	futmob	Ordinal	Q3. Do you think that life will improve, stay the same or get worse in the next 5 years for people like you?
	jobprest	Ordinal	Q200. Please think about your present job (or your last one if you don't have one now). If you compare this job to the job your father had when you were 15, would you say that the level of status of your job is (or was)… e.g. 'Much higher than your father's'
Conflict / discrimination	classconind	Numerical	Derived 'Class conflict index' indicator, based on multiple separate items, e.g. Q195. In your opinion, in South Africa how much conflict is there between the working class and the middle class?
	groupdis	Ordinal	Q59. Would you describe yourself as being a member of a group that is discriminated against in this country?

Getting ahead in life	meritind	Numerical	Derived ‘Merit factor index’, based on multiple separate items, e.g. Q164. How important is hard work for getting ahead in life?
	exogind	Numerical	Derived ‘Exogenous factors index’, based on multiple separate items, e.g. 160. How important is coming from a wealthy family for getting ahead in life?
	q68r	Ordinal	Q168. How important is a person’s race for getting ahead in life?
Political views	polideol	Categorical	Q235. In political matters, people talk of ‘the left’ and ‘the right’ or ‘liberal’ and ‘conservative’. Where would you place your views on this scale?
	anc	Ordinal	Derived ‘ANC voter’ indicator based on Q231. If there were a national election tomorrow, for which party would you vote?
Geographic identifiers	geotype	Categorical	Statistics South Africa Census 2001 enumeration area Geotype classification
	dz_code	-----	Datazone unique identifier code
	mun_name	-----	Municipality name
	prov	-----	Province name
Datazone level exposure and deprivation scores	exposure_of_poor	Ratio	aLDPxyi* exposure to inequality measure developed above in Chapter 2
	exposure_of_rich	Ratio	aLDPyxi* exposure to inequality measure developed above in Chapter 2
	inc	Ratio	Proportion of population that is classified as being deprived on the ‘Income and Material Deprivation Domain’ of the SAIMD 2001 at datazone level

7.3.3 Splitting the file into poor and non-poor

As discussed in Chapter 4, two separate but complementary measures of the 'lived experience of inequality' were developed and analysed: one relating to the experience of the poor, and the other relating to the experience of the non-poor. In order to test whether a person's experience of inequality is associated with their attitudes towards inequality and options for redress, it is important to include the inequality measure that corresponds to the person's own poor/non-poor status. In short, the measure of inequality experienced by the poor is only relevant for people who would be classified as poor, whilst the measure of inequality experienced by the non-poor is only relevant for the people who would be classified as non-poor. In order to achieve this data and model configuration it was therefore necessary to split the SASAS dataset into two distinct subsets: one consisting of respondents defined as poor and one consisting of respondents defined as non-poor.

For internal consistency within the thesis, I refer to data derived from the Income and Material Deprivation from the SAIMD 2001 in order to establish what proportion of the population is deprived and what proportion is not deprived. A total of 73% of the adult population of South Africa is defined as poor on this measure, with the remaining 27% defined as non-poor (see also: Noble et al., 2009a). This 73%-27% split is therefore adopted for the purposes of the modelling in this chapter as a means of determining what proportion of the SASAS dataset should be classed as poor and what proportion classed as non-poor.

Although SASAS does contain a question on 'income', a considerable proportion of respondents either declined to provide an answer or stated that they did not know their income. This resulted in a substantial amount of missing data in the income variable of the SASAS dataset and therefore it was not possible to use the SASAS income variable as the basis for classifying people as poor/non-poor. However, SASAS also contains a series of questions relating to people's material asset ownership which, for the purpose of this thesis, I combined together to form a composite asset index. The overall SASAS dataset was then sorted according to the respondents' score on the asset index variable, resulting in a ranking from the person with the lowest material asset ownership to the respondent with the highest material asset ownership. The SASAS dataset was then split into a 'poor' subset and a 'non-poor' subset, with the poor subset consisting of the 73% of respondents with the lowest scores on the asset index, and the non-poor subset consisting of the 27% of respondents with the highest scores on the asset index.

The poor subset contained a total of 2120 cases, while the non-poor subset contained a total of 1036 cases[39]. In the model development stage discussed below, the base dataset for the poor subset contained the measure of inequality experienced by the poor (namely, the '*exposure_of_poor*' variable listed in Table 7.4 above), whilst the base dataset for the non-poor subset contained the measure of inequality experienced by the non-poor (namely, the '*exposure_of_rich*' variable listed in Table 7.4 above). Apart from the differential specification of 'experience of inequality', the composition of variables in the poor and non-poor base datasets was identical.

[39] Note that the 73%/27% threshold was imposed on the respondents using the respondents' survey weights as the basis for the cumulative population sum.

7.4 Model development

Table 7.3 above showed how the four models developed here consisted of focusing on two dependent variables and examining these from the separate perspectives of the poor and non-poor. Models A and B focus on people's attitudes towards inequality, and utilise the '*ineqavr*' variable as the dependent response variable. Models C and D focus on people's views concerning the government's role in income redistribution, and utilise the '*govredr*' variable as the dependent response variable.

As shown in Table 7.1 and Table 7.2, both '*ineqavr*' and '*govredr*' are recorded as ordinal data, with people's responses to these two variables being captured on five-point scales relating to their stated support for or agreement with relevant survey statements.

There are typically three ways in which dependent variables of an ordinal nature are treated in statistical modelling: (i) assume that the magnitude of the differences between response options are equal and therefore treat the ordinal variable as if it were a continuous variable i.e. standard OLS regression; (ii) identify a theoretically-based or empirically-based dichotomous distinction within the ordinal scale and therefore split the possible response options into two discrete groups, i.e. logistic regression; and (iii) respect the ordinal nature of the response variable in full, i.e. ordered logit regression.[40]

In order to ensure I adopted the most appropriate modelling approach, I consulted extensively with academics from the University of Oxford's Department of Statistics (namely, Dr Dan Lunn and Dr George Nicholson). Two major recommendations emerged from those discussions: (i) my modelling approach should

[40] More specifically, fitting a 'cumulative link model'.

adopt the full ordered logit model[41]; and (ii) due to the geographical stratification inherent in the SASAS sampling methodology, multilevel models were required.

The final model development framework therefore required the specification of multilevel cumulative link ordered logit regression models. Whereas the spatial inequality analyses presented in Chapter 5 and 6 were performed using Stata, this was not possible for the modelling undertaken here in Chapter 7. The R statistical software programme was used to implement multilevel ordered logit models.

With the support of colleagues from the University of Oxford Statistics Department[42], I fitted four base models - Models A, B, C, and D. These base models each contained the full complement of explanatory variables listed in Table 7.4 above. The multilevel component of the models consisted of three levels: (i) individual respondent; (ii) datazone of residence; and (iii) municipality of residence.[43] Individual respondents nest within datazones, and datazones nest within municipalities. Datazones and municipalities were specified as levels 2 and 3 in the model in recognition of the SASAS sampling methodology (see Appendix H for further details). Ordinal and categorical variables were treated as factors, while numeric and ratio variables were standardised by subtracting the mean and dividing by the standard deviation.

Before I describe the technical approach to developing the models, it is important to state again that the key objective was to test whether there were any statistical associations between my exposure measures (and/or the neighbourhood deprivation rate) and the two dependent variables. The various other independent

[41] A problem with treating ordinal dependent variables as if they are continuous and fitting OLS models is that the residuals need to be normal and homoscedastic for the model to have validity. If the raw category numbers are used this will not be the case and the traditional approach was therefore to transform them to the normal scores associated with the appropriate quantiles. However, this has fallen into disuse since the advent of Generalised linear Models.

[42] See Appendix A

[43] Note that Province code was added as a dummy variable into the set of covariates and was therefore not entered as a hierarchical level in the same way as datazones and municipalities were.

variables in the initial starting pool were included only as controls and are not of substantive interest for this thesis. This explicit focus on the datazone level independent variables was an important factor in the choice of modelling approach which I adopted.

Starting from the models containing the full complement of selected explanatory variables, the approach adopted was to sequentially drop one variable at a time based upon the likelihood ratio (LR) test p-value until each variable remaining in the model had an LR p-value < 0.05 (see Greenland (1989) for a detailed discussion of this methodology). The only exception to this LR test p-value rule was that the explanatory variables relating to experience of inequality (‘*exposure_of_poor*’ or ‘*exposure_of_rich*’) and neighbourhood poverty rate (‘*inc*’) and the interaction term between them would be retained in the model regardless of LR test p-value. This exception was justified on the grounds that the primary objective of the modelling exercise was to assess whether experience of inequality and/or neighbourhood poverty rate played a role in influencing people’s attitudes, and whether there was any interaction between these two variables resulting in exposure to inequality accentuating the effects of poverty, and vice versa. The interaction term included in the models here is therefore equivalent to the ‘intensity’ measure that was the focus of the empirical analysis in Chapter 6 of this thesis (because ‘intensity’ was calculated as the simple mathematical product of deprivation rate and exposure score).

The sequential process entailed: (i) identifying the explanatory variable with the largest LR test p-value; (ii) fitting models with and without this particular explanatory variable; and (iii) comparing the models with/without the explanatory variable using a LR test based upon nested models fitted to the same input record specification (i.e. if there were missing values in the dropped variable, then both models were fitted to the data without the corresponding missing-value rows of the dropped variable). The explanatory variable was dropped if it had an LR test p-value > 0.05 and its LR test p-value was greater than those of all other included variables. When the variable identified as having the highest LR test p-value was categorical in nature (e.g. employment status), the LR test comparison was performed between a model with all relevant categorical levels included versus a model with none of the relevant categorical levels included.

Once this process of sequential variable reduction was complete for each model (i.e. Models A, B, C and D), attention turned to testing for any evidence of interaction effects between our explanatory variables of primary interest (experience of inequality and neighbourhood poverty rate) and other explanatory variables that remained in the model specifications at the conclusion of the sequential reduction process. Here, interaction terms were added into the model specification if the associated LR test p-value for the interaction was <0.05.

This entire model development process was developed and implemented in the R statistical software package using the 'clmm' function from the ordinal library. The final model specifications were as follows:

Model_A <- clmm(ineqavr ~ exposP + log_incS + (log_incS:exposP) + edu + hhperS + spoor + groupdis + ssocmobc + jobprest + futmob + classconindS + meritindS + exogindS + q168rS + anc + prov + (exposP:ssocmobc) + (exposP:futmob) + (log_incS:hhperS) + (1|mun_name/dz_code), weights = 2120*wt, data = poor_i, Hess = T, na.action = na.omit)

Model_B <- clmm(ineqavr ~ exposR + log_incS + (log_incS:exposR) + race + edu + empl + spoor + groupdis + ssocmobc + jobprest + classconindS + topbott100S + prov + (exposR:race) + (exposR:spoor) + (log_incS:ssocmobc) + (1|mun_name/dz_code), weights = 036*wt, data = rich_i, Hess = T, na.action = na.omit)

Model_C <- clmm(govredr ~ exposP + log_incS + (log_incS:exposP) + empl + spoor + groupdis + jobprest + classconindS + meritindS + exogindS + q168rS + polideol + ineqavrS + (exposP:empl) + (exposP:spoor) + (1 | mun_name/dz_code), weights = 2120*wt, data = poor_g, Hess = T, na.action = na.omit)

Model_D <- clmm(govredr ~ exposR + log_incS + (log_incS:exposR) + marstat + ssocmobc + futmob + anc + polideol + assetindexS + ineqavrS + (exposR:polideol) + (1|mun_name/dz_code), weights = 1036*wt, data = rich_g, Hess = T, na.action = na.omit)

where the dependent and independent variables are as described in Table 7.4 above (with the adapted notation of suffixing continuous variables with an upper-case 'S' to signify that the variable has been standardised, and renaming the variable '*exposure_of_poor*' to '*exposP*' and the variable '*exposure_of_rich*' to '*exposR*').

As explained above, the final model specifications vary in terms of the constituent control variables because the explicit interest in this chapter is on testing for associations between the datazone level independent variables and the two dependent variables and, as such, the harmonisation of the composition of the control variable set is not of interest across the different models.

7.5 Results and discussion

The main focus of this Chapter of the thesis is on testing for associations between exposure to inequality and attitudes to inequality and redress. As such, I start by looking at statistically significant results relating to exposure. As will become evident through the tables, charts and discussion below, the effect of exposure on the dependent variables largely acts through interactions with other explanatory variables. These interactions are therefore prioritised at the beginning of each section of model interpretation. After this, I present very brief summaries of any significant effects relating to the control variables, but as explained above, the role of the control variables is not the focus of this Chapter.

The analyses in this Chapter relate solely to the final models produced at the end of the methodological procedures described above. However, Appendix D contains a series of additional base models showing the coefficients and p values for our key independent variables of interest (exposure score, deprivation rate, exposure-deprivation interaction term) entered individually and in combination. These are provided for background context only. (Appendix F provides equivalent background models to accompany the non-poor models presented in Appendix E). The model outputs relating to the geographical level variance terms are presented separately in Appendix G.

In the interactions that are shown via the series of graphs in this chapter, the reference group within the categorical variable has typically been chosen to be the one which is arguably the 'most favourable' (or 'least disadvantaged') and the other categories are compared against this reference category.

The gradients of the lines shown in the series of charts in this chapter show the direction and magnitude of the relationships between different categories (within the relevant categorical variable) and the dependent response variable as the level of exposure changes.

It might be reasonable to expect that the more disadvantaged and/or vulnerable a person is and/or feels, the more likely the person is to have a strong aversion to income inequality and a strong belief that the government has a responsibility to redistribute income from the ‘haves’ to the ‘have-nots’, and Roberts (2014) found some evidence of this in his bivariate analyses. If this scenario was the case having also controlled for other factors, then the lines on the charts presented here (Figure 7.1, Figure 7.2, Figure 7.3 and Figure 7.4) would typically all be placed above the grey horizontal reference line (denoting OR=1) and would all show a positive association between exposure and the response variable. However, as shown below, this does not happen in practice.

7.5.1 Association between exposure to inequality and inequality aversion

Table 7.5: Model A - Inequality aversion – poor subgroup

Parameter	Estimate	Std. Error	z value	Pr(>|z|)	Sig.	OR	OR 95% CI lower	OR 95% CI upper	Group LR test pval
(A) Respondent's personal and household characteristics									
Educational attainment (Ref='No schooling'; N=165)									*0.0000*
Primary (N=472)	0.62	0.23	2.64	0.0084	**	1.85	1.17	2.93	
Grades 8-11 or equivalent (N=853)	1.06	0.23	4.70	0.0000	***	2.89	1.86	4.51	
Matric or equivalent (N=547)	0.71	0.23	3.03	0.0024	**	2.03	1.28	3.20	
Tertiary (N=83)	0.61	0.34	1.81	0.0710	.	1.84	0.95	3.58	
Household size (i.e. number of persons in the household)	-0.07	0.05	-1.54	0.1200		0.93	0.85	1.02	
ANC voter (N=1440)	-0.42	0.14	-3.05	0.0023	**	0.66	0.50	0.86	
(B) Respondent's attidudes and views									
Self-rated poverty status (Ref='Wealthy/very comfortable'; N=78)									*0.0170*
Reasonably comfortable (N=317)	0.06	0.30	0.21	0.8300		1.07	0.59	1.92	
Just getting by (N=882)	0.26	0.28	0.93	0.3500		1.30	0.75	2.23	
Poor (N=597)	0.34	0.29	1.17	0.2400		1.40	0.80	2.47	
Very poor (N=246)	0.80	0.32	2.47	0.0140	*	2.23	1.18	4.20	
Perceived mobility history (Ref='Upward mobility'; N=697)									*0.0037*
No mobility (N=863)	-0.09	0.12	-0.69	0.4900		0.92	0.72	1.17	
Downward mobility (N=560)	0.34	0.14	2.42	0.0160	*	1.40	1.07	1.85	
Expected future social mobility over next 5 years (Ref='Improve'; N=995)									*0.0700*
Stay the same (N=552)	-0.19	0.13	-1.46	0.1400		0.83	0.65	1.07	
Worsen (N=425)	0.18	0.15	1.21	0.2200		1.20	0.89	1.62	
Uncertain (N=148)	0.25	0.23	1.08	0.2800		1.28	0.82	2.00	
Higher job prestige than father (N=496)	0.37	0.12	3.03	0.0025	**	1.45	1.14	1.84	
Perceived class conflict index	0.11	0.06	1.96	0.0500	.	1.12	1.00	1.25	
Perceived group discrimination (N=425)	0.35	0.13	2.69	0.0072	**	1.42	1.10	1.84	
Perceived importance of 'merit' factors for getting ahead in life	0.45	0.06	7.99	0.0000	***	1.56	1.40	1.74	
Perceived importance of 'exogenous' factors for getting ahead in life	0.15	0.06	2.50	0.0120	*	1.16	1.03	1.31	
Perceived importance of 'race' for getting ahead in life	-0.18	0.06	-3.10	0.0019	**	0.83	0.74	0.94	
(C) Geographical location variables									
Province of residence (Ref='Western Cape'; N=249)									*0.0087*
Eastern Cape (N=338)	-0.46	0.43	-1.07	0.2800		0.63	0.27	1.46	
Northern Cape (N=160)	0.11	0.50	0.21	0.8300		1.11	0.42	2.95	
Free State (N=141)	0.45	0.48	0.92	0.3500		1.56	0.61	4.04	
KwaZulu Natal (N=380)	0.30	0.40	0.75	0.4500		1.35	0.61	2.98	
North West (N=117)	-0.21	0.45	-0.47	0.6400		0.81	0.34	1.96	
Gauteng (N=278)	-0.29	0.39	-0.75	0.4500		0.75	0.35	1.61	
Mpumalanga (N=191)	0.67	0.45	1.49	0.1400		1.95	0.81	4.70	
Limpopo (N=266)	0.69	0.44	1.55	0.1200		1.99	0.83	4.76	
(D) Neighbourhood level poverty and exposure variables and interactions									
Neighbourhood poverty rate	0.12	0.11	1.13	0.2600		1.13	0.91	1.40	
Exposure to inequality	0.25	0.17	1.50	0.1300		1.29	0.93	1.79	
Neighbourhood poverty rate * Exposure to inequality	0.01	0.08	0.09	0.9300		1.01	0.86	1.17	
Neighbourhood poverty rate * Household size	-0.13	0.05	-2.58	0.0099	**	0.87	0.79	0.97	
*Exposure to inequality * Perceived social mobility history (Ref='Upward mobility')*									*0.0120*
Exposure * No mobility	-0.09	0.12	-0.76	0.4500		0.91	0.72	1.15	
Exposure * Downward mobility	-0.38	0.13	-2.84	0.0044	**	0.68	0.52	0.89	
*Exposure to inequality * Expected future social mobility (Ref='Improve')*									*0.0037*
Exposure * Stay the same	0.13	0.13	1.00	0.3200		1.13	0.89	1.45	
Exposure * Worsen	0.04	0.14	0.31	0.7600		1.05	0.79	1.39	
Exposure * Uncertain	-0.65	0.20	-3.18	0.0014	**	0.52	0.35	0.78	

* significant at $p < 0.05$; ** significant at $p < 0.01$; *** significant at $p < 0.001$

In Model A there are significant interaction effects between *exposure_of_poor* and two categorical variables: 'past social mobility' (variable '*ssocmobc*' in Table 7.4) and 'expected future social mobility' (variable '*futmob*' in Table 7.4). Interaction odds ratio plots are presented below for each.

Figure 7.1 shows the interaction between 'past social mobility' over recent years and exposure to inequality. At the mean '*exposure_of_poor*' level (denoted by the black dashed vertical reference line), people who regard themselves as having suffered downward mobility over recent years show significantly greater aversion to income inequality than people who regard themselves as being upwardly mobile (p-value < 0.05)[44].

However, the relationship between perceived recent social mobility and the inequality aversion response variable can be seen to vary according to the level of exposure. The negative slope of the red line on the chart indicates that as '*exposure_of_poor*' increases, people who regard themselves as being downwardly mobile become progressively less averse to income inequality compared to people who regard themselves as having been upwardly mobile (p-value < 0.01)[45,].

So, in areas where inequality is perhaps not particularly obvious on a day-to-day basis, people who regard themselves as having suffered downward social mobility tend to have greater aversion to income inequality than people who regard themselves as having enjoyed upward social mobility. This finding could potentially reflect a greater

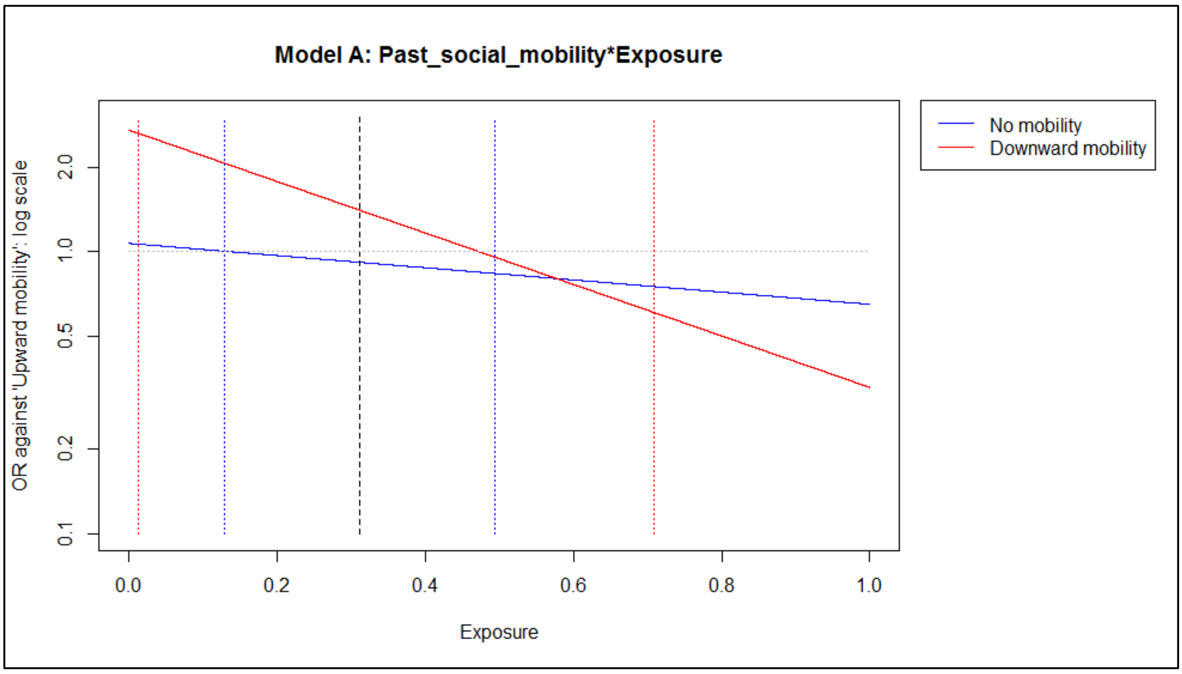
[44] Specifically, the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 1.40 times (CI: 1.07-1.85) greater for a respondent who perceives him/herself to have been downwardly mobile over recent years than for a person who is identical in every way except for perceiving him/herself to have been upwardly mobile over recent years.

[45] When the level of exposure is 1 standard deviation below the mean exposure, the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 2.06 times greater for a respondent who perceives him/herself to have been downwardly mobile over recent years than for an identical person who thinks they have been upwardly mobile. In contrast, when the level of exposure is 1 sd above the mean, the odds of a self-perceived downwardly mobile person giving a response (i.e. inequality aversion) above any particular response threshold are only 0.96 times as high as a person perceiving recent upward mobility.

optimism amongst upwardly mobile people that life is improving in various ways and therefore greater income and social status may be forthcoming, whereas for downwardly mobile people it may seem that the gap between the ‘haves’ and the ‘have nots’ is getting wider and progressively less ‘bridgeable’ thereby making the income inequality in South Africa seem increasingly unjust.

Model A: Exposure variables and interactions

Figure 7.1: Interaction of ‘past social mobility’ with ‘exposure’

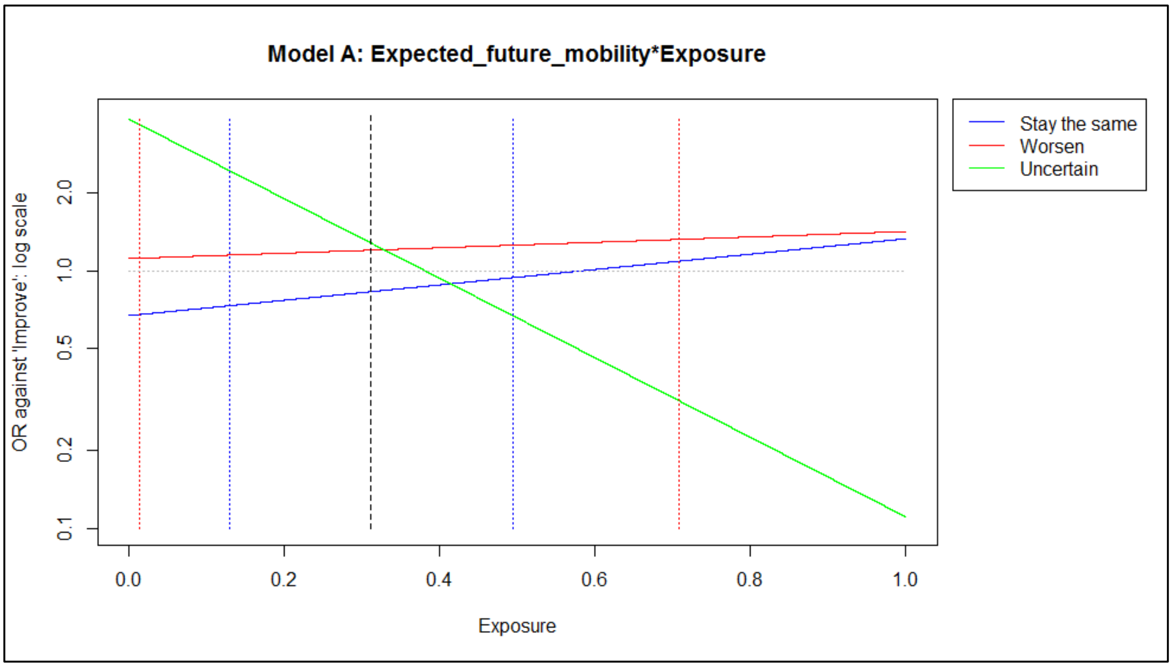


Note regarding vertical reference lines: black dash = mean exposure; blue dot = mean +/- 1 standard deviation exposure; red dot = minimum/maximum observed exposure
Note regarding horizontal reference line: grey dot = OR of 1.0.

At the other end of the ‘*exposure_of_poor*’ spectrum, where people are exposed to the obvious visual signs of income inequality on a daily basis, the results suggest that downwardly mobile people are *less averse* to income inequality than upwardly mobile people. This result is rather unexpected but could potentially be due to disillusionment amongst the upwardly mobile population with their chances of ever spanning the divide between the ‘haves’ and the ‘have nots’. It is important to note, however, that the lower aversion to income inequality of upwardly mobile people only applies where the level of ‘*exposure_of_poor*’ is approximately 1 standard deviation or more above the mean ‘*exposure_of_poor*’ and so, in the majority of cases across South Africa, upwardly mobile people will typically be less averse to income inequality than downwardly mobile people.

Figure 7.2 shows the interaction of ‘expected future social mobility’ and exposure to inequality.

Figure 7.2: Interaction of ‘expected future social mobility’ with ‘exposure’



Note regarding vertical reference lines: black dash = mean exposure; blue dot = mean +/- 1 standard deviation exposure; red dot = minimum/maximum observed exposure
Note regarding horizontal reference line: grey dot = OR of 1.0.

At the mean ‘*exposure_of_poor*’ level, people who expect their societal position will improve over the coming years are not significantly different in terms of aversion to income inequality then people who think their societal position will stay the same or worsen or who are uncertain. Furthermore, there are no statistically significant differences in how exposure to inequality affects aversion to income inequality between people who think their social position will improve, stay the same or worsen.

However, there is a statistically significant difference (p-value < 0.01) between people who think their social position will improve and people who are uncertain of their social mobility future in the way that *exposure_of_poor* affects views on inequality aversion. In geographical areas where *exposure_of_poor* is relatively low, people who are uncertain of their social mobility futures tend to have a greater aversion to income inequality than people who think their social position will improve[46]. As *exposure_of_poor* increases, however, people who are uncertain of their social mobility future become progressively less averse to income inequality compared to people who expect to enjoy improving social status[47].

So, where exposure to inequality is low, people who are uncertain about their future tend to be more averse to income inequality than people who think their situation will get better. As exposure increases, people who are uncertain about their futures become less averse to income inequality compared to people who expect an upwardly mobile future. This finding could again be related to disillusionment amongst people who are expecting to see improvements with regards to the size of the gap between the very affluent minority and the poor majority and the difficulty in bridging that gap.

[46] In geographical areas with the lowest observed *exposure_of_poor*, the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 3.70 times greater for a respondent who is uncertain about their social mobility future than for a person who is identical other than expecting their social position to improve.

[47] In areas with the highest observed *exposure_of_poor*, the odds of someone who is uncertain about their future social mobility giving a response (i.e. inequality aversion) above any particular response threshold are just 0.31 times those of a respondent who is identical other than expecting their social position to improve.

Model A: Other statistically significant effects

Although the main focus of this Chapter is on the effects of exposure to inequality on people's attitudes, the models also reveal a number of significant effects relating to the other explanatory variables. For instance, all other things being equal, there is a statistically significant interaction between neighbourhood poverty rate and household size (p-value < 0.01). The coefficient for this interaction is negative, meaning that, as neighbourhood poverty rate increases, the effect of living in a large household becomes progressively greater in terms of reducing aversion to income inequality.

With regards to education, the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 1.85 times greater for a respondent with primary schooling as their maximum educational attainment than people with no schooling (p-value < 0.01). A similar picture is seen for those with 'Grade 8-11 or equivalent' (OR = 2.89; p-value < 0.001) and 'matric or equivalent' (OR = 2.03; p-value < 0.01), again when compared to people with no schooling. There is also a similar picture for people with tertiary education (OR = 1.84; p-value < 0.1), although this does not quite meet the 0.05 significance level (possibly due to small N cases).

In terms of political affiliation, the odds of giving a response above any particular response threshold are lower for an ANC voter than for a non-ANC voter (OR = 0.66; p-value < 0.01). This could be because ANC voters perhaps value the good work that the ANC has done more highly than the persisting problems of inequality?

In terms of self-perceived poverty status, people who regard themselves as 'very poor' have higher odds of answering above a specified response threshold than people who regard themselves as 'wealthy/very comfortable' (OR = 2.23; p-value < 0.05). So ‘very poor’ people are more averse to income inequality than ‘wealthy/very comfortable’ people. With regards to employment, people who regard themselves as having a job with higher prestige than their father have higher odds of giving an above-threshold response than people who don’t (OR = 1.45, p-value < 0.01), and so are more averse to income inequality.

Further, the greater the extent of class conflict in South Africa a person perceives, the higher the odds that they will give an above-threshold response. A person with a perceived class conflict score of +1sd above mean score has odds 1.12 times greater than someone with mean class conflict score (p-value = 0.05). Perhaps those who think there is lots of class conflict also think that this conflict might be reduced by reducing income disparities? With regard to perceived inter-group discrimination, people who perceive themselves to be members of a group that are discriminated against are more averse to income inequality than people who don’t (OR = 1.42, p-value < 0.01).

The greater one’s view that ‘merit’ factors (i.e. hard work and ambition) are important for getting ahead in life, the greater the aversion to income inequality (OR for +1sd vs mean on merit index = 1.56, p-value < 0.001). However, it is also the case that the greater one’s view that ‘exogenous’ factors (e.g. coming from a wealthy family) are important for getting ahead in life, the greater the aversion to income inequality (OR for +1sd vs mean on exogenous index = 1.16, p-value < 0.05). And finally the greater one’s view that a person’s race is important for getting ahead in life, the less they are averse to income inequality (OR for +1sd vs mean on ‘q168r’ (i.e. importance of race) variable = 0.83, p-value < 0.01).

With regard to geographical differences, when compared to the Western Cape, none of the other eight provinces are statistically significant in terms of inequality aversion. However, the group effect of province is statistically significant as a whole and so province is included in the final model.

7.5.2 Association between exposure to inequality and support for government responsibility for redistribution

Table 7.6: Model C - Government Redistribution – poor subgroup

Parameter	Estimate	Std. Error	z value	Pr(>|z|)	Sig.	OR	OR 95% CI lower	OR 95% CI upper	Group LR test pval
(A) Respondent's personal and household characteristics									
Employment status (Ref='Employed full time'; N=498)									*0.0024*
Employed part time (N=150)	0.64	0.22	2.86	0.0042	**	1.90	1.22	2.94	
Unemployed, seeking work (N=688)	-0.07	0.15	-0.48	0.6300		0.93	0.70	1.25	
Unemployed not looking for work (N=230)	0.54	0.20	2.68	0.0073	**	1.71	1.16	2.53	
Pensioner (N=213)	0.15	0.20	0.76	0.4500		1.17	0.79	1.73	
Student/learner (N=201)	0.24	0.19	1.26	0.2100		1.27	0.88	1.85	
Permanently sick/disabled (N=47)	-0.11	0.36	-0.29	0.7700		0.90	0.44	1.84	
Other employment status (N=93)	-0.05	0.28	-0.18	0.8500		0.95	0.55	1.65	
(B) Respondent's attidudes and views									
Inequality aversion	0.35	0.05	6.58	0.0000	***	1.42	1.28	1.57	
Left-right ideology (Ref='Liberal'; N=801)									*0.0000*
Moderate (N=567)	-0.47	0.12	-3.82	0.0001	***	0.62	0.49	0.79	
Conservative (N=316)	-0.80	0.15	-5.26	0.0000	***	0.45	0.33	0.60	
Uncertain (N=436)	-0.29	0.15	-1.94	0.0520	.	0.75	0.56	1.00	
Self-rated poverty status (Ref='Wealthy/very comfortable'; N=78)									*0.0001*
Reasonably comfortable (N=317)	-0.46	0.31	-1.50	0.1300		0.63	0.35	1.15	
Just getting by (N=882)	-0.83	0.29	-2.91	0.0036	**	0.43	0.25	0.76	
Poor (N=597)	-0.63	0.29	-2.15	0.0310	*	0.53	0.30	0.95	
Very poor (N=246)	-0.11	0.33	-0.33	0.7400		0.90	0.47	1.71	
Higher job prestige than father (N=496)	-0.29	0.12	-2.43	0.0150	*	0.75	0.59	0.94	
Perceived class conflict index	0.19	0.06	3.37	0.0008	***	1.21	1.08	1.35	
Perceived group discrimination (N=425)	0.30	0.13	2.37	0.0180	*	1.35	1.05	1.73	
Perceived importance of 'merit' factors for getting ahead in life	0.20	0.05	3.72	0.0002	***	1.22	1.10	1.35	
Perceived importance of 'exogenous' factors for getting ahead in life	0.17	0.06	3.12	0.0018	**	1.19	1.07	1.32	
Perceived importance of 'race' for getting ahead in life	-0.13	0.05	-2.39	0.0170	*	0.88	0.79	0.98	
(C) Geographical location variables									
N/A									
(D) Neighbourhood level poverty and exposure variables and interactions									
Neighbourhood poverty rate	0.01	0.11	0.12	0.9000		1.01	0.82	1.25	
Exposure to inequality	-0.49	0.29	-1.69	0.0920	.	0.61	0.35	1.08	
Neighbourhood poverty rate * Exposure to inequality	-0.11	0.08	-1.38	0.1700		0.90	0.77	1.05	
*Exposure * Employment status (Ref='Employed full time')*									*0.0045*
Exposure * Employed part time	0.75	0.22	3.32	0.0009	***	2.11	1.36	3.27	
Exposure * Unemployed, seeking work	0.12	0.15	0.84	0.4000		1.13	0.85	1.51	
Exposure * Unemployed not looking for work	0.40	0.19	2.07	0.0390	*	1.50	1.02	2.19	
Exposure * Pensioner	0.18	0.19	0.93	0.3500		1.20	0.82	1.74	
Exposure * Student/learner	0.18	0.18	0.96	0.3300		1.19	0.83	1.70	
Exposure * Permanently sick/disabled	0.90	0.39	2.32	0.0200	*	2.45	1.15	5.23	
Exposure * Other employment status	0.67	0.28	2.45	0.0140	*	1.96	1.14	3.37	
*Exposure * Self-rated poverty status (Ref='Wealthy/very comfortable')*									*0.0020*
Exposure * Reasonably comfortable	0.02	0.28	0.08	0.9300		1.02	0.59	1.78	
Exposure * Just getting by	0.58	0.26	2.19	0.0290	*	1.78	1.06	2.99	
Exposure * Poor	0.32	0.27	1.19	0.2300		1.38	0.81	2.33	
Exposure * Very poor	0.36	0.30	1.19	0.2300		1.44	0.79	2.60	

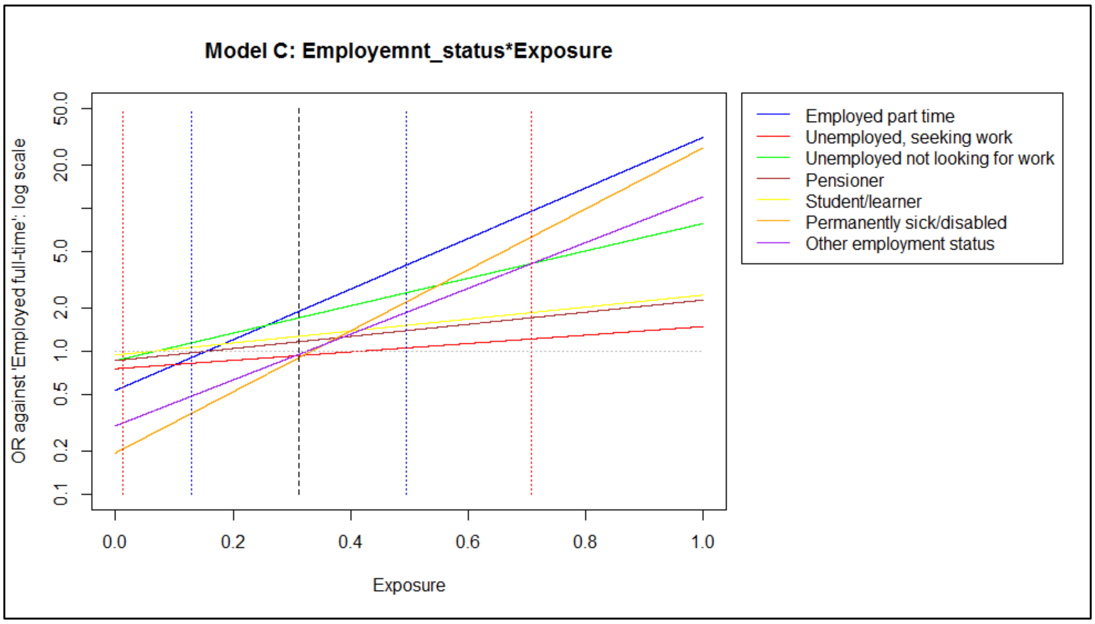
* significant at $p < 0.05$; ** significant at $p < 0.01$; *** significant at $p < 0.001$

In Model C there are significant interaction effects between *exposure_of_poor* and two categorical variables: 'employment status' (variable '*empl*' in Table 7.4) and 'self perceived poverty status' (variable '*spoor*' in Table 7.4). Interaction odds ratio plots are presented below for each.

Model C: Exposure variables and interactions

Figure 7.3 shows the interaction between employment status and exposure to inequality.

Figure 7.3: Interaction of 'employment status' with 'exposure



Note regarding vertical reference lines: black dash = mean exposure; blue dot = mean +/- 1 standard deviation exposure; red dot = minimum/maximum observed exposure
Note regarding horizontal reference line: grey dot = OR of 1.0.

At the mean '*exposure_of_poor*' level (denoted by the black dashed vertical reference line), people who are employed part time (p-value < 0.01) and people who are unemployed and not looking for work (p-value < 0.01) both show a significantly stronger agreement than people who are employed full time that it is the government's responsibility to reduce income differences between those with high incomes and those with low incomes[48].

It is also evident that four of the employment status groups show statistically significant differences in the way that *exposure_of_poor* influences the response variable when compared to the way that *exposure_of_poor* influences the response of 'Employed full time' respondents. These four groups are: employed part time (p-value < 0.001); unemployed and not looking for work (p-value < 0.05); permanently sick or disabled (p-value < 0.05); and 'other' employment status (p-value < 0.05). The interaction coefficients for these four employment status groups are all positive, meaning that as *exposure_of_poor* increases, members of the four groups become progressively more supportive of government's responsibility to redistribute income when compared to people who are employed full time[49].

The other non-significant interaction terms are also seen to be positive in direction but the lack of statistical significance means that caution must be exercised with regard to these coefficients. However, based upon the results here, it is possible to tentatively conclude that, on the whole, those respondents who might be regarded as involuntarily excluded (either fully or partially) from the labour market tend to be more

[48] At the mean '*exposure_of_poor*' level, the odds of an 'Employed part time' respondent giving a response (i.e. support for government responsibility for income distribution) above any particular threshold are 1.90 times (CI: 1.22-2.94) greater than those of a respondent who is 'Employed full time'. Similarly, a respondent who is 'Unemployed and not looking for work' has odds 1.71 times (CI: 1.16-2.53) greater than someone who is 'Employed full time'.

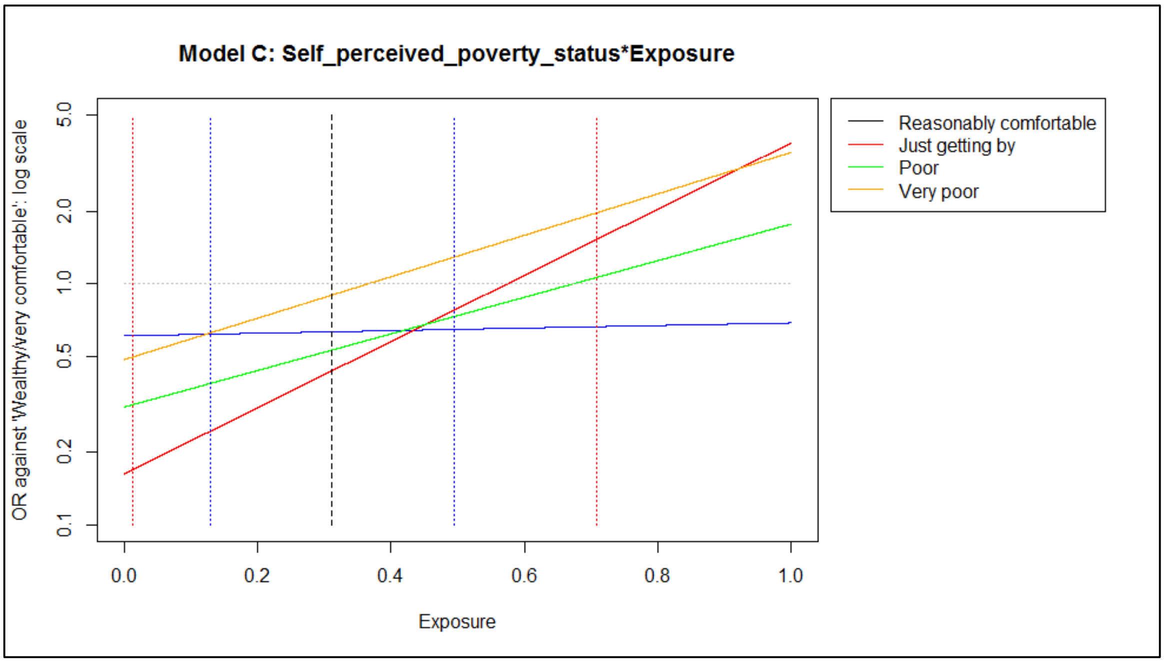
[49] At mean-1sd *exposure_of_poor*, 'Employed part time' respondents have odds just 0.90 times those of 'Employed full time' of answering above a given response threshold, whereas at mean+1sd *exposure_of_poor* the odds for 'Employed part time' are 4.00 times greater than 'Employed full time'. Similarly, at mean-1sd *exposure_of_poor*, the odds of an 'Unemployed and not seeking work' respondent of answering above a response threshold are 1.14 times greater than 'Employed full time', and at mean+1sd *exposure_of_poor* these odds increase to 2.56 times greater. Respondents who are 'Permanently sick or disabled' have odds at mean-1sd *exposure_of_poor* of just 0.37 times those of 'Employed full time', while at mean+1sd *exposure_of_poor* the odds are 2.21 times greater than 'Employed full time'. Finally, for people who state they are of 'Other employment status', the odds are 0.48 those of the 'Employed full time at mean-1sd *exposure_of_poor*, but increasing to 1,86 times greater than 'Employed full time' at mean+1sd *exposure_of_poor*. These statistically significant interaction terms are all positive in direction.

supportive of the government taking responsibility for income redistribution than people in full time employment. Furthermore, there is some evidence that as exposure to socio-economic inequality increases, people who are involuntarily excluded from the labour market become even more supportive of government role in income redistribution compared to full time employed people. These findings may be related to increased feelings of income vulnerability amongst people who are excluded from the labour market: those who are employed part time may wish to work more but may struggle to find employment opportunities and may be in relatively insecure jobs; those who are unemployed and not seeking work may be discouraged work seekers who feel it is pointless looking for jobs due to scarcity; and those who are permanently sick or disabled may feel particularly disadvantaged in the labour market due to an inability to obtain and/or retain jobs. It is not clear what employment statuses constitute the 'other' group, but these too may potentially be vulnerable or disadvantaged positions.

Figure 7.4 shows the interaction of 'self-perceived poverty status' and exposure to inequality. In terms of self-perceived poverty status, at the mean *exposure_of_poor* level people who regard themselves as 'just getting by' or as 'poor' are significantly *less supportive* than people who regard themselves as being 'wealthy/very comfortable' of the notion that the government is responsible for redistributing income from those with high incomes to those with low incomes (p-value < 0.01 for 'just getting by'; p-value < 0.05 for 'poor')[50].

[50] The odds at mean *exposure_of_poor* for people who are 'Just getting by' are only 0.43 times (CI: 0.25-0.76) those of people who regard themselves as 'Wealthy/very comfortable', while the odds at mean *exposure_of_poor* for people who regard themselves as poor are just 0.53 times (CI: 0.30-0.95) those who regard themselves as 'Wealthy/very comfortable'. The odds at mean exposure for people who regard themselves as 'Reasonably comfortable' or 'Very poor' are lower than those who regard themselves as 'Wealthy/very comfortable', but the differences are not statistically significant.

Figure 7.4: Interaction of 'self-perceived poverty status' and exposure to inequality



Note regarding vertical reference lines: black dash = mean exposure; blue dot = mean +/- 1 standard deviation exposure; red dot = minimum/maximum observed exposure
Note regarding horizontal reference line: grey dot = OR of 1.0.

In terms of the exposure gradients of the interaction terms, of the four subgroups shown on the chart, only the line for people who are 'Just getting by' is significantly different to the 'wealthy/very comfortable' group. As *exposure_of_poor* increases, people who are 'just getting by' become progressively more supportive of the government's role in income redistribution compared to people who are 'wealthy/very comfortable' (p-value < 0.05)[51].

[51] Here, it is evident that at mean-1sd *exposure_of_poor*, the odds of someone who is 'Just getting by' answering above a given response threshold are just 0.24 times those of someone who is 'Wealthy/very comfortable', whilst at mean+1sd *exposure_of_poor*, the odds are just 0.77 times those of 'Wealthy/very comfortable'. At the highest actually observed *exposure_of_poor* level, the odds for those who are 'Just getting by' are 1.52 times greater than those for the 'Wealthy/very comfortable' group, but it is clear from Figure 7.4 that across the majority of the observed *exposure_of_poor* continuum in South Africa, people who are 'Just getting by' are less supportive of government role in income redistribution than people who are 'Wealthy/very comfortable'.

The interaction coefficients for the ‘poor’ and ‘very poor’ groups are positive but of a somewhat smaller magnitude than the interaction coefficient for ‘just getting by’ and, unlike ‘just getting by’, the interaction coefficients for ‘poor’ and ‘very poor’ are not significantly different to that for the ‘wealthy/very comfortable’ group. It may be that the relatively small number of respondents in the Model_C dataset that regard themselves as ‘wealthy/very comfortable’ (n=78) is the reason why the ‘poor’ and ‘very poor’ interaction coefficients are not presented as significantly different to ‘wealthy/very comfortable’. One option might have been to recode the ‘wealthy/very comfortable’ and ‘reasonably comfortable’ respondents into a new composite category and use that as the reference group. However, after consideration, the decision was taken not to do this because there is possibly some evidence that the ‘wealthy/very comfortable’ and ‘reasonably comfortable’ groups may not necessarily be sufficiently similar to justify combining them[52].

Based upon the information conveyed graphically in Figure 7.4, and particularly the significant interaction coefficient for people who are ‘just getting by’, it is possible to tentatively suggest that increasing exposure to socioeconomic inequality is associated with progressively greater support for government income redistribution amongst the less well-off when compared to those who regard themselves as being ‘wealthy/very comfortable’.

[52] For instance, although the odds at mean *exposure_of_poor* are not quite significantly different between these two groups at the 0.05 level, the visual representation provided by Figure 4.4 suggests there may be a small yet consistent difference between the two groups which may become significant if the number of cases in the ‘Wealthy/very comfortable’ group were larger.

Model C: Other statistically significant effects

Again, although the main focus of this chapter is on the effects of exposure to inequality on societal attitudes, a number of other explanatory variables also show statistical significance. For instance, there is a positive association between inequality aversion and support for government role in income distribution. *Ineqavr* is treated as a continuous explanatory variable and so is mean-centred. A person whose response on the *ineqavr* variable is 1sd above the mean *ineqavr* response, has odds 1.42 times greater than an *ineqavr* mean respondent of giving a response on the *govredr* variable above any specified threshold (p-value < 0.001). In other words, there is a statistically significant positive association between a person's inequality aversion and their belief that it is the government's responsibility to redistribute income from the poor to the rich. In short, all other things being equal, people with greater aversion to income inequality tend to show stronger support for government role in income redistribution.

There is also a significant association between people's political ideology and their support for government role in income distribution. People who regard themselves as 'conservative' have odds just 0.45 times as large as people who regard themselves as 'liberal' of answering above a given response threshold (p-value < 0.001). People who regard themselves as 'moderate' have odds just 0.62 times those of 'liberal' (p-value < 0.001), whilst people who say they are 'uncertain' have odds just 0.75 of 'liberal' (p-value = 0.0520). The magnitude of the coefficients and ORs therefore become more extreme with each increment in political ideology, with each increment away from 'liberal' being associated with lower odds of answering above a given response threshold. In other words, all other things being equal, 'liberal' people are most supportive of government role in income redistribution, followed by people who are ‘uncertain’ and then by ‘moderate’ people, with ‘conservative’ people least supportive of government role in income redistribution.

With regards to employment factors, people who regard themselves as having a higher-prestige job than their father are significantly less supportive of government role in income redistribution than people who regard their job as having the same or less prestige than their father’s (p-value < 0.05).

In terms of perceived conflict and discrimination, there is a positive association between the level of perceived class conflict in South Africa and support for government role in income redistribution (p-value < 0.001) whilst people who perceive themselves to be part of a group that suffers discrimination are typically more supportive of government role in income redistribution than people who do not feel part of a discriminated group (p-value < 0.05).

There is a significant positive relationship between the degree to which people think that ‘merit factors’ are important for getting ahead in life and their support for government role in income redistribution (p-value < 0.001). There is also a significant positive relationship between the degree to which people think that ‘exogenous factors’ are important for getting ahead in life and their support for government role in income redistribution (p-value < 0.01). In contrast, there is a significant negative association between the degree to which people think that ‘racial group’ is important for getting ahead in life and their support for government role in income redistribution (p-value < 0.05).

7.6 Conclusions

The modelling approach developed and applied in this chapter was designed to test whether exposure to socioeconomic inequality is associated with attitudes towards inequality and government role in redress. The basic hypothesis articulated throughout this thesis is that people who experience inequality on a day to day basis may be more aware of the stark inequalities in South Africa and therefore may be more averse to inequality and more supportive of the government taking responsibility for reducing income differentials.

The analyses presented in this chapter focused primarily on the coefficients and odds ratios relating to the associations between exposure and the two selected dependent variables. The other statistically significant variables remaining in the final models are utilised only as control variables.

The models demonstrate that exposure to inequality is indeed associated with people's attitudes to inequality and options for redress, but that the associations are primarily observable through the interaction of exposure with other explanatory variables. In the statistically significant interactions involving the exposure measure, the other variable in the interaction is almost always categorical. These interactions can therefore be interpreted as measuring the extent to which exposure differentially acts on categories of the relevant categorical variable.

In relation to inequality aversion, only two statistically significant interactions with exposure were observed with the outcome measure. Firstly, as exposure to inequality increases, those respondents who perceived themselves to have experienced downward social mobility over recent years become progressively less averse to income inequality compared to respondents who perceived themselves to have experienced upward mobility, all other things being equal. Secondly, as exposure to inequality increases, those respondents who foresee an uncertain future social mobility trajectory become progressively less averse to income inequality compared to respondents who foresee an upward trajectory. As discussed above, these findings could potentially be related to increased disillusionment amongst the upwardly mobile populations of high exposure areas in terms of their chances of ever spanning the divide between the 'haves' and the 'have nots'.

In relation to support for government role in income redistribution, five statistically significant interactions with exposure were observed with the outcome measure. Four of these five significant interaction terms are categories of the employment status variable: employed part time, unemployed and not looking for work, permanently sick or disabled, and 'other'. As exposure to inequality increase, respondents in these four employment categories become progressively more supportive of the government's responsibility for income redistribution compared to respondents who were employed full-time, all other things being equal. The fifth significant interaction term in this model related to self-perceived poverty status. As exposure to inequality increases, people who perceive themselves as 'just getting by' become progressively more supportive of government responsibility for income redistribution compared to people who derive themselves to be wealthy or very comfortable. These findings are in line with the initial hypothesis articulated above that less advantaged groups of the population may be more supportive of redress measures, particularly for those who have high exposure to inequality.

These are informative findings, however, an important factor to bear in mind concerning all the analyses here is that there is generally a very strong consensus amongst the population of South Africa that income inequality is too high and that the government has a responsibility to reduce it by redistributing income from those with high incomes to those with low incomes. As such, there is not a great deal of discrimination between response intervals on the dependent variables, as the responses are clustered around the ‘agree/strongly agree’ end of the scale. Given that SASAS is an annual nationally representative survey, it would be desirable to devote a future module to further examination of social inequality issues, and to incorporate into the new module a broader set of attitudinal questions that might provide greater discrimination between response intervals and therefore greater clarity in the modelling results.

It is worth noting here that the exposure measures are, by design, weighted averages of datazone deprivation rates across bespoke spatial extents (as discussed in detail in Chapter 4). I acknowledge that using weighted averages as geographical covariates in statistical modelling has the potential to weaken the associations with the dependent variable as the weighted average removes local variation present in the underlying spatial dataset. As the degree of ‘geographical smoothing’ through the weighting process increases, one might expect the strength of the relationship with the dependent variable to progressive decrease. However, for my purpose, the process of constructing exposure scores as weighted averages is both conceptually justified and methodologically necessary and so I do not regard this as problem for this thesis.

However, the analyses here in Chapter 7 are constrained by a number of other data limitations. For instance, the exposure and deprivation measures relate to 2001 whereas the SASAS attitudinal data relate to 2009, so there is a degree of temporal inconsistency. There is also the issue of the exposure and deprivation scores only being available for the datazone in which each respondent was residing at the time of the 2009 SASAS data collection fieldwork, so it was not possible to take into consideration the extent of exposure or deprivation that respondents may have experienced previously living in other neighbourhoods – this is particularly important in the context of long distance migration patterns, such as from the rural areas to the urban areas for work opportunities. An additional consideration is the degree to which respondents experience inequality through occasional visits to other parts of the country whilst visiting friends and family. These limitations are further considered in Chapter 8 and recommendations are made for how they can be addressed through future research.

Finally, and as noted at the beginning of this chapter, this is just one example of how my exposure measures could be used as independent variables in statistical models. There are many other ways in which the measures could be used, some of which are mentioned in the final chapter.

Chapter 8: Conclusion

8.1 Thesis overview

The objective of this thesis is to explore the feasibility of producing a small area level measure of exposure to inequality. Although existing national level measures of inequality (such as the commonly quoted Gini coefficient of income inequality) are important, they are insufficiently nuanced for exploring the relationships between people's 'lived experience' of inequality and social problems such as crime. I therefore identified the need for an empirical measure that better reflected people's lived experience of inequality and was shaped by the environments in which they lived their lives. The main purpose of this thesis has been to develop a measure of the lived experience of inequality which has the potential to be used to explore associations with social outcomes at both the individual level (such as social attitudes) and the area level (such as rates of violent crime).

There is an expanding international evidence base, which was referred to in Chapter 2, concerning links between inequality and outcomes such as crime and poor health (e.g. Wikstrom, 2006a; Wilkinson and Pickett, 2010). In this context, the subject of this thesis has broad international relevance. However, I chose to focus geographically on the situation in South Africa. This is because, as demonstrated in Chapter 2, South Africa has one of the highest levels of inequality in the world (as measured conventional income inequality indices (e.g. Leibbrandt et al., 2010)) and one of the highest levels of violent crime in the world (e.g. UNODC, 2011). While some commentators have assumed a causal link between inequality and crime in South Africa and while there is a multitude of international criminological theories that suggest this might be the case, I have sought to develop a measure of geographical exposure to inequality that will allow potential relationships such as this to be examined empirically.

On the basis of my contention that a measure of the lived experience of inequality should reflect the local environments in which people live their lives, I have argued that the measure should reflect *neighbourhood* context. A person living on one side of a city may have a very different lived experience of inequality to someone living on the other side of the city, which may be very different to the experience of someone living in a smaller town, and this in turn may be very different to the experience of someone living in a rural area. In light of this, I have argued that a measure of the lived experience of inequality needs to be expressed at a *neighbourhood* level. Whilst I have acknowledged from the outset that people's experience of inequality may also be shaped by less tangible factors, such as national and international media, I argue that the localised geographical settings in which people live, work (and/or seek work), study, shop and travel through represent the principal means of being exposed to inequality. A particularly powerful visual indication of this can be seen through the two contrasting photographs included in Chapter 1 (Figure 1.1 and Figure 1.2).

In terms of the lived experience of inequality, one might expect that a poor resident of the Alexandra township in Johannesburg might have a very different lived experience of inequality to a poor resident of the former Transkei. I have sought, through this thesis, to develop an empirical measure that would enable me to quantify differences in the lived experience of inequality between residents of each small area in South Africa. This provides a new contribution to the evidence base concerning socio-economic inequality in South Africa and furthermore it has wider international relevance as a methodological approach to complement the conventional indicators of inequality such as the national Gini coefficient of income inequality.

The main requirement of my methodological approach is the availability of neighbourhood level data on socio-economic outcomes, such as a poverty rate or unemployment rate, in order to establish the socio-spatial environments in which individuals live. In this thesis I utilised the South African Index of Multiple Deprivation 2001 at Datazone level (SAIMD 2001) (Noble et al., 2009a) for this purpose which, as detailed in Chapter 3, was produced by a team including myself (see also Appendix A). The SAIMD 2001 assigned numbers – specifically, deprivation counts and rates – to every neighbourhood in South Africa and thus enabled me to observe the differences in levels of deprivation that residents of Alexandra township, for instance, might experience as they go about their daily lives (see Figure 1.1) and compare this to the situation for residents of, for instance, the former Transkei (see Figure 1.2).

The main focus of the thesis was therefore to develop a quantitative measure of the lived experience of inequality that reflected the indications of inequality depicted in the figures provided, including the photographs and accompanying thematic maps of deprivation. Given that a key motivation for this research is the need to better understand the relationships between inequality and other social outcomes, a case study was included (Chapter 7) to demonstrate how my measure of the lived experience of inequality could be used. The literature review in Chapter 2 highlighted routes by which inequality may shape people's attitudes (to themselves and to others) and that these attitudes are important factors in determining levels of crime (e.g. Wikstrom, 2006a) and poor health (e.g. Pickett and Wilkinson, 2015). In the light of that literature and my own empirical testing a case study on how people's experience of inequality might be associated with certain social attitudes was presented.

The overarching research question of the thesis is: “How does the lived experience of inequality vary spatially across South Africa and how is this associated with people’s attitudes towards inequality?” In order to address this overarching research question I specified five component sub-questions and structured the thesis around these five sub-questions. Each of the five empirical chapters (Chapters 3-7) tackles a specific sub-question. In the next section of this chapter I will review the findings of the analyses under each sub-question in turn and discuss how this relates to the overarching research question. Having reviewed the contribution of each sub-question to the overarching research question, I will offer an overall conclusion in relation to the overarching research question. Then, in the final part of this concluding chapter, I will consider the limitations of the research presented here and I make certain recommendations for how this research can be further developed in the future.

8.2 Addressing the research questions

The first research sub-question was: “To what extent is deprivation distributed unequally across neighbourhoods in South Africa?” and this was the focus of Chapter 3 of the thesis. Addressing this particular research question was a necessary first step to understanding the socio-spatial environments in which people carry out their routine daily activities and which therefore form the settings for experiencing socio-economic inequality. I interrogated the SAIMD 2001 at Datazone level and looked both at the overall composite index as well as the four constituent domains of deprivation that were conceptualised and measured as deprivation *rates*: income and material deprivation; employment deprivation; education deprivation; and living environment deprivation.

Through these analyses I demonstrated that very high proportions of the South Africa population were deprived on each of the four domains of deprivation when considered at an aggregate level, with over three-quarters of the total population of South Africa being deprived on the income and material deprivation domain and two-thirds of the total population being deprived on the living environment domain. Although the proportions deprived on the employment and education deprivation domains were lower (45% and 36% of the relevant populations respectively), they are still very high by any standard. Whilst these national totals say nothing about spatial differences within South Africa, they do nonetheless provide an important introduction to the consideration of deprivation levels within the country and provide the national benchmark against which to consider the varying levels of deprivation at sub-national level. At a very high level of sub-national disaggregation, namely disaggregation by province, I demonstrated that deprivation levels were markedly lower in the Western Cape and Gauteng than in the other seven provinces. The Western Cape and Gauteng are arguably the two provinces that are dominated to the greatest degree by their constituent metropolitan municipalities, and this disjuncture between metropolitan and non-metropolitan areas with regards to deprivation levels and patterns is a feature which recurs throughout the thesis.

The broad spatial disparities between provinces observed in relation to aggregate deprivation rates were shown to be considerably more pronounced when looking at the distribution of *datazone level* deprivation levels both between and within provinces. Using the deprivation ranks on the composite SAIMD measure I showed how the most severely deprived decile of neighbourhoods in South Africa were almost entirely located in the provinces of Eastern Cape and KwaZulu-Natal. This pattern was also borne out when the data were expressed as the proportion of each province's constituent datazones that fell within the most deprived national decile, with disproportionately large shares of datazones within the Eastern Cape and KwaZulu-Natal being classed amongst this most severely deprived group of datazones.

These findings regarding the spatial patterning of datazones on the composite SAIMD were also largely (although not unanimously) confirmed when looking at deprivation ranks across the four constituent domains of deprivation analysed in Chapter 3. Specifically, datazones in the most deprived decile of datazones on the domain measures were mostly found within the Eastern Cape and KwaZulu-Natal, while the least deprived datazones were mostly located within the Western Cape and Gauteng. When considering the entire distribution of datazones on each domain measure (i.e. not just the most/least deprived deciles), strong positive correlations were observed between the income and material deprivation, education and living environment domains (in terms of both domain scores and domain ranks). The correlations between the employment domain and the other three domains were also positive but were somewhat weaker. This suggests that a different set of dynamics may be shaping the spatial patterning of employment deprivation than are shaping the spatial patterning of income and material deprivation, education deprivation and living environment deprivation.

The contrasting results of the Western Cape and Gauteng on the one hand, and the Eastern Cape and KwaZulu-Natal on the other, suggested a possible differentiation by urban/rural status. As noted above, the Western Cape and Gauteng are largely dominated by their constituent metropolitan municipalities, whereas the Eastern Cape and KwaZulu-Natal are largely dominated by their mainly rural former homeland areas. I explored the issue of urban/rural differentiation further and identified clear and consistent patterns across the three domains of income and material deprivation, education deprivation and living environment deprivation. Specifically, on each of these three domains, the rural datazones are concentrated within the most deprived deciles while the urban datazones are concentrated within the least deprived deciles. The picture on the employment domain is similar to the other three domains at the most deprived end of the spectrum (i.e. the rural areas are disproportionately represented), but the pattern is slightly less pronounced at the least deprived end of the spectrum (where there was a notably number of rural datazones in the least deprived decile as well as the concentration of urban datazones). The mainly rural former homeland areas have a long history of disadvantage that stretches back further than their formal designation as 'homelands' or 'Bantustans' by the apartheid government (indeed before the 1913 Land Act which designated as 'native reserves' many areas which would later be designated as 'homelands') (Terreblanche, 2002) and the analyses of deprivation rates within these former homelands show that these areas are still characterised by extremely high levels of deprivation.

These findings not only highlight the broad spatial disparities in deprivation levels between urban and rural areas (and the particular case of the former homelands), but they also demonstrate the importance of the choice of the underlying measure of deprivation for determining the results of the analysis. Had I adopted employment deprivation as the sole focus for this thesis (which would have been a defensible decision given the policy emphasis on employment as a means out of poverty) then I would have produced a set of results that may not have correlated particularly well with the other forms of deprivation that afflict South Africa (e.g. Income and material deprivation, education deprivation, living environment deprivation and other dimensions). By considering all four domains of deprivation – and assessing commonalities and differences – I have been able to provide a more comprehensive account of the uneven spatial distribution of deprivation in South Africa than if the focus had been on just one dimension alone.

In short, I demonstrated in Chapter 3 that deprivation rates are typically very high across South Africa at whatever spatial level is assessed, but particularly high in the rural areas and especially so in the former homeland areas. Whilst many urban areas also registered very high levels of deprivation, the urban centres – and particularly the major metropolitan cities – were also overwhelmingly where the *least deprived* areas were located. In terms of addressing the first research question, therefore, it is clearly evident that deprivation is unevenly distributed across neighbourhoods in South Africa, with urban/rural status being a crucial determinant. In relation to the overarching research question, the findings here suggest considerable differences in the socio-spatial environments in which people conduct their daily lives and which, I contend, shape their lived experiences of inequality.

The second research sub-question was: "Can an empirical measure be developed that reflects people's lived experience of inequality?" and this was the focus of Chapter 4 of the thesis. Having demonstrated that levels of deprivation are spatially varied, I sought to build upon this to develop a new measure to reflect the differing socio-spatial environments across the country which shape people's lived experience of inequality. As noted earlier in this concluding chapter, the review of the neighbourhood level datasets in South Africa (in Chapter 2) revealed the SAMID 2001 at Datazone level to be the most appropriate for my purpose. The availability of binary categorical distinctions in the SAIMD 2001 between people defined as deprived and not deprived meant that measures of residential segregation would be best suited to my analytical purpose. I established three criteria which were necessary in order for any measure of residential segregation to be regarded as suitable as a measure of the lived experience of inequality: first, it must permit comparisons of inequality between all neighbourhoods in South Africa on a consistent basis; second, it must reflect the experience of inequality from an individual's perspective; and, third, data to calculate the measure must be available at the sub-municipality level. I adopted the framework set out by Douglas Massey and Nancy Denton (1988) in their classic study and developed a series of different empirical measures of segregation using categorical data derived from the income and material deprivation domain of the SAIMD 2001. I reviewed each of these measures both in terms of the international literature and undertook empirical testing in order to ascertain whether each measure satisfied these three criteria.

I concluded that the distance-weighted $P*$ exposure index satisfied all three criteria and represented the most suitable measure for my purpose. This index quantifies the extent to which members of one population sub-group are likely to be exposed to members of another population sub-group as they go about their daily lives. In my case, where I am looking at experience of socio-economic inequality, it reflects the extent to which a person from one end of the socio-economic spectrum is likely to be exposed to individuals from the other end of the socio-economic spectrum. It therefore represents an effective measure of the extent to which people may experience inequality as they go about their daily lives. As exposure is measured as the likelihood of interaction between a member of one population sub-group and members of another population sub-group, two separate but complementary indices are produced: one reflecting exposure to inequality as experienced by the 'poor' sub-group, and another reflecting the exposure to inequality as experienced by the 'non-poor' sub-group. The distance-weighted $P*$ exposure measure discussed by Massey and Denton is regarded as a 'geographically weighted' measure because people are assumed to be exposed to one another within a defined geographical setting that includes but is not restricted to the home neighbourhood (whereas the basic P^* exposure index considered by Massey and Denton is a 'global' measure in that it considers exposure within the home neighbourhood only). However, I acknowledged Lloyd and Shuttleworth's (2012) argument that 'local' measures of segregation offer considerable added value for exploring spatial patterns of segregation at a detailed geographical level. Given my contention that people's experience of inequality is shaped by the particular geographical settings in which they live their daily lives, I therefore commenced the development of a 'local' variant of the distance-weighted P^* exposure index[53].

The methodological development of my local distance-weighted P^* exposure index required me to specify two separate but related parameters. The first of these was to establish the extent of the geographical bounds within which a person was most likely to travel on a day-to-day basis and within which, therefore, they were likely to experience inequality. The second was to establish the likelihood of a person visiting each of the constituent neighbourhoods within the geographical bounds established in the first parameter. As I acknowledged in Chapter 4, establishing these parameters necessarily requires adopting some relatively arbitrary assumptions as the *true* values for each parameter will vary person-to-person (and indeed will probably vary over time for each individual person). Having reviewed the international literature in the context of the South African situation, I selected suitable methodological approaches for both parameters. In terms of geographical bounds, I incorporated the assumption that people may routinely travel anywhere within their home local municipality and/or within a 20km radius of their home neighbourhood. In terms of the likelihood of a person actually visiting each constituent neighbourhood within those geographical bounds, a

[53] Which, as noted here, consisted of two separate but complementary indices: one representing exposure as experienced by the 'poor' and another representing exposure as experienced by the 'non-poor'.

modified distance decay approach was used in which I began with a standard linear distance decay function and then applied a deprivation-differential re-weighting in order to better reflect the combination of needs, opportunities and barriers that I contend act to shape people's travel patterns in South Africa. As I acknowledged within Chapter 4, detailed travel pattern data did not exist within South Africa to enable me to validate my choices for establishing these two parameters, but feedback from a series of academic and policy-focused seminars in South Africa has been largely supportive of the assumptions included within this approach.

In short, I concluded that it is possible to develop an empirical measure which reflects people's lived experience of inequality in South Africa, which was the objective set out in the second research question. The innovation of a 'local' version of the distance-weighted P^* exposure index in which the two key parameters (i.e. geographical bounds and distance decay weighting) are specified to reflect exposure in South Africa represents a contribution to the evidence base concerning the lived experience of inequality in South Africa and to the study of residential segregation more broadly.

The third research sub-question was: "Where within South Africa is people's lived experience of inequality highest and lowest?" and this was the focus of Chapter 5 of the thesis. In light of the very high levels of deprivation in South Africa demonstrated in Chapter 3 (e.g. three-quarters of the population was defined as deprived on the income and material deprivation domain of the SAIMD 2001), the analytical emphasis within Chapter 5 was on my measure of exposure to inequality as experienced by the *poor* population (with analyses of exposure to inequality as experienced by the non-poor population included as Appendix C). I commenced the analyses using my exposure measure constructed using input data derived from the income and material deprivation domain of the SAIMD 2001, using all datazones in the country. The main finding was that, amongst the poor population, exposure to inequality was highest for those individuals who live within or close to areas of affluence (or, more accurately, areas of low deprivation). This is because poor individuals living in otherwise relatively affluent (low deprivation) areas are likely to come into regular contact with people from the other end of the socio-economic spectrum and they are likely to regularly observe the very visible signs of affluence (e.g. material asset ownership) of residents of these areas. On the other hand, exposure to inequality is lowest for those poor individuals who live within *and* surrounded by areas of high deprivation. This is because these poor individuals are likely to have less frequent inter-personal contact with people from the opposite end of the socio-economic spectrum and be less likely to observe the visual signs of affluence.

This broad generalisation of exposure patterns resulted in some quite marked differences in exposure levels between and within geographical areas. For instance, within the City of Johannesburg I demonstrated that, amongst the poor population, exposure was highest for poor residents of the otherwise affluent suburbs such as Sandton, and yet also very high for residents of the extremely deprived Alexandra township located nearby. Furthermore, exposure was seen to be high for poor residents of the Soweto township, but not quite as high as in either Sandton or Alexandra. I argued that these findings provide a degree of validation (in terms of face-validity) that the exposure measures are functioning as intended. The contrasting situation in the mainly rural former homelands was clearly demonstrated using King Sabata Dalindyebo local municipality as a case study. In this municipality, where almost all of the residents in the rural areas were defined as deprived on the income and material deprivation domain of the SAIMD 2001, levels of exposure to inequality were amongst the lowest in the country. Again, these findings offer a degree of validation (in terms of face validity) that the exposure measures are functioning as intended as one would *not* expect, given the extreme and widespread deprivation afflicting these communities, that poor individuals would come into frequent contact with affluent individuals or observe many visible signs of affluence that would be commonplace in somewhere such as Sandton.

The analysis of the locations of the datazones with the highest exposure to inequality (amongst the poor population) across the country revealed that most of these neighbourhoods were concentrated within the major metropolitan municipalities, and particularly within Cape Town, Johannesburg, Tshwane and Ekurhuleni. In both Cape Town and Tshwane (along with the three smaller non-metropolitan municipalities of Gamagara, Saldanha Bay and Stellenbosch), over half the datazones within the municipality were classed within the highest decile of exposure scores nationally. None of the municipalities that contained one or more datazone in the highest exposure national decile contained any datazones in the lowest exposure national decile, however many datazones within these particular municipalities did exhibit wide ranges of exposure levels (including Tshwane). This demonstrates that while there is a degree of commonality in exposure levels across datazones *within* any given municipality, notable differences in exposure do exist depending on *where* within the municipality a person lives.

At the opposite end of the exposure distribution, a group of 30 municipalities were identified within which over half of the constituent datazones in each municipality were in the *lowest* exposure national decile, and almost all of these 30 municipalities were located within the former homelands. Furthermore, in six of these municipalities every single datazone was in the lowest exposure national decile. These lowest exposure municipalities are some of the most severely deprived parts of the country where almost everyone is classed as being deprived on the income and material deprivation domain of the SAIMD 2001. However, as shown through the consideration of King Sabata Dalindyebo municipality in Chapter 3, some highly deprived municipalities also contain small geographical pockets of *lower deprivation*. The town of Umtata in King Sabata Dalindyebo is such an example. Umtata contains a number of datazones that are characterised by relatively low deprivation rates (i.e. less than 30% of population deprived) and levels of exposure to inequality are consequently somewhat higher in and around Umtata than across the rural parts of the municipality. However, the levels of inequality seen in Umtata are still far lower than seen in similarly deprived parts of major urban areas such as Cape Town, Johannesburg, Tshwane and Ekurhuleni. Whilst being wary of providing an overly generalised summary of the exposure gradients across the country, it does seem that exposure is generally highest in the major metropolitan areas, somewhat lower in the smaller non-metropolitan towns (such as Umtata), and lowest in the rural former homelands. However, as noted throughout, there are variations in datazone exposure levels within any broad grouping of areas as well as between broad groupings.

Having identified a notable difference between metropolitan and non-metropolitan municipalities in terms of exposure measured using data derived from the income and material deprivation domain of the SAIMD 2001, I then focused explicitly on the metropolitan areas. I introduced comparable exposure measures based upon data derived from the employment, education and living environment domains of the SAIMD 2001 and analysed the results. I found strong positive correlations between the four exposure measures when assessed across all metropolitan datazones, and even stronger correlations when assessed over a single metropolitan municipality. I argued that there was a conceptual basis for combining the four separate exposure measures into a single composite measure in that each separate measure could justifiably be regarded as a proxy for the phenomenon of the 'lived experience of inequality'. The strong positive correlations observed between the four separate measures offered a degree of empirical support for combination. I used factor analysis to derive the weights through which the four measures were combined into an overall composite exposure index.

Analysis of the composite exposure index across the metropolitan datazones broadly supported the earlier findings relating to the exposure measure based solely on the income and material deprivation domain. More specifically, on this composite measure I found that Cape Town and Tshwane accounted for almost all of the highest exposure decile of metropolitan datazones. In contrast, all of the datazones in Mangaung and Buffalo City were in the lower exposure end of the metropolitan distribution. Although Cape Town and Tshwane exhibited commonalities in terms of their concentrations of high exposure datazones, they differed in that Tshwane also contained a number of datazones in the lower deciles of the metropolitan distribution whereas Cape Town did not. This is further evidence that levels of exposure to inequality vary within as well as between metropolitan municipalities and further justification for adopting a 'local' measure of exposure (as I have done in this thesis) rather than using the conventional 'global' or 'geographically weighted' measures that provide a single exposure value for the entire municipality. The analysis in Chapter 5 concluded with consideration of the variations in datazone exposure levels within Cape Town, first by grouping datazones by Census Main Place across the municipality and then by grouping the datazones by Census Sub Place just for the township of Khayelitsha. These concluding analyses further illuminate the detailed geographical patterning in people's lived experience of inequality according to where they go about their daily lives.

In short, I demonstrated how exposure to inequality varies spatially across South Africa, with generalisable differences observed between major metropolitan areas, smaller non-metropolitan towns, and rural areas. The vast majority of the highest exposure datazones are in the major metropolitan areas of the Western Cape and Gauteng, while the vast majority of the lowest exposure datazones are in the rural former homelands. However, I also demonstrated sizeable differences in exposure levels *within* these types of areas and I argue that it is these detailed spatial variations between proximately located neighbourhoods that necessitates the use of a 'local' measure of exposure to inequality, as I have done here in this thesis. I argued that the new exposure measure is a powerful new indicator to help better understand other socio-economic outcomes and the individual and area levels. The new exposure measures are not meant as a replacement for poverty or deprivation rates in the analysis of other socio-economic outcomes, but rather an additional indicator to complement the more conventional poverty and deprivation indicators.

The fourth research sub-question was: "Are there any neighbourhoods across South Africa with high rates of deprivation and high lived experience of inequality?" and this was the focus of Chapter 6 of the thesis. As was observed in Chapter 2, deprivation (or poverty) and inequality have both been posited within the international literature as potential factors causing (directly and/or indirectly) other social problems such as crime (e.g. Wikstrom, 2006a). Following on from this, my contention in this thesis is that it is those neighbourhoods where deprivation is high *and* exposure to inequality is high that are potentially most at risk of (or most vulnerable to) suffering high rates of criminal offending and other forms of social disorder. As noted above, the empirical analyses of spatial patterns of deprivation in Chapter 3 and spatial patterns of exposure to inequality (amongst the poor population) in Chapter 5 highlighted a broad negative relationship at datazone level. However, some neighbourhoods were shown to have high levels of deprivation and also high levels of exposure, for example the township of Alexandra in Johannesburg. In Chapter 6 I therefore focused explicitly on identifying the spatial patterning of neighbourhoods with these twin stressors of high deprivation and high exposure to inequality.

I began by describing the method I had adopted for building a measure of community 'intensity' of exposure to inequality, this being the simple mathematical product of the deprivation rate and exposure score (constructed separately for each of the four domains of deprivation/exposure derived from the SAIMD 2001). I then analysed the results across all datazones in the country using the 'intensity' measure based upon the income and material deprivation domain of the SAIMD 2001. I introduced the results of the intensity measure by looking at the datazone level map of Johannesburg and returning to the discussion of Sandton, Alexandra and Soweto. In Alexandra the level of intensity of exposure was seen to be very high (due to both deprivation and exposure being very high), while in Soweto the level of intensity was also high but somewhat lower than Alexandra (due to deprivation and intensity also being high, but not as high as in Alexandra), while in Sandton the level of intensity was very low (because, despite the exposure score being very high here, the rate of deprivation was very low by South African standards). In other words, although poor individuals living in Sandton are very highly exposed to inequality, there are relatively few poor people actually living in Sandton and therefore the *community level* intensity of exposure is low.

When looking across the datazone intensity scores by province, it was evident that all nine provinces contained datazones with low intensity scores (i.e. lowest national decile) and all provinces except Limpopo contained datazones with high intensity scores (i.e. highest national decile). However, the upper extent of the intensity scores within the Western Cape and Gauteng was much higher than in any of the other seven provinces, indicating that this is where the most extreme levels of intensity are to be found. The major metropolitan areas of Cape Town, Ekurhuleni, Johannesburg, eThekwini and Tshwane accounted for almost all (approximately 90%) of the highest intensity national decile of datazones, with the remainder of this decile group spread over 32 other municipalities. Cape Town, Ekurhuleni, Johannesburg, eThekwini and Tshwane also ranked within the top ten municipalities nationally according to the proportion of their datazones in the highest intensity decile, providing further confirmation that community intensity of exposure is highest in these key metropolitan areas.

I concluded the analyses of community intensity of exposure to inequality by again focusing on the metropolitan datazones and using the four separate intensity measures derived from the four domains of the SAIMD 2001 (income and material deprivation, employment deprivation, education deprivation, and living environment deprivation). Once again I found strong positive correlations between the metropolitan datazones on these four measures and I proceeded to construct a single composite intensity measure by combining the four separate intensity measures using weights derived through factor analysis. When the composite intensity measure was ranked across all metropolitan datazones I found particular concentrations of high intensity metropolitan datazones within Cape Town, however, I also found particular concentrations of low intensity metropolitan datazones within Cape Town. I discussed again how the patterns of intensity are related to the underlying components of deprivation rate and exposure score and how the combination of these two factors generated the intensity scores observed within Cape Town and the other municipalities.

In short, I demonstrated that some datazones are indeed characterised by high levels of deprivation and high levels of exposure to inequality. By developing a new measure of community intensity of exposure to inequality I was able to map and analyse the spatial patterning of the high intensity and low intensity areas across South Africa and then across the metropolitan areas only. I was therefore able to address the specific research question tackled in Chapter 6. These findings also contribute to the main overarching research question in that, as I argue in Chapter 7, the intensity measure may be an important determinant of people's social attitudes over and above the separate effects of deprivation rate and exposure score. Furthermore, based upon the international literature there is a basis to assume that the high intensity neighbourhoods may be those most at risk of crime and/or other forms of social unrest[54]. Although I do not look explicitly at the spatial relationships between deprivation, exposure, intensity and crime/unrest in this thesis, this is a possibility for future research.

The fifth and final research sub-question was: "To what extent are people's attitudes to inequality and options for redress associated with their lived experience of inequality?" and this was the focus of Chapter 7. The reason why this research question was included was to provide one example of how my new exposure measures could be used to explore associations with other socio-economic outcomes. As discussed in Chapter 2, it has been contended that a person's lived experience of inequality may

[54] As noted earlier, I do not regard my exposure or intensity measures as being superior to the conventional poverty or deprivation measures. Furthermore, I do not argue that my exposure or intensity measures should necessarily be used as the basis for any sort of financial resource allocation. The purpose of my exposure and intensity measures is to help better understand the way that people's lived experience of inequality may shape other socio-economic outcomes, both at the individual level and the area level.

shape their attitudes towards themselves, others, broader society and the government and these attitudes are the direct causes of crime (e.g. Wikstrom, 2006a). I also discussed how it has been contended that these kinds of social attitudes may have a causal effect on people's mental wellbeing and wider health outcomes (e.g. Pickett and Wilkinson, 2015). In light of this, the focus in Chapter 7 was on exploring the relationships between exposure to inequality and attitudinal outcomes. I identified two particular attitudinal outcomes for the chapter: attitudes towards levels of inequality in South Africa; and attitudes towards the government's role in income redistribution. While these two outcomes do not speak *directly* to the issues of crime or poor health, they are nevertheless vital indicators of societal attitudes towards inequality and towards options for redress. As such, I argue they provide an important first step in understanding the potential relationships between the lived experience of inequality and other socio-economic outcomes. Having reviewed the international literature it was apparent that this form of analysis has not been undertaken elsewhere and therefore the research presented in this thesis makes an innovative contribution to the literature in South Africa and internationally.

For the purposes of the analysis presented in Chapter 7, I obtained a special extract of the South African Social Attitudes Survey (SASAS) at individual respondent level that contained information on the respondents' home address location. Using the address location data I merged in selected datazone level measures, specifically the datazone deprivation rate from the income and material deprivation domain of the SAIMD 2001, and the accompanying exposure scores also based on the income and material deprivation domain. With regards to the exposure scores, I merged in both variants: the exposure of poor to non-poor (which has been the primary focus in this thesis), and the exposure of non-poor to poor (which I have included in the relevant appendices to this thesis). I proceeded to split the SASAS dataset into two subsets: one containing the ‘poor’ population and another containing the ‘non-poor’ population based on an internally constructed asset index. I reviewed the SASAS questionnaire and identified the two dependent variables with which to measure outcomes and I selected a number of control variables which had been shown in separate analyses to be conceptually and empirically related to the chosen dependent variables. I proceeded to develop four separate statistical models using multilevel ordered logit regression: these four models related to the two dependent variables for each of the poor and non-poor subsets. The analytical focus was on assessing whether there were any statistically significant associations between each of the dependent variables and three datazone level variables: the deprivation rate, the exposure score, and the interaction term between the deprivation rate and exposure score (which equates to the intensity measure developed in Chapter 6). All other independent variables were included as controls only.

As per the earlier analysis in this thesis, the primary concern in Chapter 7 was on the exposure to inequality amongst the poor population of South Africa. As such, I focused the analysis on the two statistical models relating to the poor subset of respondents (with the non-poor results provided in Appendix D). The results of the models suggested that exposure to inequality may indeed play a role in shaping people’s social attitudes on inequality. In terms of inequality aversion, there was some evidence that less advantaged sub-groups of the population become progressively *less* averse to income inequality as the level of exposure to inequality increased. Whilst these results were somewhat unexpected, they may potentially be due to disillusionment amongst the *more advantaged* sub-groups (i.e. the reference groups against which the more disadvantage sub-groups were compared in the models) in relation to ever bridging the gap between the ‘haves’ and ‘have nots’. In terms of redress measures, there was some evidence that the less advantaged sub-groups became progressively more supportive of government’s responsibility for income redistribution as the level of exposure to inequality increased. These results are in line with the earlier suggestion that less advantaged groups of the population may be more supportive of redress measures, particularly for those who have high exposure to inequality.

Although the statistical model results provide new empirical insights into the associations between exposure to inequality and the two selected dependent variables, I acknowledge that the findings are potentially hampered by the skewed distribution of responses on these two dependent variables. More specifically, there is generally a very strong consensus amongst the population of South Africa that income inequality is too high and that the government has a responsibility to reduce it by redistributing income from those with high incomes to those with low incomes. This manifests as a relatively limited amount of discrimination between response intervals on the dependent variables, as the responses are clustered around the ‘agree/strongly agree’ end of the scale. Had the responses been more evenly distributed over the response categories the model results may have been more conclusive. As I note below in the final part of this concluding chapter, SASAS is an annual nationally representative survey and so as a priority for future research it would be desirable to include a broader set of attitudinal questions that might provide greater discrimination between response intervals and therefore greater clarity in the modelling results.

In summary, it was possible to use the exposure measures to explore associations between people’s lived experience of inequality and their attitudinal outcomes concerning inequality and options for redress, and this is represents a unique contribution to the literature on exposure to inequality. My analysis showed there is evidence to suggest that people’s attitudes towards inequality and options for redress are indeed associated with their lived experiences of inequality. Finally, and as I noted earlier, this particular modelling exercise is just one example of how my exposure measures could be used as independent variables in statistical models. There are many other ways in which the measures could be used, and I discuss this in the final section of this conclusion.

The overarching research question was: “How does the lived experience of inequality vary spatially across South Africa and how is this associated with people’s attitudes towards inequality?” The exposure measures enable us to move beyond the conventional indices of income inequality at broad national level and look instead at how people’s experience of inequality is contoured by the localised geographical settings in which they carry out their daily activities. I would argue that this is a major methodological advance and represents an important contribution to knowledge that is relevant to the specific case of South Africa and more broadly internationally. The methodological specification of the exposure measures was designed to reflect the context of day-to-day life in South Africa rather than simply adopt a developed country approach. I would argue that this too represents an important contribution to knowledge in South Africa and internationally. The linkage of the exposure measures (along with deprivation rates and intensity scores) to individual level SASAS data enabled me to explore associations with attitudinal outcomes that have relevance for inequality research in South Africa and which may represent the first step in assessing linkages between exposure to inequality and other socio-economic outcomes. My review of the literature suggests this is the first time local measures of exposure to inequality have been used as covariates to explain people’s social attitudes to inequality. As such, this research makes a number of unique contributions to the evidence base concerning inequality in South Africa and internationally.

In the following final part of this concluding chapter I discuss the main limitations of this research as acknowledged in earlier chapters and propose some priorities for future development of this research topic.

8.3 Limitations and priorities for future research

Although each component of this thesis is innovative and makes a unique contribution to knowledge, I also readily acknowledge the limitations of the research. In each preceding chapter of this thesis I have noted points that may be regarded as limitations as they arose during the course of the research. In this final section of the thesis I will highlight the primary limitations of my research, and make recommendations on how these may be addressed through future research.

The primary limitations of this thesis can be categorised into three main themes: (i) data time points; (ii) processes/dynamics of exposure to inequality; and (iii) associated attitudinal outcome measures (i.e. attitudes to inequality and redress). I will discuss each of these three themes in turn.

With regards to data time points, the deprivation data upon which the exposure measures were based relate to the year 2001, which means the analyses of the spatial patterns of deprivation and exposure are now quite out of date. Whilst this does not weaken the innovative nature of the methodological development process I adopted, it does mean that the utility of the results for informing evidence based policy making is somewhat limited. It also represents a limitation in terms of the disconnect between the time point of the exposure measures and the time point of the SASAS survey data used in Chapter 7. However, as other commentators have shown, the broad spatial patterning of deprivation across South Africa has not changed dramatically between 2001 and 2011, and so it is doubtful that the exposure measures for 2011 would be dramatically different to the results presented here (Noble et al., 2013). As I discussed in Chapter 2 and again in the earlier part of this conclusion, sub-national deprivation data (of the form required for exposure indices) are now available based on the 2011 Census (Noble et al., 2013; Noble et al., 2014), but these are at ward level, not datazone level. I am aware of separate research that is underway to develop a set of datazones for 2011 and once this has been completed it will become possible through future research to produce the exposure measures on a 2011 time point.

A further potential limitation concerning data time points may be that the analysis of associations between exposure and attitudinal outcomes uses a static measure of exposure (i.e. at just one point in time). Whilst static measure of exposure is appropriate as an explanatory factor, it would also be worthwhile to consider whether a measure of the *dynamics* of changing exposure might represent an additional factor associated with attitudinal outcomes. This is because people's attitudes are likely to be shaped by their experiences of inequality over a period of time (rather than at a single static snapshot) and the socio-spatial environments that shape people's exposure may change quite considerably over time when considered at a detailed geographical level. Upon completion of the 2011 exposure measures it would be good to test whether the changing dynamics of exposure are associated with other socio-economic outcomes, independently of the static measure of exposure. For example, it might be possible to add an additional independent variable to the models to reflect the magnitude of change in the neighbourhood exposure score between 2001 and 2011. Additionally, if future rounds of SASAS can incorporate a question asking people where they were living at the time of the 2001 Census (as well as ascertaining their current home residence) then it might be possible to take into account a person's exposure in a past residential location as well as exposure in the current neighbourhood.

With regards to the processes/dynamics of exposure to inequality, I acknowledge that my methodological definition of the 'lived experienced of inequality' may be regarded as quite narrow. The first criticism of this approach is that the methodology restricts people's experience of inequality to occurring within certain, quite localised, geographical bounds. It therefore excludes the potential for people to experience inequality in other parts of the country which they may visit occasionally (particularly in the context of widespread internal migration) and it excludes the potential for people to experience inequality through the national and international media. Whilst I acknowledge that the media may potentially play an important role in shaping people's lived experience of inequality, at the present time no suitable data are available in South Africa to enable me to analyse this empirically. I therefore recommend that a qualitative study would add considerable value in order to explore with individuals of different backgrounds and from different geographical locations the various factors that might shape their awareness of and exposure to inequality within South Africa, including the internet and the media.

A qualitative study would also provide valuable information on the extent to which people experience inequality through occasional trips to more distant parts of the country (i.e. outside of the defined geographical bounds specified in the exposure methodology). A potentially crucial component of this relates to internal migration, particularly between the urban and rural areas. Other research (e.g. Makiwane and Chimere-Dan, 2010; Todes et al., 2010; Tomlinson et al., 2003; Turok and Borel-Saladin, 2014) has documented the considerable levels of migration from the rural areas into the urban areas (particularly Cape Town and Gauteng) in order to seek work, and that this dynamic often involves a flow of return migration back to the rural areas. As such, I acknowledge that the socio-spatial spheres that shape people's lived experience of inequality may be much more complex than the defined geographical bounds that are necessary for the specification of the exposure measures. However, at present, no suitably detailed data exist on the complex migration patterns to enable me to explore this empirically. This is a priority for future research, and could involve the use of the national panel dataset, the National Income Dynamics Study.

Despite the potential limitations regarding the role of travel outside of the defined geographical bounds and the role of the media, the setting of some form of geographical bounds is justified in terms of defining the socio-spatial environments in which people carry out their *routine daily activities*. As such, the general principles underpinning the exposure measures are sound as a primary determinant of the lived experience of inequality. However, I acknowledge that my assumptions regarding the specification of the geographical bounds and the deprivation-adjusted weighting function are based on researcher judgement only, due to the lack of suitable empirical data to inform these parameters. For instance, the adjustment to the distance decay function is based on *deprivation* differentials on the basis that more affluent areas (proxied by lower deprivation rates) may exert greater 'pull' factors, such as offering greater labour market opportunities. However, this may not accurately reflect the varied labour market opportunities in the central business districts of major cities as these areas may not have particularly low deprivation rates (for instance, residents of the traditional central business district in Johannesburg are typically relatively deprived, as can be seen from Panel A of Figure 5.2 in Chapter 5). Using a deprivation differential approach to modifying the decay weights does not adequately reflect the 'pull' of the central business district in terms of employment opportunities. A potential improvement would be to apply a distance decay adjustment that takes better account of the concentration of labour market opportunities in areas such as central business districts. There is therefore a need for future research to explore people's routine daily travel patterns and activity spheres using a combination of qualitative and quantitative research methods. The aim here would be to better understand where people travel to on a routine basis and for what purposes, and how their travel decisions are shaped by the combination of needs, opportunities and barriers.

My recommendations for future research on the processes/dynamics of exposure to inequality also offer the possibility of assessing commonalities or differences between various population sub-groups. For instance, routine travel patterns may differ systematically according to demographic characteristics such as age, sex or population group. Similarly, people may accept longer travel distances in some geographical parts of the country than other parts. If it is possible to define the exposure parameters separately for different population sub-groups then it would become possible to develop a series of exposure measures with each one shaped in terms of a particular sub-group. When looking for associations between exposure and outcomes such as attitudes to inequality one could then use the appropriate exposure measure for each respondent dependent on the respondent's personal characteristics.

With regards to the limitations of the associated attitudinal outcome measures explored in this thesis (i.e. the overwhelming consensus on response categories leading to little discrimination between respondents), a priority for future research is to both refine the existing survey questions and to develop new survey questions in order to more comprehensively understand people's attitudes towards inequality. For instance, it might be informative to construct a series of vignettes relating to differing household scenarios and the role of the state in supporting those household members and then asking SASAS respondent whether they agree or disagree with policy options. In addition, respondents could be asked a range of questions in the more conventional format, such as giving their level of agreement/disagreement to statements such as: 'I think that rich people ought to pay more taxes and that the money generated should be used to finance social transfers to the poor'. The preferred approach to designing new SASAS questions would be to draw upon emerging findings from the proposed qualitative research which should illuminate people's views around inequality in ways that cannot be achieved through quantitative surveys. With an extended suite of inequality-related response variables it will be possible to re-analyse the associations between people's lived experience of inequality and their attitudes to inequality.

Finally, as I have stressed above, the exposure measures have the potential to be used to explore associations with a wider set of outcomes than attitudes towards inequality. For instance, a priority for future research is the application of the exposure measures to examine potential associations between the lived experience of inequality and crime-related outcomes, including both individual level outcomes (such as fear of crime) and area level outcomes (such as crime rates). As both levels of deprivation *and* levels of inequality are widely thought to be important drivers of crime, it would seem that the intensity measure developed in this thesis may offer a valuable new empirical measure with which to explore spatial patterns of crime. My final recommendation for future research is therefore an emphasis on the intensity measure as an explanatory factor for social outcomes such as crime and unrest.

Appendix A: The input of others into this thesis

All the research presented in this thesis is my own work. However, in the course of my doctoral studies I have published a range of related outputs, including peer reviewed journal articles and project reports, some of which have been jointly authored. Furthermore, my decisions regarding which methodological and analytical approaches to adopt for my research were partly shaped through discussions with academic and policy experts from South Africa and the UK. In this Appendix I clarify the extent to which my doctoral work presented in this thesis benefited from the contributions of others.

My doctoral studies have been undertaken in parallel with my employment as a Senior Research Fellow in the Centre for the Analysis of South African Social Policy (CASASP) at the University of Oxford (from 2002 to August 2015) and then at Southern African Social Policy Research Insights (SASPRI) (from 2015 to present).

During the early stages of my doctoral study I submitted a research proposal through the University of Oxford to the Economic and Social Research Council (ESRC) under the ESRC Pathfinder grant scheme. This application was successful. The title of the grant was 'Exploring the relationship between spatial inequality and attitudes to inequality in South Africa' (grant ref: ES/I034889/1). I was the Principal Investigator on the grant. In addition to myself, the project research team consisted of Professor Michael Noble (CASASP, University of Oxford) and Mr Ben Roberts and Dr Temba Masilela (Human Sciences Research Council (HSRC) in South Africa). We also appointed Dr Hope Magidimisha as a HSRC research intern to assist us. The project had three main components. Component 1 involved the construction of a spatial measure of the lived experience of inequality; Component 2 (which ran concurrently with

Component 1) involved analysing the South African Social Attitudes Survey (SASAS) to gain a better insight into people's attitudes towards inequality and options for redress; and Component 3 (which took place upon completion of Components 1 and 2) involved linking the spatial inequality measures to the SASAS survey data to test for associations between experience of inequality and attitudes to inequality. I took responsibility for Components 1 and 3, while Ben Roberts (HSRC) took responsibility for Component 2, with some assistance from Dr Magidimisha.

The measures of exposure that I developed in Chapter 4 of this thesis represents the outputs from Component 1 of the ESRC Pathfinder grant. This was entirely my own work. As part of our commitments to ESRC, we were required to write one peer reviewed journal article concerning Component 1 of the project. This paper was published in the *South African Geography Journal* as McLennan et al (2016) and included editorial contributions by Professor Michael Noble and Dr Gemma Wright.

Ben Roberts's work on Component 2 of the ESRC Pathfinder grant resulted in a paper published in *Social Indicators Research* as Roberts (2014). Although not of direct relevance to my thesis, Roberts's analyses provided a valuable foundation for the work I undertook in Component 3 of the ESRC Pathfinder.

The statistical modelling presented in Chapter 7 of this thesis represents the outputs from Component 3 of the ESRC Pathfinder grant. This was entirely my own work, although I called upon external statistical advice from senior statisticians in the University of Oxford's Department of Statistics (Dr Dan Lunn and Dr George Nicholson). Three major recommendations emerged from my discussions with the statistical experts: (i) my modelling approach should adopt the full ordered logit model; (ii) multilevel models were required due to the geographical stratification inherent in the SASAS sampling methodology; and (iii) the modelling should be done using the R statistical software package due to the computational complexity of the task. During my work on the statistical modelling element of this thesis (and for Component 3 of the ESRC Pathfinder project) I therefore consulted with the statistical experts on statistical methodology and on implementing the method using R (as I had hitherto never used R). Separately, the work of Roberts referenced above made a contribution to the selection of control variables used in the models and this contribution was further enhanced through project team discussions on advantages and disadvantages of different control variables. Although we were not required to publish a peer reviewed journal article on the subject of Component 3 of the Pathfinder, I did include a discussion of the methodology and results in the final project report (McLennan et al., 2014), which was made freely available through the CASASP website and now also through the SASPRI website (www.saspri.org). Although I wrote the relevant chapter of the final project report on the modelling exercise, this also benefited from some editorial contributions from other team members.

As noted above, the conceptual and methodological developments of my doctoral work were also shaped in part by personal interactions with leading academics and policy makers in public forums. For instance, I have presented my work to academic audiences at the University of Oxford (Department of Social Policy and Intervention) and the Human Sciences Research Council. Furthermore, this research has also been the focus of two dedicated workshops in South Africa convened by the South African Department for Science and Technology, one primarily aimed at senior policy makers and another primarily aimed at academics. Each of these public engagements has contributed to my thinking concerning the conceptualisation and measurement of my doctoral research.

Appendix B: Technical annexes on rejected residential segregation measures

Evenness

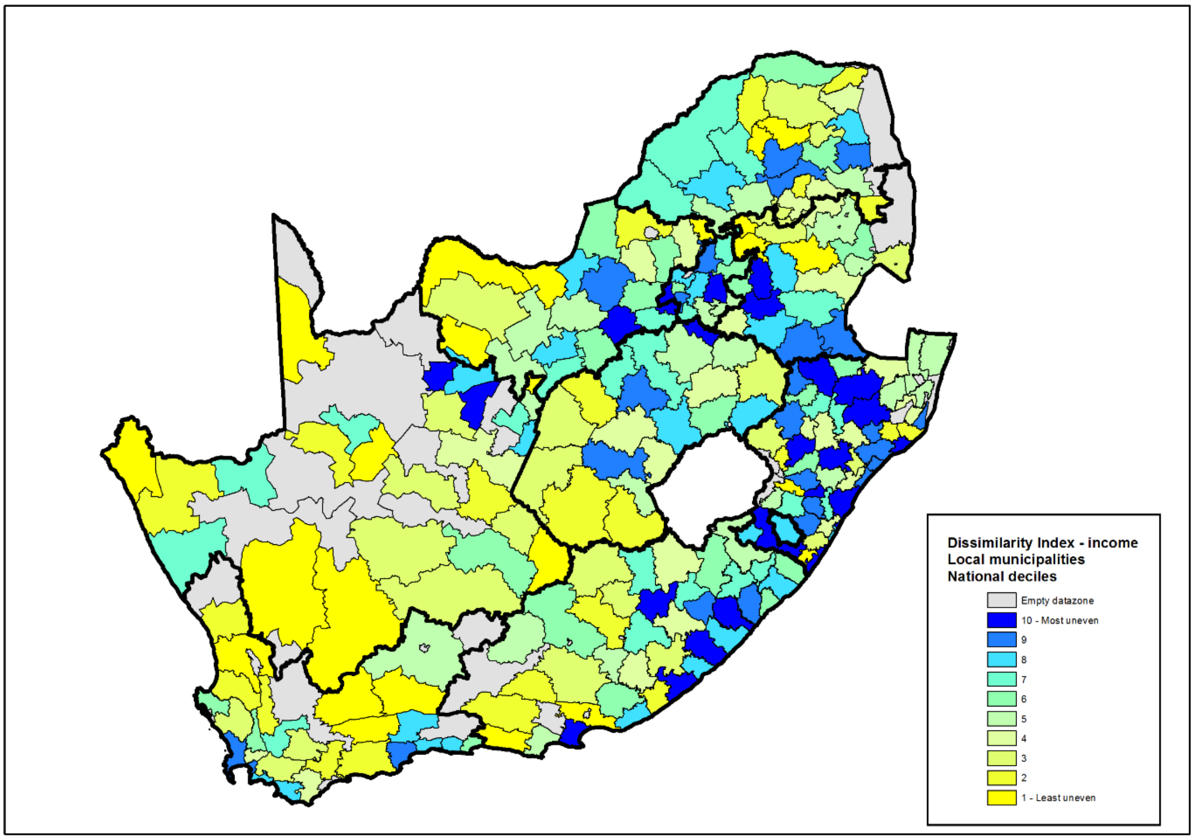
Massey and Denton recommend that the Dissimilarity Index *D* is the most appropriate measure for the Evenness dimension.

Equation (a) shows one commonly used formula for this dissimilarity index:

$$D = \sum_{i=1}^{n} \frac{t_i |p_i - P|}{2TP(1-P)} \qquad \text{(a)}$$

where t_i is the total population of areal unit i, p_i is the proportion of the total population of areal unit i that is defined as poor, T is the total population of broader geographical area, and P is the proportion of the total population of the broader geographical area that is defined as poor, where the broader geographical area is subdivided into n areal units. Figure B.1 shows a municipality level map of Dissimilarity Index scores for the whole of South Africa, with the colour-coding relating to national deciles of municipalities on this measure. The municipality level Dissimilarity Index scores are therefore constructed by aggregating data from datazone level using the above equation.

Figure B.1: Dissimilarity Index – municipality level – income – national deciles



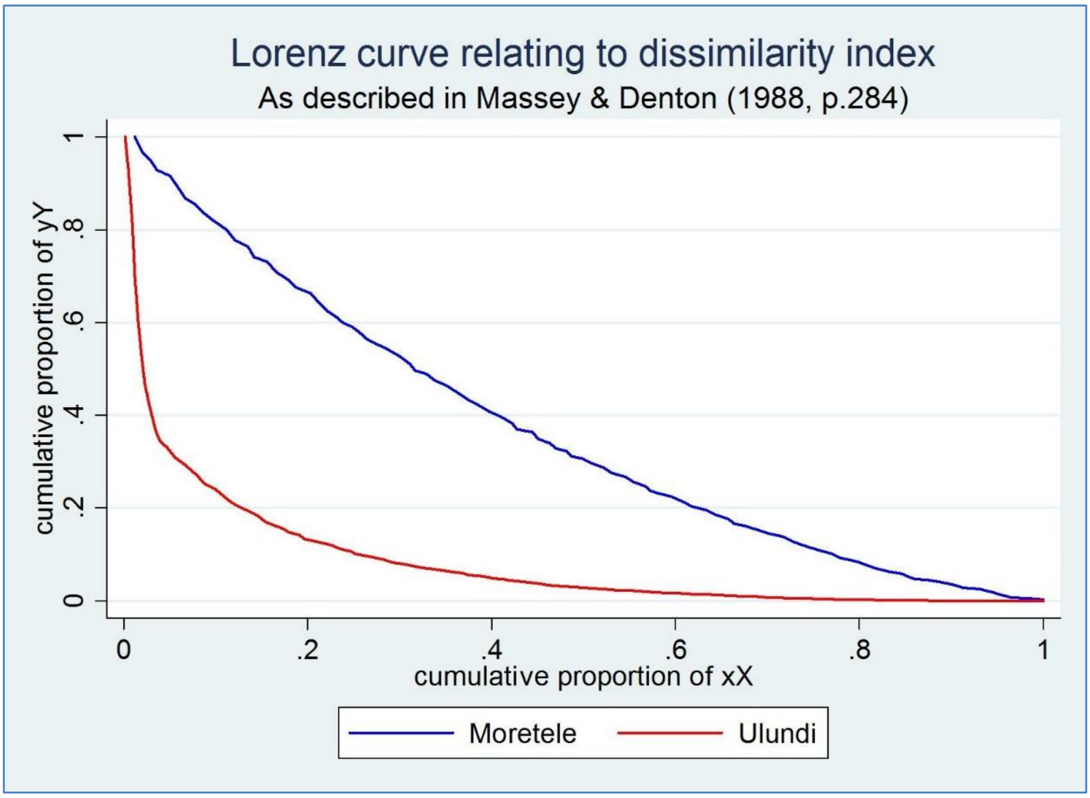
It is evident from Figure B.1 that the municipalities that score highest on the Dissimilarity Index are scattered across the country, but with a concentration of these municipalities located in KwaZulu-Natal province. Indeed, of the 23 municipalities in the most uneven decile on this measure, 12 are in KwaZulu-Natal, four are in Eastern Cape, two in Northern Care and the remaining five spread across the rest of the country. Ulundi municipality in KwaZulu-Natal province registers the highest score on the Dissimilarity Index. With regard to the 24 municipalities that constitute the least uneven on this measure, seven are located in Northern Cape (of which one straddles Northern Cape and North West), four in each of Western Cape and North West (excluding the one that straddles with Northern Cape), three are in Mpumalanga, and two are located in each of Eastern Cape, KwaZulu-Natal and Limpopo. The municipality of Karoo

Hoogland in Northern Cape provinces registers the lowest score on the Dissimilarity Index.

As Massey and Denton (1988) point out, the Dissimilarity Index can be derived from the Lorenz curve when this is configured to plot the cumulative proportion of poor population against the cumulative proportion of non-poor population across areal units within a municipality, with units ordered from smallest to largest poor proportion. *D* represents the maximum vertical distance between the curve and the diagonal line of evenness. Figure B.2 shows the Lorenz curves for the poor/non-poor classification of datazones within two selected municipalities: Ulundi in KwaZulu-Natal (which, as noted above, has the highest score on the dissimilarity index) and Moretele in North West province (which has a low score on the dissimilarity index relative to the rest of the country)[55].

[55] Although Karoo Hoogland registers the lowest score, this municipality contains only five Datazones and so the results are not well suited to presentation on a Lorenz curve for this purpose. Moretele registers the fourth lowest score on the Dissimilarity Index and contains a much larger number of Datazones thus making it more appropriate for this particular analysis.

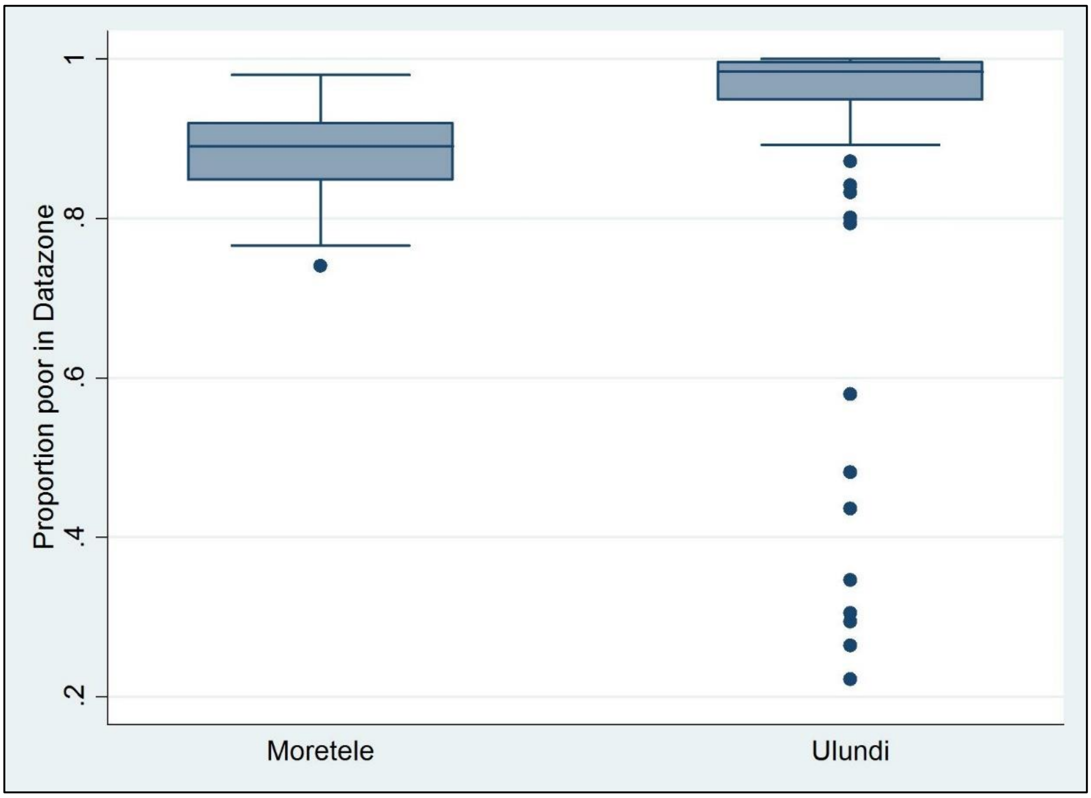
Figure B.2: Lorenz curves for Ulundi and Moretele



Note: the x-axis represents the cumulative proportion of ‘poor’ subgroup amongst the constituent datazones in the two respective municipalities; the y-axis represents the cumulative proportion of the ‘non-poor’ subgroup amongst the constituent datazones in the two respective municipalities.

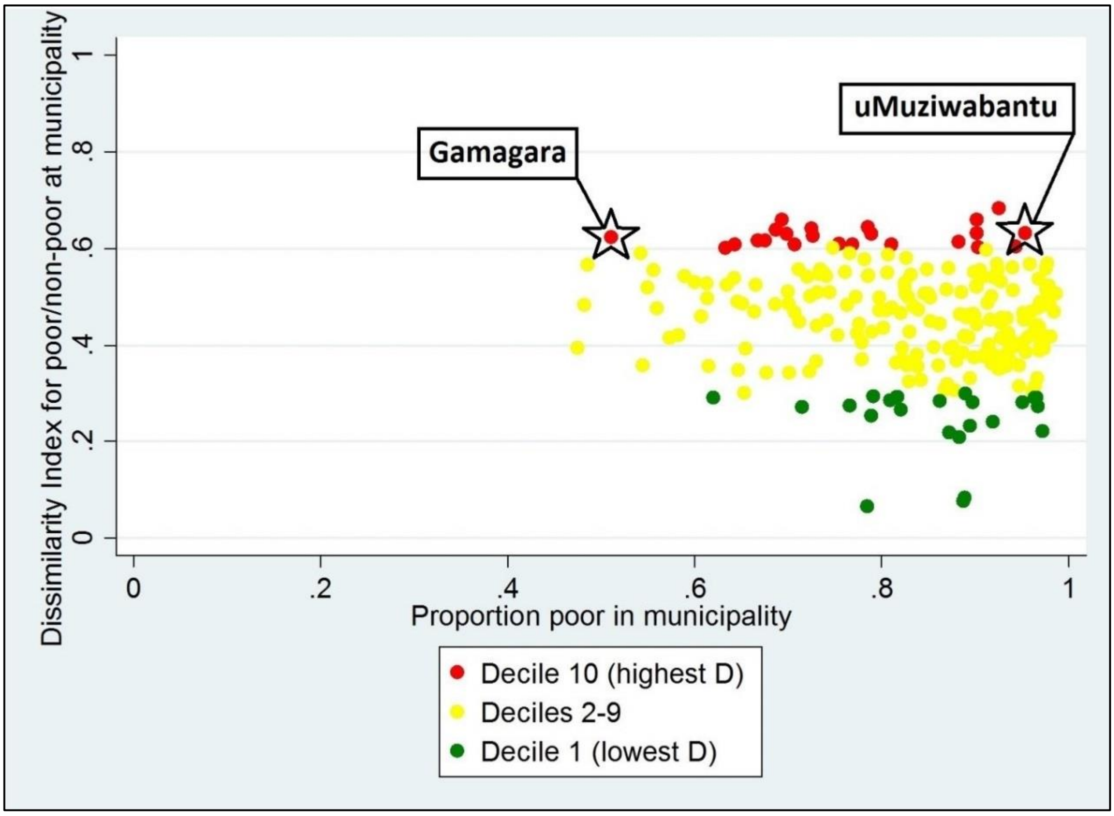
The Lorenz curves for Ulundi and Moretele depict starkly different distributions of datazone values within the respective municipalities, with the curve for Moretele being close to the diagonal line of evenness but the curve for Ulundi being a substantial departure from that line of evenness. These Lorenz curves demonstrate the relatively high degree of uniformity in the datazone level poverty rates within Moretele, contrasted with Ulundi where a relatively small number of datazones contain vastly different (specifically lower) poverty rates than the municipality as a whole. The boxplot presented in Figure B.3 presents the actual datazone level poverty rates for these two municipalities and shows the contrasting levels of variability: Moretele is composed of datazones with relatively similar levels of poverty, whereas there are a number of datazones in Ulundi with notably lower poverty rates than the vast majority of the municipality. Further analysis of the underlying poverty rate data (mapped at datazone level but not presented here) reveals that the datazones with the lowest levels of poverty are located in the town of Ulundi itself.

Figure B.3: Boxplot of Datazone level poverty rates within Moretele and Ulundi municipalities



To further unpick this point, it is possible to compare two municipalities which have very similar scores on the Dissimilarity Index but differing poverty rates, namely Gamagara in the Northern Cape and uMuziwabantu in KwaZulu-Natal. Figure B.4 shows the relationship between municipality level poverty rate (on the x axis) and the municipality level Dissimilarity Index (on the y axis), with the highest unevenness national decile depicted using red dots, the lowest unevenness national decile depicted using green dots, and the intermediate eight deciles depicted using yellow dots. Garagara and uMuziwabantu both exhibit relatively high scores on the Dissimilarity Index, but they contrast in terms of their poverty rates, with Gamagara exhibiting one of the lowest poverty rates of all municipalities in South Africa and uMuziwabantu exhibiting one of the highest poverty rates.

Figure B.4: Scatterplot of municipality level Dissimilarity Index scores against municipality level poverty rates



The dimension of evenness, including the constituent measure of the Dissimilarity Index, is rejected for the reasons detailed in Chapter 4.

Concentration

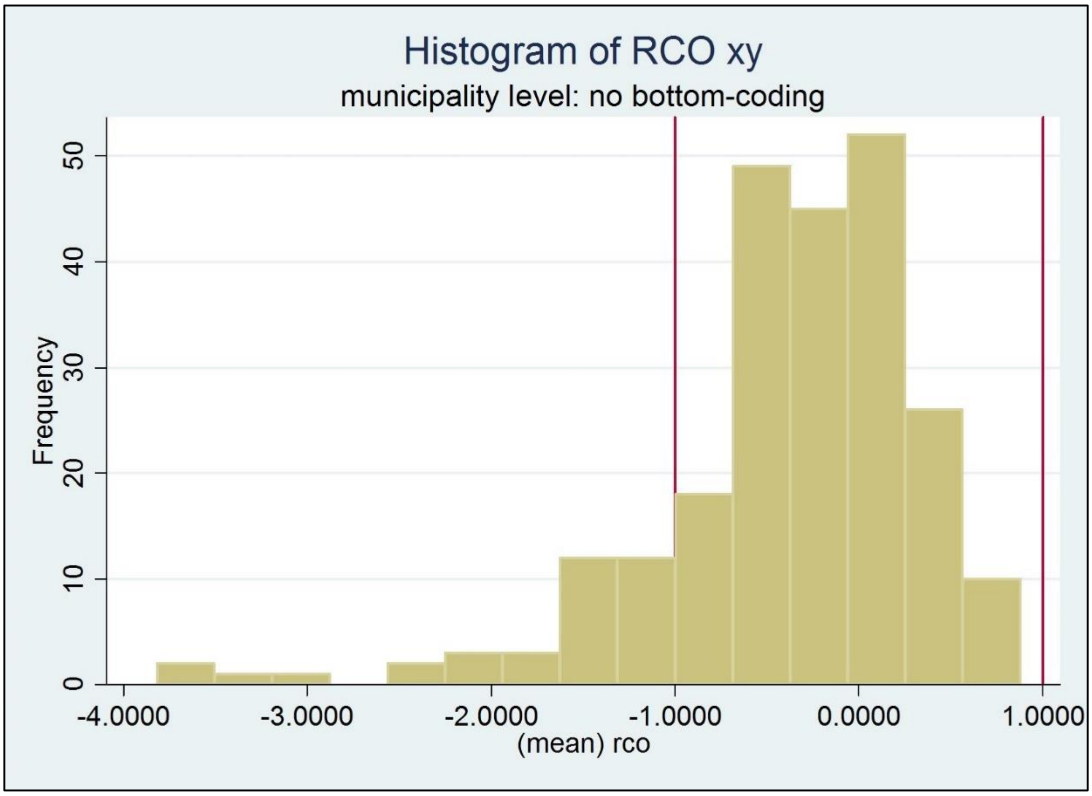
The Relative Concentration Index (RCO) is the recommended measure of concentration advocated by Massey and Denton (1988). The RCO for the degree to which population sub-group X (in my case, the poor) is concentrated compared to population sub-group Y (in my case, the non-poor) is expressed as:

$$RCO_{xy} = \frac{\left(\frac{\sum_{i=1}^{n} x_i a_i / X}{\sum_{i=1}^{n} y_i a_i / Y}\right) - 1}{\left(\frac{\sum_{i=1}^{n_1} t_i a_i / T_1}{\sum_{i=n_2}^{n} t_i a_i / T_2}\right) - 1}$$

where the areal units within a given city/municipality are ordered according to geographical size, from smallest geographical area to largest geographical area, a_i is the land area of unit i, n_1 and n_2 refer to different points in the rank ordering of areal units from smallest to largest: n_1 is the rank of the areal unit where the cumulative total population of areal units equals the total poor population of the city/municipality, summing from the smallest unit up; n_2 is the rank of the areal unit where the cumulative total population of units equals the non-poor population totalling from the largest unit down. T_1 equals the total population of areal units 1 to n_1, and T_2 equals the total population of areal units from n_2 to n. The terms x, y and t represent the numbers of poor, non-poor and total population in each area i, whilst X and Y represent the numbers of poor and non-poor in the city/municipality as a whole (Massey and Denton, 1988; Massey and Denton, 1998). Massey and Denton state that the RCO "takes the ratio of X members' to Y members' concentration and compares it with the maximum possible ratio that would be obtained if X were maximally concentrated and Y minimally concentrated, standardizing the quotient so that the index varies between -1.0 and 1.0. A score of 0 means the two groups are equally concentrated in…space. A score of -1.0 means that Y's concentration exceeds X's to the maximum extent possible, and a score of 1.0 means the converse" (Massey and Denton, 1988, p.291). In other words, the RCO is designed to measure the share of geographical space occupied by group *X* compared to group *Y*.

Figure B.5 shows a histogram of municipality level RCO scores representing the degree to which the poor population (*X*) is spatially concentrated compared to the non-poor population (*Y*). Figure B.6 shows an equivalent histogram representing the degree to which the non-poor population (*Y*) is spatially concentrated compared to the poor population (*X*).

Figure B.5: Histogram of RCO xy with no bottom-coding

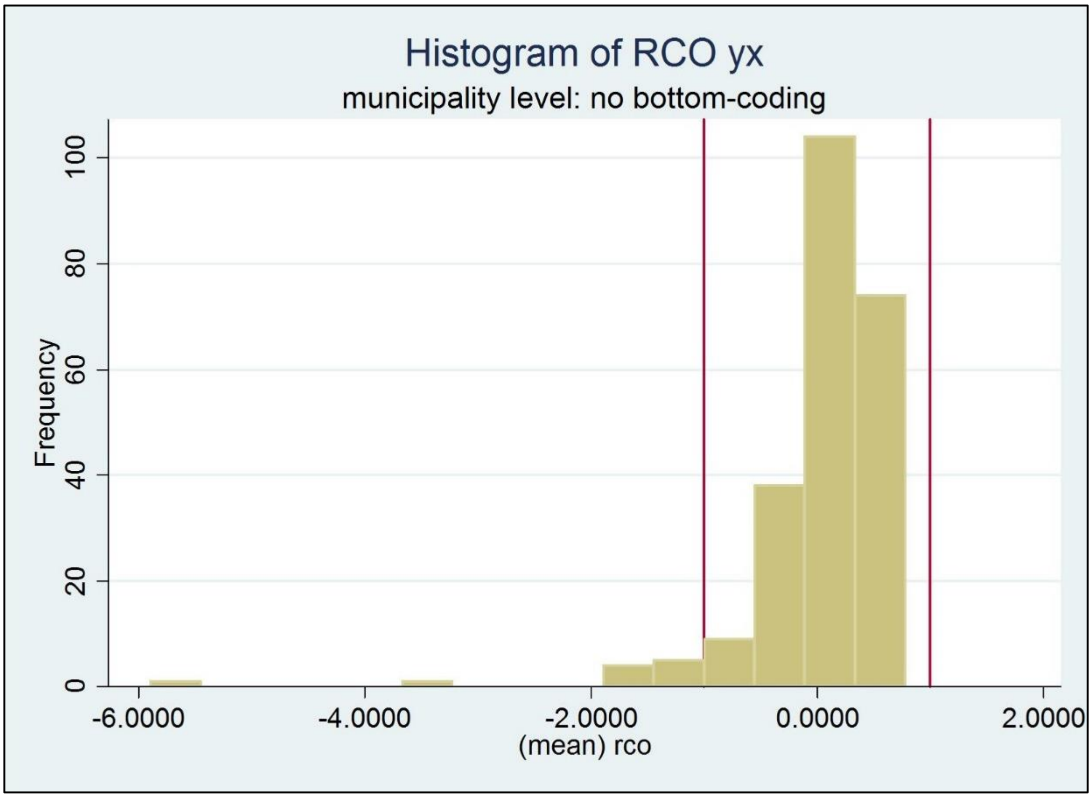


Of the 235 municipalities in South Africa, 37 have RCOxy scores lower than -1.0. In Umtshezi, Abaqulusi and Ulundi in KwaZulu-Natal, and King Sabata Dalindyebo in the Eastern Cape, the RCOxy value computed is less than -3.0. As cited above, Massey and Denton stated that in order for an municipality to generate a RCO value of less than -1.0, two conditions had to be satisfied: first, the municipality must have a “very small” number of X members; and, second, the areas in which the X members live must be “very large” in a geographical sense (Massey and Denton, 1998, p.1126-1127). However, as was demonstrated in Chapter 3 of this thesis, none of the municipalities in South Africa contain very small poor populations, which brings into further question the configuration of Massey and Denton’s RCO index. In King Sabata Dalindyebo municipality, for example, the population-weighted average areal size of datazones inhabited by the poor population is 17.2 km^2, while the equivalent value for the non-poor population is 4.7 km^2. If the poor population was maximally concentrated the average areal size of datazones inhabited by the poor would be 10.1 km^2, while if the non-poor population was minimally concentrated the average areal size of datazones inhabited by the non-poor would be 69.5 km^2. The numerator of the RCOxy as specified by Massey and Denton would therefore be represented as [(17.2 / 4.7) – 1 = 2.66] while the denominator would be represented by [(10.1 / 69.5) – 1 = -0.85]. The overall RCOxy for King Sabata Dalindyebo consists of dividing 2.66 by -0.85 to give a final RCOxy value of -3.11. To test whether the RCO values for municipalities such as King Sabata Dalindyebo were due solely to the presence of very large (in a geographical sense) datazones, I repeated the RCO computation process but this time imposing a series of ever-more stringent constraints on the spatial size of datazones that were permitted to be included in the calculation. RCOs were computed for ten different iterations, with the first iteration dropping all datazones that were 50 km^2 or larger in size; the second iteration dropping all datazones that were 45 km^2 or larger in size; the third dropping all datazones 40 km^2 or larger in size, and so on a 5 km^2 stepped reduction until the final iteration consisted of dropping all datazones that were 5 km^2 or larger in size. The results of this iterative testing for the King Sabata Dalindyebo municipality revealed that the RCOxy value remained consistently above the -1.0 bound in each iteration from 50 km^2 through to 15 km^2, after which the RCOxy values with the 10km^2 and 5km^2 constraints resulted in RCO values within the -1.0 to 1.0 range.

Whilst applying the spatial size constraints resulted in more ‘acceptable’ RCOxy values for King Sabata Dalindyebo municipality, this was not the case for other municipalities. For instance, the municipality of uMuziwabantu in KwaZulu-Natal received a RCOxy value of 0.54 when the computation was performed without any spatial size constraints, but this value changed to -0.54 when the 50 km^2 spatial size constraint was imposed. When the spatial constraint reached 30 km^2 the RCOxy value for this municipality fell to less than -2.0 and upon application of the 10 km^2 constraint the value reached as low as -2.9. These analyses suggest that the RCOxy does not give a robust measure of residential concentration based upon the pattern of poor and non-poor population in South Africa mapped by the data used in this thesis.

These RCO analyses were repeated to examine the pattern of results concerning the spatial concentration of non-poor compared to poor, in other words computing a RCOyx to complement the RCOxy discussed above. Figure B.6 shows a histogram of RCOyx scores for municipalities with no spatial constraints imposed. Of the 235 municipalities in South Africa, 11 have RCOyx scores lower than -1.0. In The Big 5 False Bay in KwaZulu-Natal and Ga-Segonyana in North West/Northern Cape, the RCO value computed is less than -3.0.

Figure B.6: Histogram of RCO yx with no bottom-coding



Centralisation

ACE is defined as:

$$ACE = \left(\sum_{i=1}^{n} X_{i-1}A_i\right) - \left(\sum_{i=1}^{n} X_i A_{i-1}\right)$$

where the n areal units are ordered by increasing distance from the central business district, X_i and Y_i are the respective cumulative proportions of *X*'s and *Y*'s population in areal unit *i*, and A_i refers to the proportion of land area through unit *i* (Massey and Denton, 1988, p.292-293). As noted in Chapter 4, measures of centralisation are rejected to the difficulty in accurately defining what constituents a CBD.

Appendix C: Spatial patterns of exposure to inequality amongst the 'non-poor' population

A similar range of analyses to those presented in Chapter 5 have also been carried out with regards to the exposure of the non-poor population to the poor population. However, the conceptual justification for an intensity measure is less apparent for the non-poor subgroup of the population, so the following analyses focus solely on the $aLDPyx_i$*measure of exposure.

Figure C.1 shows the national deciles of the $aLDPyx_i$* distribution at datazone level for the whole of South Africa. It is strikingly evident from this map that the places where the non-poor population is most exposed to the poor population are predominantly located in the more rural areas where the levels of poverty are highest. This is of course as one might expect because as the poverty rate rises, the ratio of poor to non-poor people increases and therefore the likelihood of a non-poor person being exposed to a poor person similarly rises.

The cartographic representation of the results shown in Figure C.1 is backed-up by the tabular data presented in Tables C.1 and C.2.

Table C.1 shows the twelve municipalities with the largest absolute numbers of datazones in the highest exposure decile nationally on the $aLDPyx_i$* measure. These twelve municipalities account for over half the datazones in the highest exposure decile, with the remaining datazones in the highest exposure decile being located across 34 different municipalities. Nine of the twelve municipalities listed in Table C.1 are located in the Eastern Cape province, with the other three municipalities being located in KwaZulu-Natal.

Table C.2 shows the twenty municipalities with the highest proportions of their constituent datazones in the highest exposure decile on the *aLDPyx$_i$** measure. It is evident that six of the municipalities are composed entirely of datazones in the highest exposure decile, with a further ten municipalities being characterised by over 90% of their datazones being within the highest exposure decile.

Figure C.1: National Datazone deciles of *aLDPyx$_i$**

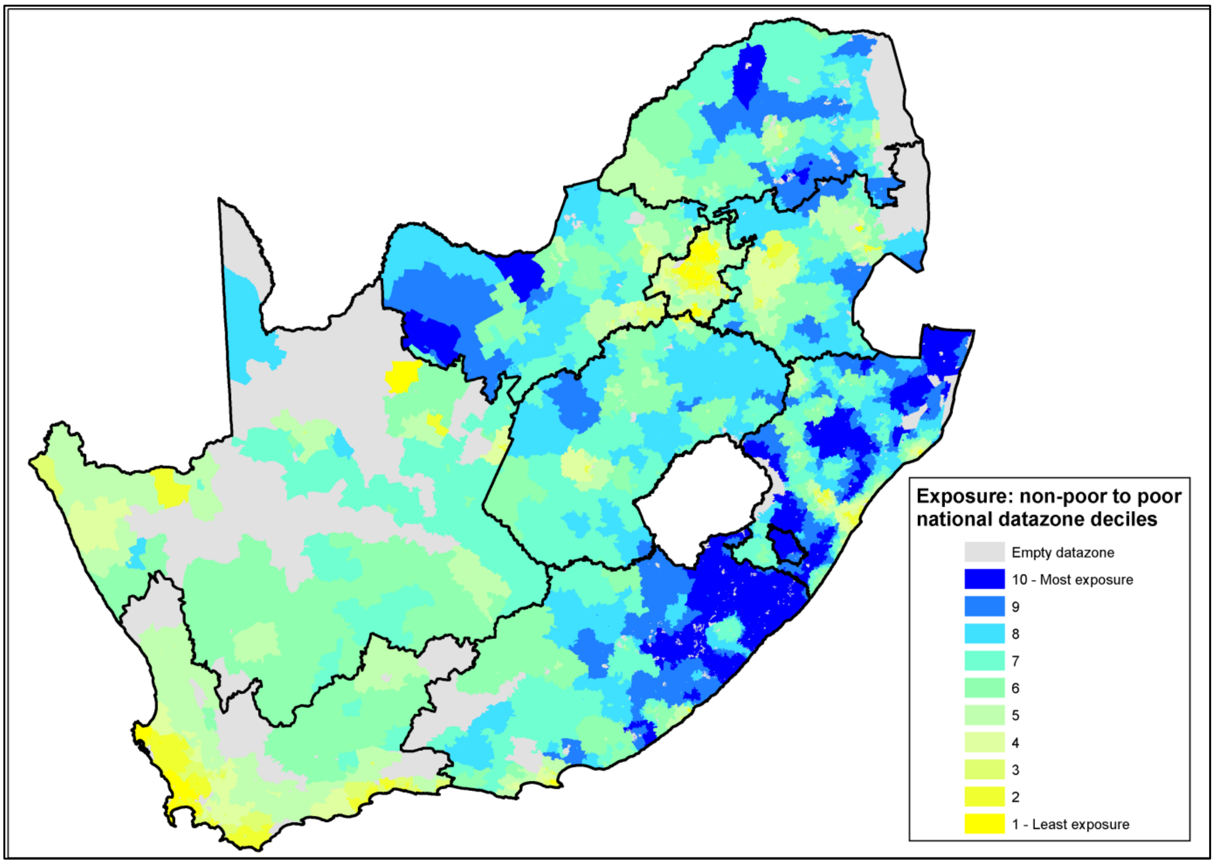


Table C.1: Location of Datazones in the 10% highest exposure decile nationally on the *aLDPyx$_i$ measure**

Municipality	**Number**	**Percentage (column total)**
Umzimvubu, EC	129	5.8%
Qaukeni, EC	124	5.6%
Mbhashe, EC	119	5.4%
Mbizana, EC	113	5.1%
Intsika Yethu, EC	99	4.5%
Nyandeni, EC	92	4.2%
Nongoma, KZN	90	4.1%
Mhlontlo, EC	90	4.1%
Msinga, KZN	81	3.7%
Jozini, KZN	81	3.7%
Engcobo, EC	73	3.3%
Port St Johns, EC	72	3.2%
Other (34 muns)	1053	47.5%
Total in the 10% highest exposure decile nationally	2216	100.0%

Table C.2: The twenty municipalities with the highest proportions of Datazones in the 10% highest *aLDPyx*$_i$* decile nationally

Municipality	Number of Datazones in the municipality	Number of Datazones in the 10% highest exposure decile nationally	Percentage of municipality Datazones in the 10% highest exposure decile nationally
Ntabankulu, EC	68	68	100.0%
Qaukeni, EC	124	124	100.0%
Msinga, KZN	81	81	100.0%
Nkandla, KZN	67	67	100.0%
Setla-Kgobi, NW	54	54	100.0%
Port St Johns, EC	72	72	100.0%
Elundini, EC	72	71	98.6%
Intsika Yethu, EC	101	99	98.0%
Mbhashe, EC	122	119	97.5%
Engcobo, EC	75	73	97.3%
Ingwe, KZN	54	51	94.4%
Nongoma, KZN	96	90	93.8%
Impendle, KZN	15	14	93.3%
Mhlontlo, EC	97	90	92.8%
Moshaweng, NW	41	38	92.7%
Mbizana, EC	123	113	91.9%
Jozini, KZN	91	81	89.0%
Blouberg, LP	83	72	86.7%
Indaka, KZN	57	48	84.2%
Maphumulo, KZN	60	50	83.3%

These results reflect the reality that in places such as Ntabankulu in the Eastern Cape almost every person is counted as being deprived according to the definition of ‘income and material deprivation’ adopted in the SAIMD 2001. Poverty rates are so high in places such as these that any non-poor people living in the area are surrounded almost entirely by poor people, and therefore they would have very high levels of exposure to socio-economic inequality as they go about their daily lives.

Of the twenty municipalities listed in Table C.2, nine are located in the Eastern Cape, a further eight are located in KwaZulu-Natal, two are located in North West province, and one is located in Limpopo.

The measures of exposure as experienced by the *non-poor* population show markedly different patterns than the measures of exposure experienced by the poor. Overwhelmingly, the areas where exposure of the non-poor to the poor is highest are located in the largely rural municipalities where levels of poverty and deprivation are extremely high. In these areas, people who are non-poor will have extremely high levels of exposure to inequality because almost everyone else in their local area will be poor.

Appendix D: Additional base models for the ‘poor’ population subgroup

Model A

Model A: Test 1	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposP	0.0792	0.0926	0.8550	0.3920	

Model A: Test 2	Estimate	Std. Error	z value	Pr(>|z|)	Sig
log_incS	-0.0833	0.0732	-1.1390	0.2550	

Model A: Test 3	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposP	0.0284	0.1116	0.2550	0.7990	
log_incS	-0.0705	0.0887	-0.7940	0.4270	

Model A: Test 4	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposP	-0.0355	0.1140	-0.3110	0.7558	
log_incS	-0.1501	0.0960	-1.5640	0.1179	
exposP:log_incS	0.1387	0.0686	2.0210	0.0432	*

Model A: Test 5	**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**	**Sig**
exposP	0.1119	0.1426	0.7840	0.4328	
log_incS	0.0973	0.1073	0.9070	0.3645	
exposP:log_incS	0.0268	0.0767	0.3500	0.7264	
edu2	0.5182	0.2311	2.2420	0.0249	*
edu3	0.9343	0.2231	4.1870	0.0000	***
edu4	0.6422	0.2311	2.7790	0.0054	**
edu5	0.4652	0.3346	1.3900	0.1645	
hhperS	-0.1207	0.0439	-2.7490	0.0060	**
spoor2	0.2380	0.2961	0.8040	0.4215	
spoor3	0.4414	0.2739	1.6120	0.1070	
spoor4	0.5398	0.2844	1.8980	0.0576	.
spoor5	1.0053	0.3195	3.1470	0.0017	**
groupdis1	0.3626	0.1301	2.7880	0.0053	**
ssocmobc2	-0.0707	0.1224	-0.5780	0.5635	
ssocmobc3	0.3906	0.1390	2.8100	0.0050	**
jobprest1	0.4040	0.1208	3.3450	0.0008	***
futmob2	-0.2509	0.1252	-2.0040	0.0451	*
futmob3	0.1535	0.1510	1.0160	0.3095	
futmob4	0.3200	0.2248	1.4240	0.1546	
classconindS	0.1279	0.0570	2.2450	0.0248	*
meritindS	0.4570	0.0552	8.2830	0.0000	***
exogindS	0.1474	0.0597	2.4680	0.0136	*
q168rS	-0.1709	0.0578	-2.9570	0.0031	**
anc1	-0.4129	0.1358	-3.0410	0.0024	**
prov2	-0.5108	0.4282	-1.1930	0.2328	
prov3	0.0028	0.4999	0.0060	0.9956	
prov4	0.2882	0.4861	0.5930	0.5533	
prov5	0.1353	0.4078	0.3320	0.7400	
prov6	-0.3828	0.4511	-0.8490	0.3961	
prov7	-0.4731	0.3966	-1.1930	0.2329	
prov8	0.5191	0.4504	1.1530	0.2491	
prov9	0.5578	0.4473	1.2470	0.2124	

Model C

Model C: Test 1	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposP	0.0367	0.0948	0.3880	0.6980	

Model C: Test 2	Estimate	Std. Error	z value	Pr(>|z|)	Sig
log_incS	-0.0530	0.0769	-0.6900	0.4900	

Model C: Test 3	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposP	-0.0008	0.1156	-0.0070	0.9950	
log_incS	-0.0534	0.0933	-0.5720	0.5670	

Model C: Test 4	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposP	0.0273	0.1189	0.2300	0.8180	
log_incS	-0.0158	0.1010	-0.1570	0.8760	
exposP:log_incS	-0.0697	0.0730	-0.9550	0.3400	

Model C: Test 5	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposP	0.0994	0.1199	0.8290	0.4072	
log_incS	-0.0055	0.1060	-0.0520	0.9584	
exposP:log_incS	-0.1070	0.0774	-1.3830	0.1667	
empl2	0.6060	0.2213	2.7380	0.0062	**
empl3	-0.0603	0.1470	-0.4100	0.6818	
empl4	0.5072	0.1991	2.5470	0.0109	*
empl5	0.1562	0.1973	0.7910	0.4287	
empl6	0.2157	0.1861	1.1590	0.2464	
empl7	-0.3344	0.3378	-0.9900	0.3222	
empl8	-0.1337	0.2678	-0.4990	0.6176	
spoor2	-0.3618	0.2817	-1.2840	0.1990	
spoor3	-0.6655	0.2605	-2.5550	0.0106	*
spoor4	-0.4344	0.2705	-1.6060	0.1082	
spoor5	0.0703	0.3042	0.2310	0.8172	
groupdis	0.3177	0.1257	2.5270	0.0115	*
jobprest	-0.2811	0.1181	-2.3800	0.0173	*
classconindS	0.1837	0.0555	3.3060	0.0009	***
meritindS	0.2222	0.0518	4.2890	0.0000	***
exogindS	0.1711	0.0545	3.1410	0.0017	**
q168rS	-0.1262	0.0541	-2.3340	0.0196	*
polideol2	-0.4713	0.1235	-3.8170	0.0001	***
polideol3	-0.7568	0.1508	-5.0180	0.0000	***
polideol4	-0.2858	0.1487	-1.9220	0.0546	.
ineqavrS	0.3588	0.0522	6.8770	0.0000	***

Appendix E: Modelling attitudes to inequality amongst the ‘non-poor’ population

In Chapter 7, I discussed the conceptual and methodological underpinnings of the statistical modelling undertaken for this thesis. I described how I produced four separate models, consisting of a model for each of the two dependent variables for each of the poor and non-poor subsets of the survey dataset. In Chapter 7 I focused my discussion of the results on the poor population subset. In this Appendix I present the equivalent set of results for the non-poor subset.

The attitudes of the non-poor are important to consider alongside the attitudes of the poor, because the non-poor are likely to include those with the greatest political as well as economic power. The following section of this chapter presents and discusses the outputs from Models B and D, namely the effect of exposure to inequality on inequality aversion (Model B) and on support for government role in redistribution (Model D).

Association between exposure to inequality and inequality aversion – non-poor subgroup

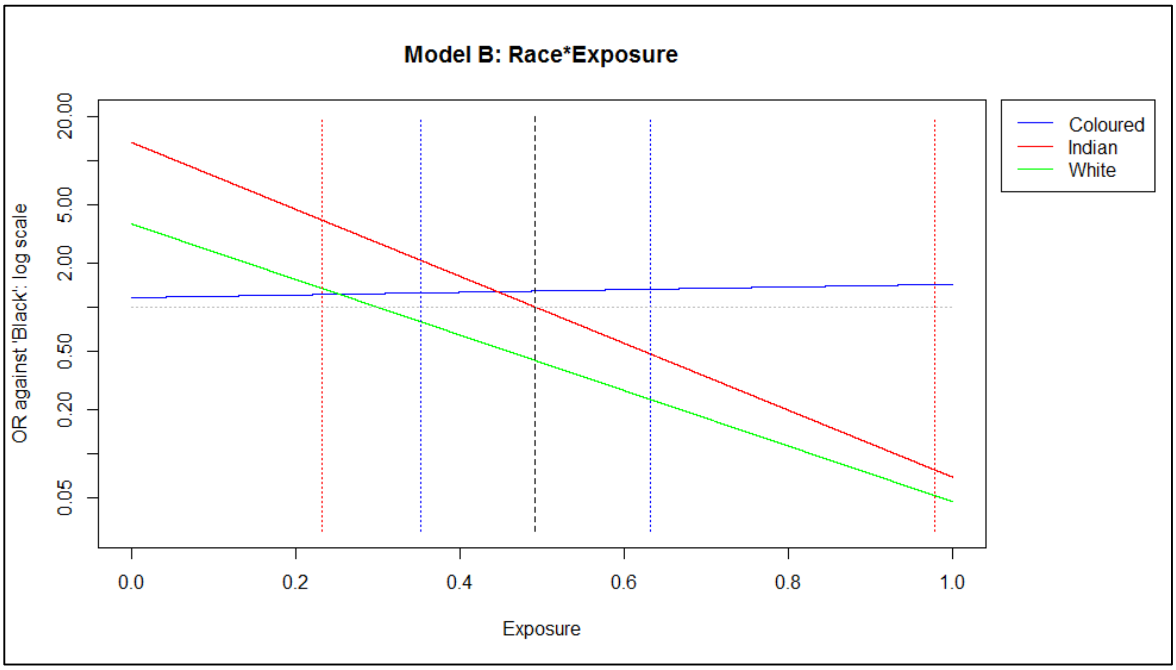
Table E.1: Model B – Inequality Aversion – non-poor subgroup

Parameter	Estimate	Std. Error	z value	Pr(>|z|)	Sig.	OR	OR 95% CI lower	OR 95% CI upper	Group LR test pval
(A) Respondent's personal and household characteristics									
Educational attainment (Ref='No schooling'; N=8)									*0.0099*
Primary (N=33)	-0.56	1.86	-0.30	0.7600		0.57	0.02	21.89	
Grades 8-11 or equivalent (N=246)	-0.84	1.80	-0.47	0.6400		0.43	0.01	14.57	
Matric or equivalent (N=389)	-0.50	1.79	-0.28	0.7800		0.61	0.02	20.14	
Tertiary (N=360)	0.02	1.79	0.01	0.9900		1.02	0.03	34.33	
Employment status (Ref='Employed full time'; N=483)									*0.0000*
Employed part time (N=52)	-0.93	0.34	-2.73	0.0063	**	0.40	0.20	0.77	
Unemployed, seeking work (N=111)	-0.20	0.28	-0.70	0.4800		0.82	0.47	1.43	
Unemployed not looking for work (N=106)	1.16	0.30	3.83	0.0001	***	3.20	1.76	5.80	
Pensioner (N=113)	0.29	0.30	0.99	0.3200		1.34	0.75	2.41	
Student/learner (N=127)	-0.39	0.24	-1.62	0.1100		0.68	0.42	1.09	
Permanently sick/disabled (N=12)	0.58	0.98	0.59	0.5500		1.78	0.26	12.08	
Other employment status (N=32)	-0.38	0.38	-1.01	0.3100		0.69	0.33	1.43	
Population Group (Ref='Black'; N=263)									*0.0002*
Coloured (N=194)	0.26	0.30	0.87	0.3800		1.29	0.72	2.31	
Indian (N=250)	0.01	0.35	0.02	0.9800		1.01	0.51	1.99	
White (N=329)	-0.83	0.23	-3.57	0.0004	***	0.43	0.28	0.69	
(B) Respondent's attidudes and views									
Self-rated poverty status (ref='Wealthy/very comfortable'; N=231)									*0.0000*
Reasonably comfortable (N=505)	0.40	0.20	2.03	0.0420	*	1.49	1.01	2.18	
Just getting by (N=276)	0.95	0.25	3.85	0.0001	***	2.57	1.59	4.16	
Poor (N=22)	1.88	0.50	3.77	0.0002	***	6.57	2.47	17.48	
Very poor (N=2)	-3.69	3.65	-1.01	0.3100		0.03	0.00	32.06	
Subjective social class position	0.38	0.10	3.80	0.0002	***	1.46	1.20	1.78	
Perceived social mobility history (Ref='Upward mobility'; N=430)									*0.0170*
No mobility (N=366)	0.38	0.18	2.15	0.0310	*	1.46	1.03	2.07	
Downward mobility (N=240)	0.56	0.21	2.62	0.0088	**	1.75	1.15	2.67	
Higher job prestige than father (N=391)	0.34	0.17	2.00	0.0450	*	1.40	1.01	1.95	
Perceived class conflict index	0.47	0.08	5.72	0.0000	***	1.59	1.36	1.87	
Perceived group discrimination (N=269)	0.52	0.20	2.58	0.0098	**	1.68	1.13	2.48	
(C) Geographical location variables									
Province of residence (Ref='Western Cape'; N=180)									*0.0001*
Eastern Cape (N=91)	-0.72	0.49	-1.48	0.1400		0.49	0.19	1.26	
Northern Cape (N=75)	-0.43	0.57	-0.77	0.4400		0.65	0.21	1.97	
Free State (N=96)	-0.94	0.52	-1.81	0.0700	.	0.39	0.14	1.08	
KwaZulu Natal (N=237)	0.54	0.41	1.33	0.1800		1.72	0.77	3.82	
North West (N=10)	-0.01	0.82	-0.01	0.9900		0.99	0.20	4.91	
Gauteng (N=269)	-0.78	0.33	-2.39	0.0170	*	0.46	0.24	0.87	
Mpumalanga (N=46)	1.70	0.60	2.82	0.0048	**	5.48	1.68	17.88	
Limpopo (N=32)	-0.95	0.70	-1.36	0.1700		0.39	0.10	1.52	
(D) Neighbourhood level poverty and exposure variables and interactions									
Neighbourhood poverty rate	-0.32	0.14	-2.21	0.0270	*	0.73	0.55	0.97	
Exposure to inequality	0.27	0.23	1.20	0.2300		1.31	0.84	2.04	
Neighbourhood poverty rate * Exposure to inequality	-0.08	0.07	-1.14	0.2600		0.93	0.81	1.06	
*Neighbourhood poverty rate * Perceived social mobility history (Ref='Upward mobility')*									*0.0025*
Neighbourhood poverty rate * No mobility	0.04	0.15	0.26	0.8000		1.04	0.77	1.40	
Neighbourhood poverty rate * Downward mobility	0.61	0.19	3.22	0.0013	**	1.84	1.27	2.67	
*Exposure * Population Group (Ref='Black')*									*0.0052*
Exposure * Coloured	0.03	0.24	0.12	0.9000		1.03	0.64	1.66	
Exposure * Indian	-0.74	0.37	-2.01	0.0440	*	0.48	0.23	0.98	
Exposure * White	-0.61	0.20	-3.01	0.0026	**	0.54	0.37	0.81	
*Exposure * Self-rated poverty status (ref='Wealthy/very comfortable')*									*0.0001*
Exposure * Reasonably comfortable	0.15	0.17	0.85	0.4000		1.16	0.83	1.62	
Exposure * Just getting by	0.72	0.21	3.41	0.0006	***	2.05	1.36	3.10	
Exposure * Poor	-1.56	0.61	-2.56	0.0100	*	0.21	0.06	0.69	
Exposure * Very poor	-3.00	3.05	-0.98	0.3300		0.05	0.00	19.69	

* significant at $p < 0.05$; ** significant at $p < 0.01$; *** significant at $p < 0.001$

Exposure variables and interactions

Figure E.1: Interaction of 'race' with 'exposure'



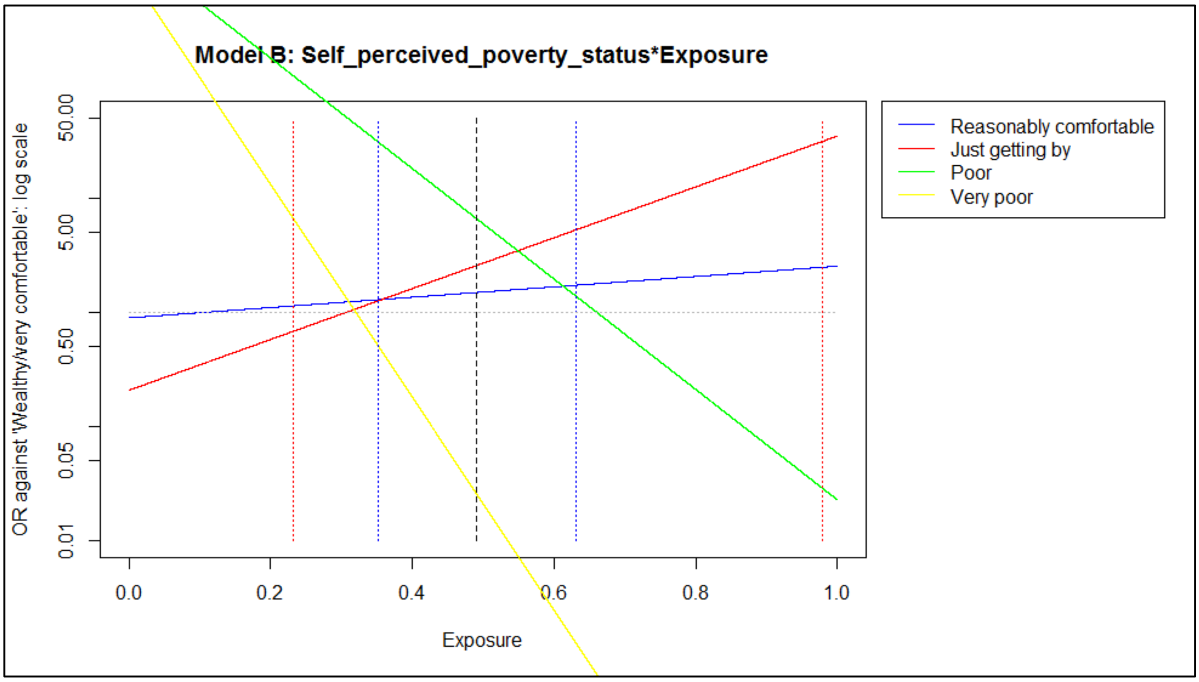
Note regarding vertical reference lines: black dash = mean exposure; blue dot = mean +/- 1 standard deviation exposure; red dot = minimum/maximum observed exposure
Note regarding horizontal reference line: grey dot = OR of 1.0.

At the mean '*exposure_of_rich*' level, the odds of a White respondent giving a response (i.e. inequality aversion) above any particular threshold are just 0.43 times (CI: 0.28-0.69) those of a Black African respondent. In contrast, at the mean *exposure_of_rich* level there are no statistically significant differences in the odds of a Coloured or Indian respondent compared to a Black African respondent.

The relationship between *exposure_of_rich* and the response variable can be seen to vary by race group. So, although there was not a significant difference between Black African and either Coloured or Indian respondents at mean *exposure_of_rich*, there is a significant difference with regards to the way in which exposure affects responses for these two groups when compared to the Black African group. When the level of exposure is 1 standard deviation below the mean exposure, the odds of giving a response (i.e. inequality aversion) above any particular threshold are 2.10 times greater for Indian respondents than Black African respondents, all other things being equal. In contrast, when the level of exposure is 1 standard deviation above the mean, the odds for an Indian respondent are just 0.48 times those of a Black African respondent. In addition to exhibiting the statistically significant difference between White and Black African respondents at mean *exposure_of_rich* noted above, the White population also displays a significantly different relationship between exposure and response, when compared to the Black African population. So, when the level of exposure is 1 standard deviation below the mean exposure, the odds of a White respondent giving a response (i.e. inequality aversion) above any particular threshold are just 0.80 times those of a Black African respondent, all other things being equal. When the level of exposure is 1 standard deviation above the mean exposure, the odds of a White respondent giving a response (i.e. inequality aversion) above any particular threshold are just 0.24 times those of a Black African respondent, all other things being equal.

Coloured respondents show broadly the same pattern of response as Black respondents irrespective of the level of exposure, all other things being equal. For both Indian and White respondents, increasing levels of *exposure_of_rich* is associated with lower aversion to income inequality relative to Black African respondents. So non-poor White and Indian people who live in areas where inequality is really obvious have much lower odds of being averse to income inequality than non-poor Black African people living in the same areas, all other things being equal. Non-poor White people living in areas where inequality is least obvious have notably higher odds of being averse to income inequality than non-poor Black people living in the same area, all other things being equal.

Figure E.2: Interaction of 'self-perceived poverty status' with 'exposure'



Note regarding vertical reference lines: black dash = mean exposure; blue dot = mean +/- 1 standard deviation exposure; red dot = minimum/maximum observed exposure
Note regarding horizontal reference line: grey dot = OR of 1.0.

At the mean '*exposure_of_rich*' level, the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 1.49 times (CI: 1.01-2.18) greater for a respondent who regards him/herself as being 'reasonably comfortable' than for a person who is identical in every way except for regarding him/herself as 'wealthy/very comfortable'. This difference does not vary significantly in terms of the level of *exposure_of_rich*. In other words, these two groups do exhibit differences in terms of inequality aversion, but these differences seem to be independent of the level of exposure.

There is a more pronounced difference in response between non-poor people who regard themselves as 'just getting by' and non-poor people who regard themselves as 'wealthy/very comfortable'. At mean *exposure_of_rich*, the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 2.57 times (CI: 1.59-4.16) greater for someone who is 'just getting by' than for someone who is identical in every way except for being ‘wealthy/very comfortable’. Furthermore, this relationship can be seen to vary significantly according to the level of *exposure_of_rich* experienced: at mean+1sd *exposure_of_rich*, the odds of giving a response above any particular threshold by someone who is ‘just getting by’ are 5.27 times greater than someone who is ‘wealthy/very comfortable’, all other things being equal, whereas at mean-1sd *exposure_of_rich* the odds are just 1.26 (and at the minimum actual observed *exposure_of_rich* level the odds are in fact just 0.68 those of the ‘wealthy/very comfortable’ respondents, all other things being equal.

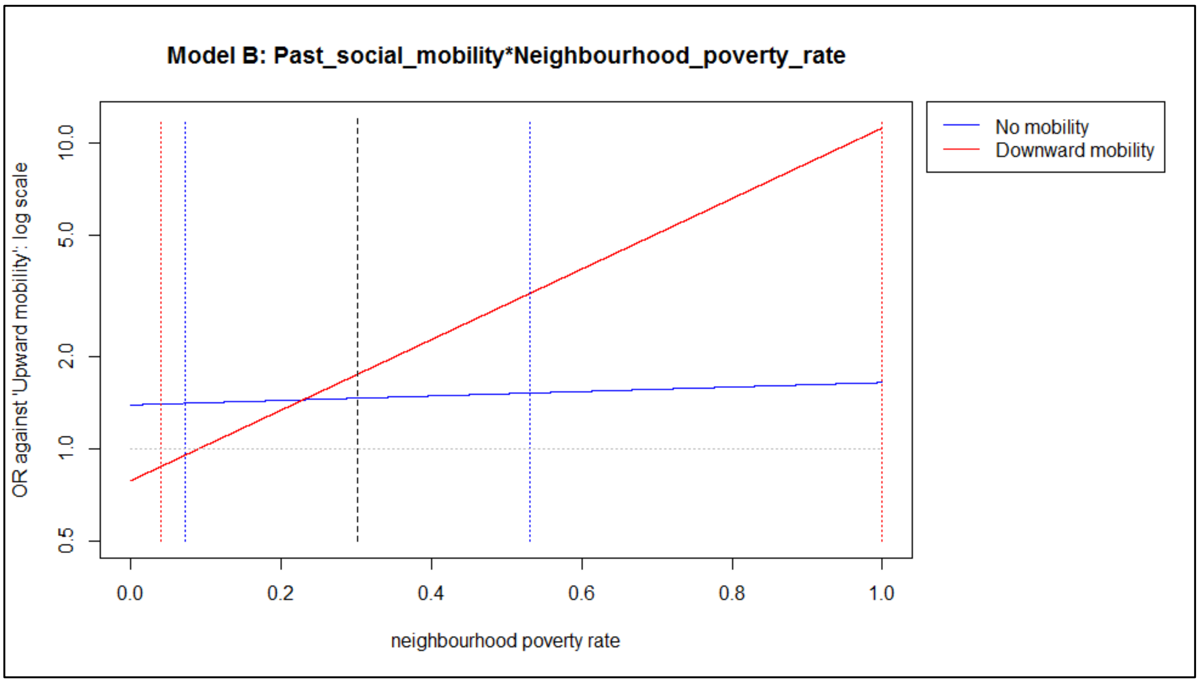
The contrast with the ‘wealthy/very comfortable’ respondents is starker still for respondents who regard themselves as being poor (remembering again that everyone in this sample subset is defined empirically as ‘non-poor’). At mean the *exposure_of_rich* level, the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 6.57 times (CI: 2.47-17.48) greater for a respondent who regards him/herself as being ‘poor’ than for a person who is identical in every way except for regarding him/herself as ‘wealthy/very comfortable’. This relationship can also be seen to vary according to the level of *exposure_of_rich* experienced and, in this case, the relationship consists of decreasing odds ratios with increasing *exposure_of_rich*[56]. So, in areas where non-poor people are relatively less exposed to socio-economic inequality, those that regard themselves as ‘poor’ are substantially more likely to be averse to income inequality than people who regard themselves as ‘wealthy/very comfortable’, but in areas where the non-poor are highly exposed to inequality, people who regard themselves as ‘poor’ are much less averse to income inequality than people

[56] So, at mean-1sd *exposure_of_rich*, the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 31.10 times (CI: 1.01-2.18) greater for a respondent who regards him/herself as being ‘poor’ than for a person who is identical in every way except for regarding him/herself as ‘wealthy/very comfortable’. At mean+1sd *exposure_of_rich*, however, the odds are just 1.38 times greater, all other things being equal.

to regard themselves as ‘wealthy/very comfortable’. There are only 22 non-poor people who regard themselves as ‘poor’, but the coefficient (and OR) is still significant.

Although the interaction coefficient for the non-poor respondents who regard themselves as being ‘very poor’ is steeper still, there are only 2 people in this group and therefore the results are not significant.

Figure E.3: Interaction of ‘perceived social mobility history’ with ‘neighbourhood poverty rate’



Note regarding vertical reference lines: black dash = mean poverty rate; blue dot = mean +/- 1 standard deviation poverty rate; red dot = minimum/maximum observed poverty rate
Note regarding horizontal reference line: grey dot = OR of 1.0.

At the mean ‘neighbourhood poverty rate’ level for the non-poor subset of respondents (note that this is different to the mean neighbourhood poverty rate for the poor subset), the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 1.46 times (CI: 1.03-2.07) greater for a respondent who regards him/herself as having experienced ‘no mobility’ over recent years than for a person who is identical in every way except for regarding him/herself as having experienced ‘upward mobility’. This difference does not vary significantly in terms of the neighbourhood poverty rate. In other words, these two groups do exhibit differences in terms of inequality aversion, but these differences seem to be independent of the level of poverty in the home neighbourhood.

At the mean ‘neighbourhood poverty rate’ level for the non-poor subset, respondents who perceive themselves to have been ‘downwardly mobile’ over recent years also differ significantly in their response regarding inequality aversion when compared to the ‘upwardly mobile’ respondents. In this case, at mean neighbourhood poverty rate, the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 1.75 times (CI: 1.15-2.67) greater for a respondent who regards him/herself as having experienced ‘downward mobility’ over recent years than for a person who is identical in every way except for regarding him/herself as having experienced ‘upward mobility’. Furthermore, this relationship varies according to the neighbourhood poverty rate. In areas where the neighbourhood poverty rate is mean-1sd (again based upon the distribution of neighbourhood poverty rates based upon data for non-poor subset only), the odds of a ‘downwardly mobile’ person responding above any particular threshold are 0.95 those of an ‘upwardly mobile’ person, whereas at mean+1sd neighbourhood poverty rate, the odds are 3.40 times greater than those of an ‘upwardly mobile’ person, all other things being equal. Focusing at the extreme of the observed poverty rate spectrum, if we were to take two non-poor people who were identical other than their perceived social mobility history over recent years, and assume they both live in a neighbourhood where almost everyone else is poor (i.e. extremely high neighbourhood poverty rate), then the person who thought they’d been ‘downwardly mobile’ would have odds of answering above a given response threshold that are 11.20 times greater than the odds for the person who thought they'd been 'upwardly mobile'.

Other statistically significant effects

In addition to the findings concerning the effects of exposure to inequality on people's aversion to income inequality, a number of other explanatory variables also exhibit statistically significant relationships with the dependent variable. With regard to employment status, the odds of a non-poor person who is 'employed part time' giving a response (i.e. inequality aversion) above any particular response threshold are just 0.40 those of a non-poor 'employed full time' respondent. In contrast, the odds of giving a response (i.e. inequality aversion) above any particular response threshold are 3.20 times greater for a non-poor respondent who is 'unemployed and not looking for work' than a non-poor respondent who is 'employed full time'.

The subjective social class position variable 'topbott100' is a continuous variable, ranging from 0 to 100. As with all continuous variables in model it is mean centred. A respondent who places him/herself at mean+1sd in the social class position scale has odds of answering above a given response threshold that are 1.46 times greater than an otherwise identical person who places her/herself at the mean social class position. In addition, respondents who perceive themselves as having a higher-prestige job than their father have odds of responding above a given threshold that are 1.40 times greater than respondents who perceive themselves to have job of equal or lower prestige than their father.

The class conflict index 'classconind' is a continuous variable, ranging from 1 to 5 in increments of 0.5 (i.e. it is the simple average of two separate questions on five-point scales). It has been mean-centred for the model. A respondent with a perceived class conflict score of mean+1sd has odds of answering above a given response threshold that are 1.59 times greater than an otherwise identical person with a mean perceived class conflict score. Furthermore, respondents who perceive themselves as being part of a group that is discriminated against have odds of responding above a given threshold that are 1.68 times greater than respondents who do not perceive themselves to be part of a discriminated group.

Finally, Model D examines the effects of exposure to inequality on the attitudes of the non-poor population towards it being the government's responsibility to redistribute income from those with high incomes to those with low incomes.

Association between exposure to inequality and support for Government Redistribution – non-poor subgroup

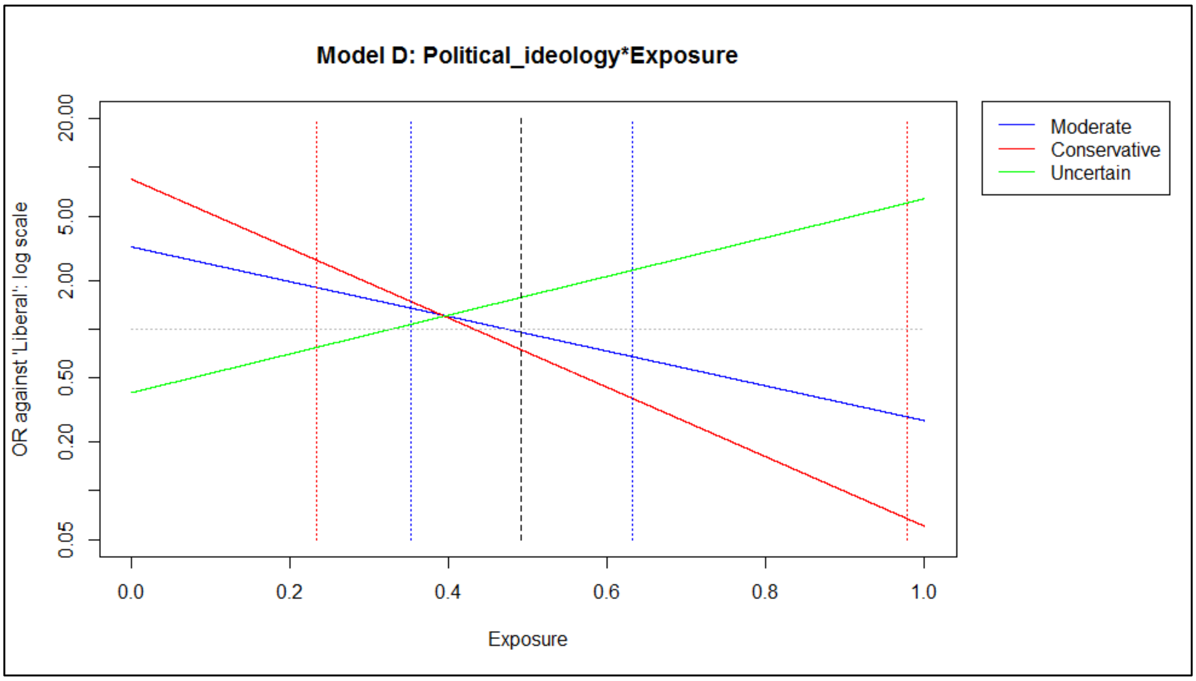
Table E.2: Model D – Government Redistribution – non-poor subgroup

Parameter	Estimate	Std. Error	z value	Pr(>|z|)	Sig.	OR	OR 95% CI lower	OR 95% CI upper	Group LR test pval
(A) Respondent's personal and household characteristics									
Marital status (Ref='Married'; N=600)									*0.0000*
Widowed (N=60)	1.10	0.43	2.58	0.0098	**	3.00	1.30	6.89	
Separated / divorced (N=49)	-1.08	0.34	-3.20	0.0014	**	0.34	0.17	0.66	
Never married (N=327)	0.40	0.16	2.50	0.0130	*	1.50	1.09	2.05	
Household Asset Index	-0.44	0.08	-5.31	0.0000	***	0.64	0.55	0.76	
ANC voter (N=326)	0.46	0.17	2.73	0.0063	**	1.58	1.14	2.20	
(B) Respondent's attidudes and views									
Inequality aversion	0.45	0.07	6.21	0.0000	***	1.56	1.36	1.80	
Left-right ideology (Ref='Liberal'; N=324)									*0.0260*
Moderate (N=346)	-0.05	0.18	-0.26	0.7900		0.95	0.67	1.35	
Conservative (N=201)	-0.30	0.22	-1.34	0.1800		0.74	0.48	1.15	
Uncertain (N=165)	0.45	0.21	2.11	0.0350	*	1.57	1.03	2.39	
Perceived social mobility history (Ref='Upward mobility'; N=430)									*0.0019*
No mobility (N=366)	0.28	0.17	1.68	0.0930	.	1.32	0.95	1.83	
Downward mobility (N=240)	0.66	0.19	3.53	0.0004	***	1.93	1.34	2.78	
Expected future social mobility over next 5 years (Ref='Improve'; N=491)									*0.0100*
Stay the same (N=255)	0.43	0.18	2.45	0.0140	*	1.54	1.09	2.17	
Worsen (N=211)	0.64	0.22	2.95	0.0032	**	1.90	1.24	2.91	
Uncertain (N=79)	0.30	0.28	1.07	0.2900		1.35	0.78	2.34	
(C) Geographical location variables									
N/A									
(D) Neighbourhood level poverty and exposure variables and interactions									
Neighbourhood poverty rate	0.25	0.13	1.95	0.0510	.	1.29	1.00	1.66	
Exposure to inequality	-0.12	0.20	-0.62	0.5300		0.88	0.60	1.30	
Neighbourhood poverty rate * Exposure to inequality	0.14	0.07	2.02	0.0440	*	1.15	1.00	1.33	
*Exposure * Left-right ideology (Ref='Liberal')*									*0.0000*
Exposure * Moderate	-0.35	0.17	-2.05	0.0400	*	0.71	0.51	0.98	
Exposure * Conservative	-0.69	0.21	-3.24	0.0012	**	0.50	0.33	0.76	
Exposure * Uncertain	0.39	0.18	2.09	0.0360	*	1.47	1.02	2.11	

* significant at $p < 0.05$; ** significant at $p < 0.01$; *** significant at $p < 0.001$

Exposure variables and interactions

Figure E.4: Interaction of 'Political Ideology' with 'exposure'



.

Note regarding vertical reference lines: black dash = mean exposure; blue dot = mean +/- 1 standard deviation exposure; red dot = minimum/maximum observed exposure
Note regarding horizontal reference line: grey dot = OR of 1.0

With regards to political ideology, at the mean '*exposure_of_rich*' level, there are no statistically significant differences between people who regard themselves as 'liberal' (i.e. the reference group) and people who regard themselves as 'moderate' or 'conservative' in terms of support for government role in income redistribution. However, people who are 'uncertain' of their political ideology show a significantly higher support for government role in income redistribution than people who regard themselves as 'liberal' (OR=1.57; CI=1.03-2.39).

Although the 'main effects' (i.e. at mean '*exposure_of_rich*') for 'moderate' and 'conservative' political ideologies are not significantly different to the 'liberal' ideology, there are significant differences in the way that *exposure_of_rich* is associated with support for government income redistribution for each of ‘moderate’, ‘conservative’ and ‘uncertain’ when compared against ‘liberal’. As *exposure_of_rich* increases, people who regard themselves as being ‘moderate’ or ‘conservative’ exhibit progressively *less support* for government role in income redistribution when compared to ‘liberal, whereas people who are ‘uncertain’ of their ideology exhibit progressively *more support* when compared to ‘liberal’.

Other statistically significant effects

In addition to the findings concerning the effects of exposure to inequality on people’s support for government’s role in income redistribution, a number of other explanatory variables also exhibit statistically significant relationships with the dependent variable For instance, there is a significant association between people’s marital status and their support for government role in income distribution. When compared against the ‘married’ reference category, those who are ‘widowed’ or ‘never married’ are more supportive of government role in income redistribution (OR=3.0 and OR=1.58 respectively), while people who are ‘divorced/separated’ are less supportive (OR=0.34). There is also a significant negative association between household asset ownership and support for government role in redistribution. People with higher levels of asset ownership are less supportive of government role in income redistribution than people with lower levels of asset ownership. With regards to political affiliation, ANC voters are more supportive of government role in redistribution than non-ANC voters (OR=1.58).

In terms of social mobility dynamics, those who believe they have suffered downward social mobility in recent years are more supportive of government role in income redistribution than people who believe they've enjoyed upward mobility (OR=1.93). Furthermore, people who expect their social position to worsen or stay the same are significantly more supportive of government role in income redistribution that people who expect to enjoy future upward social mobility (OR=1.54 and OR=1.90 respectively).

There is a significant positive interaction between the two continuous area level variables of neighbourhood poverty rate and exposure to inequality. A person living in an area with poverty rate and exposure scores 1 standard deviation above the mean on both these explanatory variables has odds 1.15 times greater than a person living in an area with mean scores on both variables of answering above any given '*govredr*' response threshold. Finally, there is also a significant positive association between a person's aversion to inequality ('*ineqavr*') and their support for government role in income redistribution (OR=1.56).

Summary

In Model B, as exposure to inequality increases amongst the non-poor population, we see greater aversion to inequality amongst the Black African population group than amongst either the coloured or white population groups; we see greater aversion to inequality amongst people who regard themselves as 'just getting by' relative to people who regard themselves as 'wealthy/very comfortable'; and we see greater aversion to inequality amongst people who perceive themselves to have been downwardly mobile in relation to people who perceive themselves to have been upwardly mobile. In this sense, this gives a broad indication that the less advantaged categories of the non-poor group tend to have greater aversion to inequality as exposure to inequality rises, all other things being equal. Finally, in Model D, an increase in exposure amongst the non-poor population was associated with a greater support for government role in redistribution amongst people with a liberal political ideology than amongst people who regard themselves as moderate or conservative, all other things being equal.

Appendix F: Additional base models for the 'non-poor' population subgroup

Model B

Model B: Test 1	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposR	-0.0049	0.1067	-0.0460	0.9630	

Model B: Test 2	Estimate	Std. Error	z value	Pr(>|z|)	Sig
log_incS	-0.0368	0.0799	-0.4610	0.6450	

Model B: Test 3	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposR	0.0460	0.1393	0.3300	0.7410	
log_incS	-0.0591	0.1046	-0.5650	0.5720	

Model B: Test 4	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposR	0.0603	0.1487	0.4060	0.6850	
log_incS	-0.0543	0.1062	-0.5120	0.6090	
exposR:log_incS	-0.0162	0.0602	-0.2700	0.7870	

Model B: Test 5	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposR	0.1532	0.1643	0.9320	0.3514	
log_incS	-0.1319	0.1137	-1.1600	0.2459	
exposR:log_incS	-0.0212	0.0614	-0.3460	0.7297	
race2	0.0560	0.2819	0.1990	0.8425	
race3	-0.1812	0.3290	-0.5510	0.5818	
race4	-0.7458	0.2167	-3.4420	0.0006	***
edu2	-0.3200	1.8193	-0.1760	0.8604	
edu3	-0.5130	1.7596	-0.2920	0.7706	
edu4	-0.2916	1.7512	-0.1670	0.8677	
edu5	0.2760	1.7580	0.1570	0.8753	
empl2	-0.7022	0.3297	-2.1300	0.0332	*
empl3	-0.1149	0.2729	-0.4210	0.6737	
empl4	1.0900	0.2982	3.6550	0.0003	***
empl5	0.2800	0.2909	0.9620	0.3359	
empl6	-0.2944	0.2340	-1.2580	0.2083	
empl7	0.1251	0.9472	0.1320	0.8949	
empl8	-0.4019	0.3558	-1.1300	0.2586	
spoor2	0.2849	0.1881	1.5140	0.1299	
spoor3	0.7300	0.2341	3.1180	0.0018	**
spoor4	1.7994	0.4836	3.7200	0.0002	***
spoor5	-0.5058	0.9723	-0.5200	0.6029	
groupdis1	0.4637	0.1924	2.4100	0.0160	*
ssocmobc2	0.4157	0.1733	2.3990	0.0164	*
ssocmobc3	0.5227	0.2066	2.5300	0.0114	*
jobprest1	0.3730	0.1636	2.2800	0.0226	*
classconindS	0.4428	0.0787	5.6230	0.0000	***
topbott100S	0.3285	0.0964	3.4070	0.0007	***
prov2	-0.5147	0.4602	-1.1180	0.2634	
prov3	-0.4820	0.5335	-0.9030	0.3663	
prov4	-0.8700	0.4906	-1.7730	0.0762	.
prov5	0.6400	0.3898	1.6420	0.1006	
prov6	-0.1579	0.7770	-0.2030	0.8390	
prov7	-0.6361	0.3031	-2.0990	0.0358	*
prov8	1.4463	0.5674	2.5490	0.0108	*
prov9	-0.7158	0.6664	-1.0740	0.2828	

Model D

Model D: Test 1	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposR	0.2141	0.1296	1.6520	0.0986	.

Model D: Test 2	Estimate	Std. Error	z value	Pr(>|z|)	Sig
log_incS	0.3352	0.0933	3.5930	0.0003	***

Model D: Test 3	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposR	-0.1454	0.1599	-0.9090	0.3632	
log_incS	0.4059	0.1222	3.3210	0.0009	***

Model D: Test 4	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposR	-0.1982	0.1704	-1.1630	0.2448	
log_incS	0.3863	0.1240	3.1150	0.0018	**
exposR:log_incS	0.0583	0.0684	0.8520	0.3941	

Model D: Test 5	Estimate	Std. Error	z value	Pr(>|z|)	Sig
exposR	-0.2137	0.1708	-1.2510	0.2110	
log_incS	0.2206	0.1279	1.7250	0.0846	.
exposR:log_incS	0.0996	0.0688	1.4470	0.1480	
marstat2	1.0361	0.4181	2.4780	0.0132	*
marstat3	-1.0458	0.3363	-3.1100	0.0019	**
marstat4	0.2974	0.1577	1.8860	0.0593	.
ssocmobc2	0.3074	0.1644	1.8700	0.0614	.
ssocmobc3	0.7332	0.1851	3.9610	0.0001	***
futmob2	0.4500	0.1753	2.5670	0.0103	*
futmob3	0.6865	0.2161	3.1770	0.0015	**
futmob4	0.4826	0.2773	1.7400	0.0818	.
anc1	0.4725	0.1669	2.8300	0.0047	**
polideol2	0.0109	0.1736	0.0630	0.9501	
polideol3	-0.2250	0.2188	-1.0290	0.3037	
polideol4	0.4685	0.2144	2.1850	0.0289	*
assetindexS	-0.4275	0.0824	-5.1890	0.0000	***
ineqavrS	0.4476	0.0712	6.2860	0.0000	***

Appendix G: Geographical level variance terms for final models

The multilevel model output generated in Models A, B, C and D, include information on the geographical level variance terms. This appendix presents this output for the four final model specifications.

Model A

Groups	Name	Variance	Std.Dev.
dz_code:mun_name	(Intercept)	0.4714	0.6866
mun_name	(Intercept)	0.1442	0.3798

Number of groups: dz_code:mun_name 416, mun_name 158

Model B

Groups	Name	Variance	Std.Dev.
dz_code:mun_name	(Intercept)	0.7301	0.8544
mun_name	(Intercept)	0.0000	0.0000

Number of groups: dz_code:mun_name 264, mun_name 78

Model C

Groups	Name	Variance	Std.Dev.
dz_code:mun_name	(Intercept)	0.7720	0.8786
mun_name	(Intercept)	0.2661	0.5159

Number of groups: dz_code:mun_name 416, mun_name 158

Model D

Groups	Name	Variance	Std.Dev.
dz_code:mun_name	(Intercept)	1.1514	1.0731
mun_name	(Intercept)	0.4617	0.6795

Number of groups: dz_code:mun_name 265, mun_name 78

Appendix H: Background to SASAS

As discussed in Appendix A, parts of the research presented in this thesis also represent outputs from an ESRC Pathfinder grant, on which I was the Principal Investigator. Ben Roberts from the Human Sciences Research Council (HSRC) in South Africa was a Co-Investigator on the grant and played an integral role in the analysis of social attitudes around inequality using the South African Social Attitudes Survey (SASAS). Roberts is programme manager for SASAS and, as such, has an unrivalled knowledge of the survey dataset. In our jointly authored end-of-project report from the Pathfinder grant, Roberts provided valuable background information on SASAS. I have reproduced Roberts's description of SASAS here in order to provide additional context for the reader:

> "[The] South African Social Attitudes Survey (SASAS) [is] a nationally representative, repeated cross-sectional survey that has been conducted by the Human Sciences Research Council (HSRC) on an annual basis since 2003. Designed as a time series, SASAS aims to provide a long-term account of the speed and direction of change in underlying public values and the social fabric of modern South Africa. The sample for each year's round of interviewing consists of a cross-section of individuals aged 16 years and older residing in private households, hostels and other structures, regardless of nationality or citizenship. Each probability sample is based on a sub-sample of 500 Population Census enumeration areas (EAs), stratified by province, geographical sub-type and majority population group. A split sample design is used, with two different questionnaire versions being administered to 3,500 target respondents each year in order to accommodate increased thematic content. Apart from a standard set of demographic and background variables, the questionnaires contain a core module, which is repeated each round, with the aim of monitoring change and continuity in a variety of socio-economic and socio-political variables over time. In addition to the core module, each round of interviewing accommodates rotating modules on specific themes, the aim being to provide detailed attitudinal evidence to inform policy and academic debate. The series relies on face-to-face interviewing using a paper-and-pencil survey design, and the English source questionnaires are translated into and administered in other official languages. The study design and research tools were approved by the HSRC's Research Ethics Committee (REC). Participants are asked for written informed consent, while written permission for young South Africans less than 18 years is also secured from their parents/guardians. Through SASAS, the country has been a formal member of International Social Survey Program (ISSP) since 2003." (McLennan et al., 2014, p.51-52)

Appendix I: Glossary of selected technical terms

Atkinson Index

A classical measure of income inequality. There are different variants of the index which place the emphasis on inequality across different parts of the income distribution: Atkinson class A(e) for e = 0.5, 1, 2. The more positive e > 0 (the 'inequality aversion parameter') is, the more sensitive A(e) is to income differences at the bottom of the distribution[57].

Clustering

One of the five dimensions of residential segregation identified by Massey and Denton (1988). Clustering refers to the degree to which neighbourhoods populated by a particular group of the population are geographically contiguous or proximately located. Clustering therefore concerns the spatial distribution of neighbourhoods with respect to one another i.e. taking into account the spatial inter-relationships between neighbourhoods.

[57] Summarised from the Stata help file for the ineqdeco.ado command.

Centralisation

One of the five dimensions of residential segregation identified by Massey and Denton (1988). Centralisation refers to the extent to which members of a particular population sub-group are residentially located close to the centre of an urban area

Concentration

One of the five dimensions of residential segregation identified by Massey and Denton (1988). Concentration indices measure the extent to which a certain population sub-group occupies a small or large proportion of the actual physical space within a geographical area. A group can be said to be residentially concentrated if it occupies a small proportion of the overall physical space

Datazone

A small area level statistical geography for South Africa developed by researchers at the Centre for the Analysis of South African Social Policy (CASASP) at the University of Oxford. The datazones were built using 2001 Census Enumeration Areas. Datazones have a population of approximately 2000.

Deprivation

Townsend defined people as deprived if "they lack the types of diet, clothing, housing, household facilities and fuel and environmental, educational, working and social conditions, activities and facilities which are customary" (Townsend, 1979, p.131).

Evenness

One of the five dimensions of residential segregation identified by Massey and Denton (1988). Evenness refers to the differential proportional distribution of two social groups among neighbourhoods within the specified broader geographical extent, such as within a city or other local government administrative area. A group is said to be segregated if it is unevenly distributed over the constituent neighbourhoods relative to another group

Exposure

One of the five dimensions of residential segregation identified by Massey and Denton (1988). Exposure, refers to the likelihood of potential contact and possible interaction between members of two different groups of the population

General Entropy (GE) measures

These are inequality measures typically applied to income distributions. The GE measures permit decomposition of inequality into *between-group* and *within-group* components. Different variants can be specified by changing the constant which has the effect of changing the part of the income distribution afforded greatest emphasis in the calculation of the measures: Generalized Entropy class GE(a) for a = -1, 0, 1, 2. The more positive a is, the more sensitive GE(a) is to income differences at the top of the distribution; the more negative a is, the more sensitive it is to differences at the bottom of the distribution. GE(0) is the mean logarithmic deviation, GE(1) is the Theil index, and GE(2) is half the square of the coefficient of variation[58].

Gini coefficient

"Gini index measures the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution. A Lorenz curve plots the cumulative percentages of total income received against the cumulative number of recipients, starting with the poorest individual or household. The Gini index measures the area between the Lorenz curve and a hypothetical line of absolute equality, expressed as a percentage of the maximum area under the line. Thus a Gini index of 0 represents perfect equality, while an index of 100 implies perfect inequality." Source: World Bank http://data.worldbank.org/indicator/SI.POV.GINI

Inequality

Inequality refers to the unequal distribution of opportunities or outcomes (including financial resources such as income) amongst members of a given population.

Poverty

Townsend defined people as poor if "they lack the resources to obtain the types of diet, participate in the activities and have the living conditions and amenities which are

[58] Summarised from the Stata help file for the ineqdeco.ado command.

customary, or at least widely encouraged or approved in the societies to which they belong" (Townsend, 1979, p.31),

Residential Segregation

Residential segregation measures express the degree to which members of one population category are spatially segregated from those of another category (or categories). Indices of residential segregation typically quantify the extent to which population composition varies *between* neighbourhoods that lie within a larger geographical area (e.g. within a city, local municipality or some other geographical bounds).

SAIMD 2001

The South African Index of Multiple Deprivation 2001. This multidimensional index consist of five component domains which are each measured at datazone level using indicators derived from the South African 2001 Census.

Spatial inequality

Spatial inequality refers to the unequal distribution of opportunities or outcomes (or of people experiencing those opportunities or outcomes) between geographical areas within a given spatial constraint (e.g. between neighbourhoods within a country).

References

ADELZADEH, A. 2004. *South Africa Human Development Report 2003. The challenge of sustainable development in South Africa: unlocking people's creativity,* Oxford: Oxford University Press.

AFRICAN NATIONAL CONGRESS 1994. The Reconstuction and Development Programme: A Policy Framework. Umanyano Publications.

AGUERO, J., CARTER, M. R. & MAY, J. 2007. Poverty and inequality in the first decade of South Africa's democracy: what can be learnt from panel data from KwaZulu-Natal? *Journal of African Economies,* 16**,** 782-812.

ALDERMAN, H., BABITA, M., DEMOMBYNES, G., MAKHATHA, N. & OZLER, B. 2002. How low can you go? Combining census and survey data for mapping poverty in South Africa. *Journal of African Economies,* 11**,** 169-200.

ALDERMAN, H., BABITA, M., LANJOUW, J., LANJOUW, P., MAKHATHA, N., MOHAMED, A., OZLER, B. & QABA, O. 2000. Combining census and survey data to construct a poverty map of South Africa. *Measuring poverty in South Africa.* Pretoria: Statistics South Africa.

ALEXANDER, P. 2010. Rebellion of the poor: South Africa's service delivery protests - a preliminary analysis. *Review of African Political Economy,* 37**,** 25-40.

ANSELIN, L. 1995. Local Indicators of Spatial Association. *Geographical Analysis,* 27**,** 93-115.

ATKINSON, A. B. 2015a. *Inequality*: Harvard University Press.

ATKINSON, A. B. 2015b. *Inequality* - *what can be done?*: LSE International Inequalities Institute.

AVENELL, D., NOBLE, M. & WRIGHT, G. 2009. *South African datazones: A technical report about the development of a new statisitcal geography for the analysis of deprivation in South Africa at the small area level*: Centre for the Analysis of South African Social Policy, University of Oxford.

BARNES, H., NOBLE, M., DIBBEN, C., METH, C., WRIGHT, G. & CLUVER, L. 2007a. *South Africa Microdata Scoping Study,* Oxford: Centre for the Analysis of South African Social Policy, University of Oxford.

BARNES, H., NOBLE, M., WRIGHT, G. & DAWES, A. 2009. A Geographical Profile of Child Deprivation in South Africa. *Child Indicators Research,* 2**,** 181-199.

BARNES, H., WRIGHT, G., NOBLE, M. & DAWES, A. 2007b. *The South African Index of Multiple Deprivation for Children: Census 2001,* Cape Town: HSRC Press.

BECKER, G. S. 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy,* 76**,** 169.

BECKER, G. S. 1974. Crime and punishment: An economic approach. *Essays in the Economics of Crime and Punishment.* NBER.

BEEGLE, K. G., CHRISTIAENSEN, L., DABALEN, A. L. & GADDIS, I. 2016. *Poverty in a rising Africa,* Washington, D.C.: World Bank Group.

BEINART, W. 1994. *Twentieth-century South Africa,* Oxford: Oxford University Press.

BENSON, T. 2002. The poverty targeting efficiency of public works programs: An application of poverty mapping in Malawi. *International Food Policy Research Institute, Washington DC*.

BERG, A., OSTRY, J. D. & ZETTELMEYER, J. 2012. What makes growth sustained? *Journal of Development Economics,* 98**,** 149-166.

BHORAT, H. & KANBUR, R. 2005. *Poverty and Well-Being in Post-Apartheid South Africa: An overview of data, outcomes and policy,* Cape Town: Development Policy Research Unit, University of Cape Town.

BHORAT, H., NAIDOO, P. & VAN DER WESTHUIZEN, C. 2006. *Shifts in Non-Income Welfare in South Africa: 1993-2004*: Development Policy Research Unit.

BHORAT, H., POSWELL, L. & NAIDOO, P. 2004. *Dimensions of Poverty in Post-Apartheid South Africa 1996-2001*, Cape Town: Development Policy Research Unit.

BLAU 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*: Free Press.

BOTES, L. J. 2007. *The new struggle: service delivery-related unrest in South Africa: two case studies*: Centre for Development Support, University of the Free State.

BOURGUIGNON, F. 1999. Crime as a Social Cost of Poverty and Inequality: A Review Focusing on Developing Countries. *Facets of Globalization: International and Local Dimensions of Development. World Bank Discussion Paper No. 415.* World Bank.

BRANTINGHAM, P. J. & BRANTINGHAM, P. L. 1981. *Environmental criminology*: Sage Publications Beverly Hills, CA.

BURGER, R., VAN DER BERG, S., VAN DER WALT, S. & YU, D. 2004. Geography as destiny: Considering the spatial dimensions of poverty and deprivation in South Africa. *African development and poverty reduction: The macro-micro linkage.* Somerset West, South Africa.

BURSIK, R. J. 1988. Social disorganisation and theories of crime and delinquency: problems and prospects. *Criminology,* 26**,** 519-552.

CHRISTOPHER, A. J. 1988. Roots of urban segregation: South Africa at Union, 1910. *Journal of Historical Geography,* 14**,** 151-169.

CHRISTOPHER, A. J. 1989. Spatial variations in the application of residential segregation in South African cities. *Geoforum,* 20**,** 253-267.

CHRISTOPHER, A. J. 1994. *The Atlas of Apartheid,* London: Routledge.

CHRISTOPHER, A. J. 2001. Urban Segregation in Post-apartheid South Africa. *Urban Studies,* 38**,** 449-466.

CITY OF CAPE TOWN 2005. *A population profile of Khhayelitsha: Socoi-economic information from the 2001 Census*: City of Cape Town.

COHEN, L. E. & FELSON, M. 1979. Social change and crime rate trends: A routine activity approach. *American sociological review***,** 588-608.

COOK, G. P. 1991. Cape Town. *In:* LEMON, A. (ed.) *Homes apart: South Africa's segregated cities.* London: Paul Chapman Publishing.

DABLA-NORRIS, M. E., KOCHHAR, M. K., SUPHAPHIPHAT, M. N., RICKA, M. F. & TSOUNTA, E. 2015. *Causes and consequences of income inequality: a global perspective*: International Monetary Fund.

DEMOMBYNES, G. & OZLER, B. 2006. Crime and local inequality in South Africa. *In:* BHORAT, H. & KANBUR, R. (eds.) *Poverty and Policy in Post-Apartheid South Africa.* Cape Town: HSRC Press.

DENTON, N. A. 1994. Are African Americans Still Hypersegregated? *Residential apartheid: the American legacy,* 2**,** 49-81.

DEPARTMENT OF HEALTH AND SOCIAL SERVICES 1999. *Poverty in the Western Cape: An analysis of poverty in the Western Cape as enumerated in teh 1996 Census,* Cape Town: Department of Health and Social Services.

DEPARTMENT OF THE PREMIER OF THE WESTERN CAPE 2005. *Measuring the state of development in the Province of the Western Cape,* Cape Town: Western Cape: Department of the Premier.

DINKELMAN, T. 2011. The effects of rural electrification on employment: New evidence from South Africa. *The American Economic Review***,** 3078-3108.

DORLING, D. 2014a. Class segregation. *In:* LLOYD, C. D., SHUTTLEWORTH, I. & WONG, D. W. S. (eds.) *Social-spatial segregation: Concepts, processes and outcomes.* Bristol: Policy Press.

DORLING, D. 2014b. *Inequality and the 1%,* London: Verso Books.

DORLING, D. & REES, P. 2003. A nation still dividing: the British census and social polarisation 1971-2001. *Environment and Planning A,* 35**,** 1287-1313.

DUNCAN, O. D. & DUNCAN, B. 1955. A Methodological Analysis of Segregation Indexes. *American Sociological Review,* 20**,** 210-217.

DUNCAN, O. D. & LIEBERSON, S. 1959. Ethnic segregation and assimilation. *American Journal of Sociology***,** 364-374.

EGAN, K. L., ANDERTON, D. L. & WEBER, E. 1988. Relative Spatial Concentration among Minorities: Addressing Errors in Measurement. *Social Forces,* 76**,** 1115-1121.

ELBERS, C., LANJOUW, J. O. & LANJOUW, P. 2003. Micro–level estimation of poverty and inequality. *Econometrica,* 71**,** 355-364.

ELBERS, C., LANJOUW, P., MISTIAEN, J. A. & ÖZLER, B. 2008. Reinterpreting between-group inequality. *The Journal of Economic Inequality,* 6**,** 231-245.

FEDDERKE, J. W., MANGA, J. & PIROUZ , F. 2003. Challenging Cassandra: Household and Per Capita Household Income Distribution in the October

Household Surveys 1995-1999: Income and Expenditure Surveys 1995 and 2000 and the Labour Force Survey 2000. University of the Witwatersrand.

FEINSTEIN, C. H. 2005. *An Economic History of South Africa: Conquest, Discrimination and Development,* Cambridge: Cambridge University Press.

FEITOSA, F. F., CAMARA, G., MONTEIRO, A. M. V., KOSCHITZKI, T. & SILVA, M. P. 2007. Global and local spatial indices of urban segregation. *International Journal of Geographical Information Science,* 21**,** 299-323.

FINN, A. 2015. A National Minimum Wage in the Context of the South African Labour Market. Cape Town: SALDRU.

GALSTER, G. C. 1984. On the Measurement of Metropolitan Decentralization of Blacks and Whites. *Urban Studies,* 21**,** 465-470.

GHOSH, M. & RAO, J. 1994. Small area estimation: an appraisal. *Statistical science***,** 55-76.

GRAHAM, C. 2002. Public attitudes matter: a conceptual frame for accounting for political economy in safety nets and social assistance policies. *The World Bank, SPDP Series*.

GREENLAND, S. 1989. Modeling and variable selection in epidemiologic analysis. *American journal of public health,* 79**,** 340-349.

HABIB, A. 2012. SA has competing policies on creating greater equality. *Business Day*, 11 September 2012.

HARDOON, D., FUENTES-NIEVA, R. & AYELE, S. 2016. An Economy for the 1%: How privilege and power in the economy drive extreme inequality and how this can be stopped.

HENTSCHEL, J. & LANJOUW, P. 1996. *Constructing an indicator of consumption for the analysis of poverty: principles and illustrations with reference to Ecuador*: World Bank Publications.

HILLS, J. 2004. *Inequality and the State*: OUP Oxford.

HIRSCHOWITZ, R., ORKIN, M. & ALBERTS, P. 2000. Key baseline statistics for poverty measurement. *Measuring poverty in South Africa.* Pretoria: Statistics South Africa.

HOOGEVEEN, J. G. & ÖZLER, B. 2006. Poverty and inequality in post-apartheid South Africa: 1995-2000. *Poverty and policy in post-apartheid South Africa*, 59-94.

HOOVER, E. M. J. 1941. Interstate Redistribution of Population, 1850-1940. *The Journal of Economic History,* , 1, 199-205.

HORN, A. 2005. Measuring multi-ethnic spatial segregation in South African cities. *South African Geographical Journal,* 87, 58-72.

JAMES, D. R. & TAUEBER, K. E. 1985. Measures of Segregation. *Sociological Methodology,* 15, 1-32.

JOHNSTON, R. J. & PATTIE, C. 2011. Social networks, geography and neighbourhood effects. *In:* SCOTT, J. & CARRINGTON , P. (eds.) *The SAGE Handbook of Social Network Analysis.* SAGE Publications Ltd.

JOST, J. T. & BANAJI, M. R. 1994. The role of stereotyping in system-justification and the production of false consciousness. *British journal of social psychology,* 33, 1-27.

KALATI, N. & MANOR, J. 1999. Elite perceptions of poverty: South Africa. *IDS Bulletin,* 30, 117-126.

KALATI, N. & MANOR, J. 2005. Elite perceptions of poverty and poor people in South Africa. *Elite Perceptions of Povery & Inequality. London and New York: Zed Books*.

KAUFMANN D, KRAAY A & M., M. 2010. *Worldwide Governance Indicators (WGI) Project*.

KAWACHI, I., KENNEDY, B. P. & WILKINSON, R. G. 1999. Crime: social disorganization and relative deprivation. *Social science & medicine,* 48**,** 719-731.

KINGDON, G. & KNIGHT, J. 2007. Community, comparisons and subjective wellbeing in a divided society. *Journal of Economic Behavior & Organization,* 64 69–90.

KLASEN, S. 2000. Measuring Poverty and Deprivation in South Africa. *Review of Income and Wealth,* 46**,** 33-58.

KLUEGEL, J. R. & SMITH, E. R. 1986. *Beliefs about inequality: Americans' views of what is and what ought to be*: Transaction Publishers.

LEIBBRANDT, M., FINN, A. & WOOLARD, I. 2012. Describing and decomposing post-apartheid income inequality in South Africa. *Development Southern Africa,* 29**,** 19-34.

LEIBBRANDT, M., POSWELL, L., NAIDOO, P., WELCH, M. & WOOLARD, I. 2005. *Measuring Recent Changes in South African Inequality and Poverty using 1996 and 2001 Census Data*, Cape Town: Development Policy Research Unit, University of Cape Town.

LEIBBRANDT, M., WOOLARD, I., FINN, A. & ARGENT, J. 2010. *Trends in South African Income Distribution and Poverty since the Fall of Apartheid*: OECD Publishing.

LEMON, A. 1991. The apartheid city. *In:* LEMON, A. (ed.) *Homes apart: South Africa's segregated cities.* London: Paul Chapman Publishing.

LERNER, M. J. 1965. Evaluation of performance as a function of performer's reward and attractiveness. *Journal of Personality and Social Psychology,* 1**,** 355.

LESTER, A., NEL, E. & BINNS, T. 2000. *South Africa: Past, present and future,* London: Pearson Education Limited.

LICHTER, D. T., PARISI, D. & TAQUINO, M. C. 2012. The Geography of Exclusion: Race, Segregation, and Concentrated Poverty. *Social Problems,* 59**,** 364-388.

LIEBERSON, S. 1981. An Asymmetrical Approach to Segregation. *In:* PEACH, C., ROBINSON, V. & SMITH, S. (eds.) *Ethnic Segregation in Cities.* Croom Helm.

LIEBERSON, S. & CARTER, D. K. 1982. Temporal Changes and Urban Differences in Residential Segregation: A Reconsideration. *American Journal of Sociology,* 88**,** 296-310.

LLOYD, C. D. 2010. Exploring population spatial concentrations in Northern Ireland by community background and other characteristics: an application of geographically weighted spatial statistics. *International Journal of Geographical Information Science,* 24**,** 1193-1221.

LLOYD, C. D. & SHUTTLEWORTH, I. 2012. Residential segregation in Northern Ireland in 2001: assessing the value of exploring spatial variations. *Environment and Planning A,* 44**,** 52-67.

LOGAN, J. R. & STULTS, B. J. 2011. *The Persistence of Segregation in the Metropolis: New Findings from the 2010 Census*.

MAKIWANE, M. & CHIMERE-DAN, D. 2010. The people matter: The state of the population in the Eastern Cape. *Eastern Cape Department of Social Development, East London*.

MASSEY, D. S. & DENTON, N. A. 1988. The dimensions of residential segregation. *Social Forces,* 67**,** 281-315.

MASSEY, D. S. & DENTON, N. A. 1989. Hypersegregation in U.S. Metropolitan Areas: Black and Hispanic Segregation along Five Dimensions. *Demography,* Vol. 26**,** 373-391.

MASSEY, D. S. & DENTON, N. A. 1993. *American apartheid: Segregation and the making of the underclass*: Harvard University Press.

MASSEY, D. S. & DENTON, N. A. 1998. The Elusive Quest for the Perfect Index of Concentration: Reply to Egan, Anderton, and Weber. *Social Forces,* 76**,** 1123-1133.

MASSEY, D. S. & EGGERS, M. L. 1993. The Spatial Concentration of Affluence and Poverty during the 1970s. *Urban Affairs Review* 29**,** 299-315.

MASSEY, D. S. & FISCHER, M. J. 2000. How segregation concentrates poverty. *Ethnic and Racial Studies,* 23**,** 670–691.

MASSEY, D. S. & FISCHER, M. J. 2003. The Geography of Inequality in the United States, 1950-2000. *Brookings-Wharton Papers on Urban Affairs***,** 1-40.

MASSEY, D. S., ROTHWELL, J. & DOMINA, T. 2009. The Changing Bases of Segregation in the United States. *The annals of the American Academy of Political and Social Science***,** 74-90.

MASSEY, D. S., WHITE, M. J. & PHUA, V.-C. 1996. The Dimensions of Segregation Revisited. *Sociological Methods & Research* 25**,** 172-206.

MATTES, R., BRATTON, M. & DAVIDS, Y. D. 2003. *Poverty, survival and democracy in Southern Africa,* Cape Town/Accra/East Lansing.

MAY, J. 1998. *Poverty and Inequality in South Africa,* Durban: Praxis Publishing.

MCINTYRE, D., MUIRHEAD, D., GILSON, L., GOVENDER, V., MBATSHA, S., GOUDGE, J., WADEE, H. & NUTUTELA, P. 2000. *Geographic patterns of deprivation and health inequalities in South Africa: Informing public resource allocation strategies*: University of Cape Town, University of Witwatersrand, London School of Tropical Hygiene and Medicine, and National Department of Health.

MCKENZIE, D. J. 2003. *Measuring Inequality with Asset Indicators*: BREAD (Bureau for Research in Economic Analysis of Development).

MCLENNAN, D., NOBLE, M. & WRIGHT, G. 2016. Developing a spatial measure of exposure to socio-economic inequality in South Africa. *South African Geographical Journal,* 98**,** 254-274.

MCLENNAN, D., ROBERTS, B., NOBLE, M., MASILELA, T. & MAGIDIMISH, H. 2014. *Exploring the relationship between spatial inequality and attitudes to inequality in South Africa.*: Centre for the Analysis of South African Social Policy (CASASP), University of Oxford.

METH, C. 2006. *What was the poverty headcount in 2004 and how does it compare to recent estimates by van der Berg et al?*: Southern Africa Labour and Development Research Unit, University of Cape Town.

METH, C. 2007. Flogging a dead horse: Attempts by van der Berg et al to measure changes in poverty and inequality.

METH, C. & DIAS, R. 2003. *Increases in poverty in South Africa, 1999-2002,* Durban: University of Natal.

MORGAN, B. S. 1975. The segregation of socio-economic groups in urban areas: a comparative analysis. *Urban Studies,* 12**,** 47-60.

MORGAN, B. S. 1983. An Alternate Approach to the Development of a Distance-Based Measure of Racial Segregation. *American Journal of Sociology,* 88**,** 1237-1249.

NDP 2012. National Development Plan 2030: Our Future Make it Work. *In:* PRESIDENCY (ed.). Pretoria, RSA.

NEVES, D. & DU TOIT, A. 2010. Working on the margins: Poverty and economic marginality in South Africa. HeinOnline.

NGP 2010. New Growth Path. *In:* DEPARTMENT OF ECONOMIC DEVELOPMENT (ed.). RSA.

NOBLE, M., BABITA, M., BARNES, H., DIBBEN, C., MAGASELA, W., NOBLE, S., NTSHONGWANA, P., PHILLIPS, H., RAMA, S., ROBERTS, B., WRIGHT, G. & ZUNGU, S. 2006a. *The Provincial Indices of Multiple Deprivation for South Africa 2001,* Oxford: Centre for the Analysis of South African Social Policy.

NOBLE, M., BABITA, M., BARNES, H., DIBBEN, C., MAGASELA, W., NOBLE, S., NTSHONGWANA, P., PHILLIPS, H., RAMA, S., ROBERTS, B., WRIGHT, G. & ZUNGU, S. 2006b. *The Provincial Indices of Multiple Deprivation for South Africa 2001: Technical Report,* Oxford: Centre for the Analysis of South African Social Policy.

NOBLE, M., BARNES, H., SMITH, G. A. N., MCLENNAN, D., DIBBEN, C., ANTTILA, C., SIGALA, M., AVENELL, D., SMITH, T. & MOKHTAR, C. 2005. *Northern Ireland Multiple Deprivation Measure 2005,* Belfast: Northern Ireland Statistics and Research Agency.

NOBLE, M., BARNES, H., WRIGHT, G., MCLENNAN, D., AVENELL, D., WHITWORTH, A. & ROBERTS, B. 2009a. *The South African Index of*

Multiple Deprivation 2001 at Datazone level, Pretoria: Department of Social Development.

NOBLE, M., BARNES, H., WRIGHT, G. & ROBERTS, B. 2009b. Small area indices of multiple deprivation in South Africa. *Social Indicators Research,* 95**,** 281-297.

NOBLE, M., NTSHONGWANA, P. & SURENDER, R. 2008. *Attitudes to work and social security in South Africa*: HSRC Press Cape Town.

NOBLE, M., SMITH, G., PENHALE, B., WRIGHT, G., DIBBEN, C., OWEN, T. & LLOYD, M. 2000a. *Measuring multiple deprivation at the small area level: The Indices of Deprivation 2000,* London: Department of the Environment, Transport and the Regions.

NOBLE, M., SMITH, G. A. N., WRIGHT, G., DIBBEN, C. & LLOYD, M. 2001. *Northern Ireland Multiple Deprivation Measure 2001,* Belfast: Northern Ireland Statistics and Research Agency.

NOBLE, M., SMITH, G. A. N., WRIGHT, G., DIBBEN, C., LLOYD, M. & PENHALE, B. 2000b. *Welsh Index of Multiple Deprivation 2000,* Cardiff: National Assembly for Wales.

NOBLE, M. & WRIGHT, G. 2013a. Ring of fire: Socially perceived necessities in informal urban settlements in South Africa. *Policy & Politics,* 41**,** 259-277.

NOBLE, M. & WRIGHT, G. 2013b. Using indicators of multiple deprivation to demonstrate the spatial legacy of Apartheid in South Africa. *Social Indicators Research,* Vol. 112**,** 187-201.

NOBLE, M., WRIGHT, G., DIBBEN, C., SMITH, G. A. N., MCLENNAN, D., ANTTILA, C., BARNES, H., MOKHTAR, C., NOBLE, S., AVENELL, D., GARDNER, J., COVIZZI, I. & LLOYD, M. 2004. *The English Indices of*

Deprivation 2004, London: Neighbourhood Renewal Unit, Office of the Deputy Prime Minister.

NOBLE, M., WRIGHT, G., LLOYD, M., DIBBEN, C., SMITH, G., RATCLIFFE, A., MCLENNAN, D., SIGALA, M. & ANTTILA, C. 2003. *Scottish Index of Deprivation 2003: Summary Report* Edinburgh: Scottish Executive.

NOBLE, M., ZEMBE, W., WRIGHT, G. & AVENELL, D. 2013. Multiple Deprivation and Income Poverty at Small Area Level in South Africa in 2011. *Cape Town: Southern African Social Policy Research Institute and Southern African Social Policy Research Insights (SASPRI).*

NOBLE, M., ZEMBE, W., WRIGHT, G., AVENELL, D. & NOBLE, S. 2014. *Income Poverty at Small Area Level in South Africa in 2011.,* Cape Town: SASPRI.

OPENSHAW, S. 1984. *The modifiable areal unit problem*: Geo Abstracts University of East Anglia.

OPENSHAW, S. & BAXTER, R. 1977. Algorithm 3: a procedure to generate pseudo-random aggregations of N zones into M zones, where M is less than N. *Environment and Planning A,* 9**,** 1423-1428.

ORTON, M. & ROWLINSON, K. 2007. *Public attitudes to economic inequality,* Bristol, UK: The Policy Press for the Joseph Rowntree Foundation.

PARNELL, S. M. & PIRIE, G. H. 1991. Johannesburg. *In:* LEMON, A. (ed.) *Homes apart: South Africa's segregated cities.* London: Paul Chapman Publishing.

PARRY, K. & VAN EEDEN, A. 2015. Measuring racial residential segregation at different geographic scales in Cape Town and Johannesburg. *South African Geographical Journal,* 97**,** 31-49.

PICKETT, K. E. & WILKINSON, R. G. 2015. Income inequality and health: A causal review. *Social Science & Medicine,* 128**,** 316-326.

PIKETTY, T. 1995. Social mobility and redistributive politics. *The Quarterly journal of economics,* 110**,** 551-584.

PIKETTY, T. 2014. *Capital in the twenty-first century*: Harvard University Press.

POSEL, D. R. & CASALE, D. M. 2011. Relative Standing and Subjective Well-Being in South Africa: The Role of Perceptions, Expectations and Income Mobility. *Social Indicators Research,* 104**,** 195-223.

REARDON, S. F. & FIREBAUGH, G. 2002. Measures of multigroup segregation. *Sociological methodology,* 32**,** 33-67.

REARDON, S. F. & O'SULLIVAN, D. 2004. Measures of Spatial Segregation. *Sociological Methodology,* 34**,** 121–162.

REPUBLIC OF SOUTH AFRICA 1996. The Constitution of the Republic of South Africa Act 108 or 1996. Constitutional Assembly.

ROBERTS, B. 2006. The happy transition? Attitudes to poverty and inequality after a decade of democracy. *In:* PILLAY, U., ROBERTS, B. & RULE, S. (eds.) *South African Social Attitudes: Changing Times, Diverse Voices.* Cape Town, South Africa: HSRC Press.

ROBERTS, B. J. 2014. Your place or mine? Beliefs about inequality and redress preferences in South Africa. *Social Indicators Research,* 118**,** 1167-1190.

ROWLINGSON, K. 2011. *Does income inequality cause health and social problems?*

SAMPSON, R. J. 2006. How Does Community Context Matter? Social Mechanisms and the Explanation of Crime. *The explanation of crime: Context, mechanisms and development.* Cambridge: Cambridge University Press.

SAMPSON, R. J. & GROVES, W. B. 1989. Community structure and crime: Testing social-disorganization theory. *American journal of sociology,* 94**,** 774-802.

SAMPSON, R. J., MORENOFF, J. D. & EARLS, F. 1999. Beyond social capital: Spatial dynamics of collective efficacy for children. *American sociological review,* 64**,** 633-660.

SAMPSON, R. J., RAUDENBUSH, S. W. & EARLS, F. 1997. Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science,* 277**,** 918-924.

SAMPSON, R. J., RAUDENBUSH, S. W. & EARLS, F. 1998. *Neighborhood Collective Efficacy--: Does it Help Reduce Violence?*: US Department of Justice, Office of Justice Programs, National Institute of Justice.

SAUNDERS, P. 2010. *Beware False Prophets: Equality, the Good Society and The Spirit Level*: Policy Exchange.

SCHNELL , I. & YOAV, B. 2001. The Sociospatial Isolation of Agents in Everyday Life Spaces as an Aspect of Segregation. *Annals of the Association of American Geographers,* 91**,** 622-636.

SHAW, C. R. & MCKAY, H. D. 1942. *Juvenile Delinquency in Urban Areas,* Chicago: University of Chicago Press.

SHORROCKS, A. F. 1980. The class of additively decomposable inequality measures. *Econometrica: Journal of the Econometric Society***,** 613-625.

SIMKINS, C. 2005. *What Happened to the Distribution of Income in South Africa between 1995 and 2001?,* Johannesburg: University of the Witwatersrand.

SNOWDON, C. 2010. *The Spirit Level Delusion,* Ripon: Little Dice.

STATISTICS SOUTH AFRICA 2002. *Earning and spending in South Africa: Selected findings and comparisons from the income and expenditure surveys of October 1995 and October 2000*.

STATISTICS SOUTH AFRICA 2004a. *Census 2001 Metadata: Geography hierarchy and attributes,* Pretoria: Statistics South Africa.

STATISTICS SOUTH AFRICA 2004b. *Census 2001: Concepts and Definitions*.

STATISTICS SOUTH AFRICA 2012. *Income and Expenditure of Households 2010/2011*.

STATISTICS SOUTH AFRICA & HSRC 2007. *Using the 2001 Census: Approaches to analysing data,* Pretoria.

STIGLITZ, J. 2012. *The price of inequality,* London: Penguin UK.

SURENDER, R., NOBLE, M., WRIGHT, G. & NTSHONGWANA, P. 2010. Social assistance and dependency in South Africa: An analysis of attitudes to paid work and social grants. *Journal of Social Policy,* 39**,** 203-221.

TERREBLANCHE, S. 2002. *A History of Inequality in South Africa 1652-2002,* Scottsville: University of Natal.

TOBLER, W. R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic geography***,** 234-240.

TODES, A., KOK, P., WENTZEL, M., VAN ZYL, J. & CROSS, C. Contemporary South African urbanization dynamics. Urban forum, 2010. Springer, 331-348.

TOMLINSON, R., ABRAHAMS, G. & GILDENHUYS, B. 2003. The changing nature of South African housing demand. *South Africa: Mega-Tech*.

TOWNSEND, P. 1979. *Poverty in the United Kingdom: A survey of household resources and standards of living* London: Penguin Books.

TOWNSEND, P., PHILLIMORE, P. & BEATTIE, A. 1988. *Health and deprivation: inequality and the North,* London: Routledge.

TUROK, I. & BOREL-SALADIN, J. 2014. Is urbanisation in South Africa on a sustainable trajectory? *Development Southern Africa,* 31**,** 675-691.

UNDP 2003. *South Africa: The challenge of sustainable development*.

UNDP 2015. *Human Development Report 2015*.

UNITED NATIONS 1995a. Copenhagen Declaration on Social Development. *World Summit for Social Development.* Copenhagen, March.

UNITED NATIONS 1995b. Report of the World Summit for Social Development. *World Summit for Social Development.* Copenhagen.

UNODC 2011. *Global Study on Homicide: Trends, Contexts, Data*: United Nations Office on Drugs and Crime (UNODC).

VAN DER BERG, S., BURGER, R., BURGER, R., LOUW, M. & YU, D. 2005. *Trends in poverty and inequality since the political transition,* Stellenbosch: Bureau for Economic Research.

VAN DER BERG, S. & LOUW, M. 2004. Changing patterns of South Afircan Income Distribution - towards time series estimates of distribution and poverty. *South African Journal of Economics,* 72**,** 546-572.

VAN DER BERG, S., LOUW, M. & YU, D. 2008. Post-transition poverty trends based on an alternative data source. *South African Journal of Economics,* 76**,** 58-76.

VOGT, W. P. 1999. *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences,* Thousand Oaks, CA: SAGE Publications.

WEISBURD, D., GROFF, E. R. & YANG, S.-M. 2014. The importance of both opportunity and social disorganization theory in a future research agenda to advance criminological theory and crime prevention at places. *Journal of research in crime and delinquency*.

WHITE, M. J. 1983. The Measurement of Spatial Segregation. *American Journal of Sociology,* 88**,** 1008-1018.

WIKSTROM, P.-O. H. 2004. Crime as alternative: Towards a cross-level situational action theory of crime causation. *In:* MCCORD, J. (ed.) *Beyond empiricism: Institutions and intentions in the study of crime.* New Brunswick, NY: Transaction.

WIKSTROM, P.-O. H. 2006a. Individuals, settings, and acts of crime: Situational mechanisms and the explanation of crime. *The explanation of crime: Context, mechanisms and development.* Cambridge: Cambridge University Press.

WIKSTROM, P.-O. H. 2006b. Situational action theory. *In:* WILCOX, F. C. A. P. (ed.) *Encyclopaedia of criminological theory.* Beverly Hills, CA: Sage.

WIKSTRÖM, P.-O. H. 2005. The social origins of pathways in crime: Towards a developmental ecological action theory of crime involvement and its changes. *Integrated developmental and life-course theories of offending,* 14**,** 211-245.

WIKSTRÖM, P.-O. H. 2014. Why crime happens: A situational action theory. *In:* MANZO, G. (ed.) *Analytical Sociology.* Chichester, United Kingdom: John Wiley & Sons, Ltd.

WIKSTRÖM, P.-O. H. & TREIBER, K. H. 2009. Violence as Situational Action. *International Journal of Conflict and Violence,* Vol 3.

WILKINSON, R. 1996. Health and society: The afflictions of inequality. London, England: Routledge.

WILKINSON, R. G. & PICKETT, K. 2010. *The Spirit Level: Why Equality Is Better for Everyone,* London: Allen Lane.

WILSON, G. D. & PATTERSON, J. R. 1968. A new measure of conservatism. *British Journal of Social and Clinical Psychology*, 7**,** 264-269.

WITTENBERG, M. & LEIBBRANDT, M. 2015. *Measuring inequality by asset indices: A general approach with application to South Africa*.

WOLPE, H. 1988. *Race, Class and the Apartheid State,* Paris: Unesco Press.

WONG, D. W. 1998. Measuring multiethnic spatial segregation. *Urban Geography,* 19**,** 77-87.

WONG, D. W. S. 2002. Modeling Local Segregation: a Spatial Interaction Approach. *Geographical & Environmental Modelling,* 6**,** 81-97.

WOOLARD, I. D. & WOOLARD, C. D. 2006. *Earnings Inequality in South Africa, 1995-2003*: HSRC Press.

WRIGHT, G., BARNES, H., NOBLE, M. & DAWES, A. 2009a. *The South African Index of Multiple Deprivation for Children 2001 at Datazone level,* Pretoria: Department of Social Development.

WRIGHT, G. & NOBLE, M. 2009. *The South African Index of Multiple Deprivation 2007 at Municipality level,* Pretoria: Department of Social Development.

WRIGHT, G. & NOBLE, M. 2013. Does Widespread Lack Undermine the Socially Perceived Necessities Approach to Defining Poverty? Evidence from South Africa. *Journal of Social Policy,* 42**,** 147-165.

WRIGHT, G., NOBLE, M., BARNES, H. & NOBLE, S. 2009b. *The South African Index of Multiple Deprivation for Children 2007 at Municipality level,* Pretoria: Department of Social Development.