MASTER OF ECONOMIC SCIENCE
RESEARCH REPORT

Poverty in South Africa:

An Analysis of Former Vs Non-Former Homeland Areas

A Research Report submitted in partial fulfilment of the

Degree of Master of Economic Science in the

School of Economic and Business Sciences,

University of the Witwatersrand

by

Name: Lesego Masenya

Student No: 716839

Supervised by:

Professor Prudence Magejo

Word Count: 9 933

Date: 29 March 2019

"***The financial assistance of the National Treasury and the National Research Fund (grant number 112119) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and should not be attributed to the funders"***

Declaration

I, Lesego Masenya, declare that this dissertation titled, "*Poverty in South Africa: An Analysis of Former Vs Non-Former Homeland Areas"* has not been previously accepted in substance for any degree nor is not being concurrently submitted for any degree at another university. I further state that this degree is the result of my own independent work, except where otherwise stated. All sources used or referenced in the study have been explicitly acknowledged.

Signed   Date: 29 March 2019

____

Abstract

The objective of the study is to analyse the effect former homeland status on poverty in South Africa. The study uses 2011 Census community profiles data from Statistics South Africa and cartographic data. Two methodologies are used in order to identify the effect of former homeland status on poverty, i.e., Ordinary Least Squares (OLS) and Regression Discontinuity Design (RDD). Notably, the RDD model is the main model as it formally identifies the treatment effect by comparing former and non-former homelands within a quasi-experimental framework. The results indicate that former homeland areas experience higher poverty levels relative to non-former homeland areas. The analysis shows that a large portion of the "raw" poverty differential is explained by differences in observed characteristics between former and non-former homeland areas. The remaining difference is attributable to former homeland status. The 'scarring effect' is small but statistically significant. Thus, the results call for government intervention aimed at reducing differences in observed characteristics of former and non-former homeland areas. The study notes that such mechanisms will narrow the difference in poverty rates but might not close it entirely since part of the difference is structural and depends on the rate at which the 'scarring effect' fades overtime.

Acknowledgements

I would like to thank my supervisor, Professor Prudence Magejo. I owe her a great sense of gratitude for the invaluable support she offered throughout my research and constantly availing herself as often as I needed her. She guided me patiently and consistently, and without her support, I never would have been able to complete this thesis. I would also like to express my gratitude to the lecturers and administrators from the School of Economic and Business Sciences who were always willing to offer their assistance where possible. Finally, a big thank you to my family and friends for constantly supporting me throughout my studies.

1. Introduction

As is well known - residential segregation served as strong pillar of racial discrimination during apartheid. A key point to consider is the marginalisation of the Black majority from mainstream socio-economic and political activities through the creation of Bantustans or homeland areas. This resulted in overcrowded, poverty-stricken communities with little economic opportunities. Fortunately, the dismantling of the Apartheid regime opened up room for several remedial measures e.g. Black Economic Empowerment, social grants, housing, education and public health policies (Jensen and Zenker, 2015). Though some level of success has been achieved from these rehabilitative measures, the contemporary relevance of Apartheid through its legacies still poses a complex challenge to South Africa's socio-economic transformation agenda. To add, South Africa's racial profile on socio-economic outcomes is still reflective of the Apartheid hierarchy (Mariotti and Fourie, 2014) – despite government's concerted efforts towards redressing past injustices (Gwanya, 2010). As such, with apartheid still being a topical discussion, the bulk of the studies (e.g. Van der Berg, 2011; Leibbrandt, Woolard and Woolard, 2012; Bhorat and Van der Westhuizen 2013; Harris et al 2011; Banerjee et al., 2008) unanimously note that the problems of inequality, poverty and unemployment remain high and persistent among Black South Africans.

Previous studies (e.g. Burger, Van der Berg, van der Walt and Yu; 2017; Gwanya, 2010; Turok, 1994; Turok, 2012) allude to the fact that South Africa's poverty and inequality patterns are deeply rooted in the repressive institutions of the Apartheid regime. Limited scholarship on the long-term effects of Apartheid policies - particularly spatial policies – coupled with data constraints make these dynamics hard to understand. However, through recent advances in the digitization of data, availability of information on the Apartheid era can now be facilitated (Mariotti and Fourie, 2014). The exploitation of this type of data has enabled recent studies to examine the development and failure of industrial zones in homelands, the effects of homelands on inter-ethnic trust, social capital, political outcomes and the evolution of spatial inequality (see Kerby, 2014; Abel, 2015; De Kadt and Larreguy, 2014; Bastos and Bottan, 2014; von Fintel, 2018). It is against this background that this paper adds to a small but growing literature on the long-term effects of Apartheid policies with a particular focus on the homeland policy. To this end, this study investigates the effect of former homeland status on poverty. While previous studies (Aliber, 2003; Carter and May; 1999) suggest that poverty is high in former homeland areas, they do not explicitly quantify the magnitude of the differential neither do they pin-point if the former homeland policy per-se ('scarring effect') underpin observed patterns. The present study uses the 2011 Census community profiles and employs OLS and a regression discontinuity design to identify this effect.

The rest of this research report is structured as follows; section 2 provides a brief background to poverty and homeland policy in South Africa, conceptual framework on poverty as well as a review of the empirical literature. The methodology is discussed in section 3, followed by the data description in section 4. Section 5 presents the empirical results, while section 6 concludes.

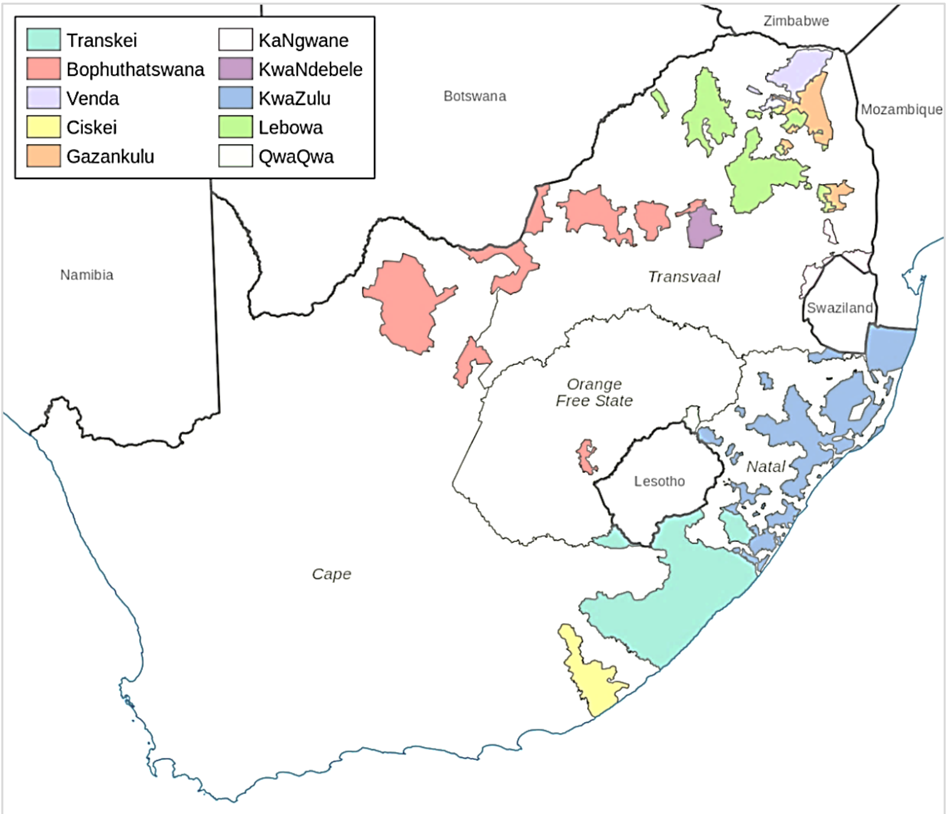
2. Background and Literature Review

2.1. Background: Homeland policy

The 1913 Natives Land Act denied Blacks the right to purchase or lease from whites when outside of the areas demarcated by the Act. This laid the foundation for the apartheid policy of homeland areas (Pienaar and Von Fintel; 2013). Following the establishment of the apartheid state in 1948, the bulk of the legislation put in place was detrimental to the black population. Some of these include; the 1950 Population Registration Act, the 1950 Groups Areas Act, the 1951 Bantu Authorities Act, the 1959 Promotion of Bantu Self-Governing Act, the 1953 Bantu Education Act, the 1953 Reservation of Separate Amenities Act, the 1961 Urban Bantu Councils Act, the 1970 Bantu Homelands Citizen's Act and the use of 'pass laws'.

It was during the 1960s and the early 1980s that the apartheid government executed a ruthless policy of 'resettlement' to coerce people to migrate to their designated urban 'group areas' and rural 'homelands' (Bastos and Bottan; 2014). This period saw over two million people evicted and relocated to homelands (Terreblanche, 2002; Turok, 1976; Turok, 2012). This resulted in overcrowding in homeland areas; the Black population comprised of more than 70 percent of the country's population but only 13 percent of the total land was allocated to the 10 African homelands. As indicated by Figure 1, homeland areas were mainly located in the periphery of the country thereby marginalizing the Black populace from mainstream economic activities (Hopkins, 2001; Leibbrandt, Woolard and Woolard 2000).

Figure 1: A map of South African homeland areas pre-1994



Source: South Africa Gateway: Alexander (2018)

It is also worth noting that homeland areas were located in the least productive regions agriculturally (Lahiff and Cousins, 2005; Porter and Phillips-Howard, 1997). The poor physical environment complimented by underinvestment by the apartheid government resulted in poor economic outcomes among Black South Africans (Burger et al, 2017; Lahiff and Cousins, 2005; Leibbrandt, Woolard and Woolard 2000; Makgetla, 2010; Turok, 2012). With very little economic opportunities available in homeland areas, the advancement of the manufacturing sector in urban areas - which was labour intensive - lead to an increase in rural-urban migrant flows. To restrict this industrial development and further confine the Black population to homeland areas, the government created industrial zones which Turok (2012) notes, was a ploy to reduce rural-urban migration. This, as the literature explains, was a way of ensuring that the Black population continued to remain oppressed, and were unable to access economic opportunities (Burger et al, 2017; Turok, 2012). The discontinuation of industrial zones in the late 1980s onwards resulted in an increased migration of people to industrialized areas with more jobs opportunities (David, et al., 2018; Harrison, 2013).

2.2. Conceptual framework: Poverty measurement

The study of poverty dates back to the 1800s, and is still a topical issue to date. Notably, a wealth of information has resulted from studies that examine poverty, thus leading to further studies that are diverse and are of great importance (Davids, 2010). However, the lack of distinction between conceptualizing, defining and measuring poverty is frequently misunderstood when conducting research on poverty (Davids, 2010; Noble, Ratcliffe and Wright, 2004). As such, this conceptual framework briefly discusses the different definitions of poverty and associated measurement issues.

When defining poverty based on absolute deprivation, the United Nations[1] (1995) adopted two definitions: absolute poverty and overall poverty. The former is characterised by a severe scarcity of basic human needs e.g. food, safe drinking water, health, shelter, education and information (Lok-Dessalien; 2002; Ravallion, Datt and van de Walle, 1991; Cutler, 1984); while the latter describes a lack of income and productive resources to ensure sustainable livelihoods (Hanmer, Pyatt and White, 1999; Kanbur, 1987; Hagenaars and Praag, 1985). With there being a need to incorporate various dimensions to the definition such as lack of opportunities, insecurity and social exclusion, poverty is now widely understood as a multidimensional phenomenon (Sen, 1997; Townsend, 1979).

Focusing on poverty as a single dimensional problem, previous literature often operationalizes it using the money metric approach where the poor are identified based on an absolute or relative poverty line. The money metric approach is often based on a method pioneered by Rowntree (1902) which focuses on income and consumption (Davids, 2010). This method defines poor people by making use of the poverty datum line – a threshold derived from income, and perceives poor people as those earning below the threshold and non-poor as earning above the line (Davids, 2010; van Praag, Hagenaars and van Weerden, 1982). The relative approach serves as a criticism to the absolute approach which fails to link poverty status to the average welfare in society (Townsend, 1979). Townsend (1979) argues that the use of absolute poverty lines may not serve as a real measure of poverty, especially in developed countries where people are able to meet their basic needs but can still be considered as poor (Davids, 2010). For the purposes of

[1] http://www.unesco.org/new/en/social-and-human-sciences/themes/international-migration/glossary/poverty/

the study, poverty is operationalized from a single dimension perspective and is assessed using the money-metric and absolute rather than relative poverty lines.

With a poverty line, the extent of poverty can be examined by constructing poverty measures. The most popular poverty measures are those proposed by Foster-Greer-Thorbecke (1984) derived as follows:

$$P_\alpha = \frac{1}{n}\sum_{i=1}^{q}\left[\frac{z-y_i}{z}\right]^\alpha \qquad (1)$$

where z denotes the poverty line, n is the total population, q is the number of poor households, y_i represents income or expenditure while α is the poverty aversion measure (the higher the value, the greater the weight placed on the most deprived). When $\alpha = 0$, equation (1) yields the poverty head count measure which is simply the incidence of poverty (i.e., proportion of the population whose income/expenditure lies below the poverty line). When $\alpha = 1$, this yields the difference in poverty rates measuring the mean aggregate income of each individual in the population around the poverty line. When $\alpha = 2$, the poverty severity measure is obtained. Due to data constraints, this study will focus on the poverty head count measure.

2.3 Previous studies

Evaluating the role that historical institutions play in explaining development is key in understanding current economic outcomes (e.g. poverty, inequality, employment) globally (Dell, 2010). Notably, several studies exist on the international arena. For instance, studies conducted in Europe, Africa, South America and the US show how slavery (Nunn, 2008); colonialism (Acemoglu et al., 2001); and *mita* (forced labour system in Peru) (Dell, 2010) have a direct impact on present-day overall welfare. Of all the studies mentioned, the one closest in comparison in terms of historical persistence is Dell's (2010) paper. Her study evaluates the long-term effects of the *mita* to determine if there were any prevalent effects to date. Not all districts in Peru were subjected to participate in the *mita,* as such, the regression discontinuity model was estimated (a similar model estimated in this study). The results obtained indicate that due to the *mita,* there has been a decline in household consumption and an increase in stunted growth in children, a sign of poverty. To date, regions affected by the *mita* are less integrated to the road infrastructure and have lower educational attainment. As a result; residents within the affected areas are more likely to be subsistence farmers as a means of survival. These are effects that highlight the prevalent effects of the *mita* and show how the *mita* has impacted on economic development. While the above mentioned studies, with the exception Dell (2010), offer little guidance in distinguishing the underpinning mechanisms, what is clear is that historical institutions have an impact on contemporary socio-economic outcomes.

In South Africa, a number of studies exist that consider the effect of apartheid on socio-economic outcomes. In particular, studies that consider the homeland policy (e.g. Kerby, 2014; Abel, 2015; De Kadt and Larreguy, 2014; Bastos and Bottan, 2014; von Fintel, 2018) focus on outcomes other than poverty i.e., inter-ethnic trust, social capital, political outcomes and the evolution of spatial inequality. From the extant literature, the effect of former homeland status on poverty and hence the difference in poverty rates between former and non-former homeland areas is not well-established. Yet, this knowledge is pertinent for policy especially within this age of socio-economic transformation.

The lack of studies that zero in on estimating differences in the poverty rates -between former and non-former homeland areas does not negate the existence of existing poverty studies in South Africa. Interestingly, there are quite several studies that are concerned with poverty in the Republic. However, these consider the impact of apartheid policies in a broad sense via race or broad geographical units. To date, studies exist that consider the evolution of poverty (e.g., Adato et. al., 2006; Bhorat and Kanbur, 2006; Carter and May 1999); its determinants (Leibbrandt, Woolard and Woolard, 2000); distribution across various dimensions such as age, gender and race (e.g., Burger, et al., 2017; Aliber, 2003; Woolard, 2012; Yamauchi, 2005; David et al, 2018; Noble and Wright, 2013); and across space (Hoogeveen and Özler, 2006). A common finding across these studies is that Black South Africans suffer from high levels of poverty despite many government corrective measures (e.g. social grants and BEE policies) and nationwide infrastructure programs that have resulted in improved access basic amenities.

For studies incorporating a spatial dimension, Naudé, McGillivray and Rossouw, (2009) constructed a local level vulnerability index for 354 magisterial districts over the period 1996 - 2001. The study found high levels of vulnerability in areas that are remote and isolated from main cities. This highlights the lack of economic progress in former homeland areas. Another study, Burger et al. (2017) evaluates poverty by location and found evidence which gives an understanding of the enduring apartheid legacy. The results suggest that there is correlation between place and poverty; provinces that do not include substantial parts of former homeland areas (i.e., Western Cape, Gauteng and the Northern Cape) exhibit low levels of deprivation. It is also in these provinces that urban-rural migration is peeking[2]. This is due to the employment opportunities that exist, better educational and health care services and wage differences (Lee, 1966; Mlambo, 2018; Ntshidi, 2017). However, those individuals who choose not to migrate may do so due to structural rigidities and affordability that may make it difficult to relocate. These include higher cost of living, insufficient skills sets and poverty (Ntshidi, 2017). As such, migration from former homelands into non-former homelands is not as easy as crossing the homeland boundary borders.

Overall, previous studies unanimously agree that past apartheid policies have and are still contributing immensely to observed poverty patterns. This study complements the extant literature by explicitly measuring the effect of former homeland status on poverty. Accordingly, the magnitude of the poverty rate difference attributable to the homeland policy, if any, will be identified.

3. Methodology

The study employs two methodologies to identify the effect of former homeland status on poverty; Ordinary Least Squares (OLS) estimation and Regression Discontinuity Design (RDD). The study is also able to attribute the poverty rate differences to the 'scarring effect' – defined as the long term effect that unemployment has on the future labour force (Nilsen and Reiso, 2011). However, 'scarring' in the context of the study simply refers to the effect of former homeland area status on poverty. Using South Africa's cartographic data, main places located (100%) inside and (100%) outside former homelands are identified. The former will be collectively referred to as former homeland areas and the latter as non-former homeland areas hereafter. Notably, there are some main places (1,549 out of 14, 039) straddling former homeland boundaries; these are excluded in the analysis to operationalise a clean sharp RDD.

[2] http://www.statssa.gov.za/publications/Report-03-04-02/Report-03-04-02.pdf

Ordinary Least Squares

First, the following poverty function is estimated using OLS:

$$P_i = \alpha_0 + \lambda HOM_i^d + \sigma X_i + \theta_R + \mu_i \qquad (2)$$

Where P_i is the poverty headcount measure for main place *i;* HOM_i^d is a dummy variable = 1 for former homeland areas and 0 otherwise; X_i is the vector containing main place characteristics as defined in Table A.1 in the Appendix. α_0, σ, λ and η are parameters to be estimated and μ_i is the error term. The coefficient of interest, λ, is expected to be positive; this would suggest that former homeland status is associated with higher levels of poverty.

Equation (2) is likely to yield biased results as it uses all observations including those located further from former homeland areas (e.g. main places in the Western Cape Province where there are no former homeland areas). Indeed, comparing main places that are too different from each other (e.g. differences in terrain, local governments etc.) might confound the estimates. To account for this source of confounding factors, the study restricts estimates by distance around former homeland border. Thus, OLS regressions are estimated around a 5km, 10km, 15km and 20km radius. This approach mimics RD design which is the second approach used to formally identify the treatment effect.

Regression Discontinuity Design

A proper identification of the effect of former homeland status on poverty is achieved by use of a sharp RD design. This exploits a sharp discontinuity in the outcome variable at a given threshold of a running variable (Imbens and Lemieux, 2008; Lee and Lemieux, 2010). For validity, two key identifying assumptions should be satisfied: the local randomisation assumption (which requires treatment to be as good as random at the cut-off) and the local continuity assumption (which requires potential outcomes to be continuous around the threshold). The former assumption is satisfied if communities do not have perfect control or cannot manipulate their location relative to the cut-off point (Lee, 2008). Arguably, the assumption appears reasonable in this application – observed main place formations are considered to be exogenous and were created after a long period of historical apartheid policies (c.f. Abel, 2015, Bakker et al., 2016). Moreover, the large-scale removals and resettlement of Africans to the homelands during the 1960s and 1970s was an exogenous shock to the living and social conditions of Africans (Abel, 2015). It is in this context that homeland status is used as a natural experiment.

To operationalise the sharp RD design, distance to/from former homeland boundary is used as a running variable, *z*. Distance of each main place to a former homeland boundary, $Dist_i$, is measured in kilometers (km) as the shortest distance of a main place (centroid) to the closest former homeland border. Using this running variable, treatment assignment is determined via hard-thresholding, i.e., all main places whose distance is above a known cut-off point, z_0 are treated (i.e., main places in former homeland areas) while main places that are below z_0 are non-treated (i.e., main places in non-former homeland areas):

$$Treat_i = \begin{cases} 1 \text{ if } Dist_i \geq z_0 \\ \\ 0 \text{ if } Dist_i < z_{0.} \end{cases} \tag{3}$$

To identify if a discontinuity exists at the cut-off point, the study plots, as is standard in the literature, the average value of the poverty rate against the running variable (see Imbens and Lemieux's, 2008 for an exposition). The study then employs non-parametric regressions to identify the treatment effect. This is achieved by estimating local linear regressions. In the simplified case when using a rectangular kernel, local linear regressions can be thought of simply estimating a linear regression on both sides of the threshold while restricting the observations to those falling within a certain distance of the threshold using the following model (Imbens and Lemieux, 2008 and Lee and Lemieux, 2010; Jacob and Zhu, 2012):

$$P_i = \alpha + \beta Treat_i + \gamma(z_i - z_0) + u_i \tag{4}$$
$$\text{for } z_0 - h \leq z \leq z_0 + h \text{ with h > 0.}$$

where P_i is as previously defined, *h* denotes the bandwidth with the optimal value is identified by Imbens and Kalyanaraman's (2009) approach. α, β and γ are parameters to be estimated and u_i is the error term. The parameter of interest β captures the effect of former homeland status on poverty.

Tests for RDD validity and corrections

Robustness checks are conducted to assess if covariates, X, (that are likely to have an impact on poverty) are balanced around the cut-off and checks for local continuity. If an imbalance is observed in the covariates, estimates will be invalid. To account for covariate imbalance, Frölich and Huber (2018) propose conditioning on X to identify the treatment effect that will be identified as follows:

$$\lim_{\varepsilon \to 0} E[Y^1 - Y^0 | \, X, D(z_0 + \varepsilon) > D(z_0 - \varepsilon), Z = z_0] \tag{5}$$

where $Y_i^1 - Y_i^0$ is the individual treatment effect and z_0 is the threshold. Given the appealing nature of identifying an unconditional effect for policy, Frölich and Huber (2018) propose an estimator that identifies the unconditional effect:

$$\lim_{\varepsilon \to 0} E[Y^1 - Y^0 | \; D(z_0 + \varepsilon) > D(z_0 - \varepsilon), Z = z_0], \tag{6}$$

by first controlling for X and thereafter average over all X. The estimation of the unconditional effect (equation6) as a more significant estimator is reinforced by at least three reasons. First, it is easier to convey a small number of summary measures to policy makers than it is to do so for many estimated effects at every value of X when focusing on evidence-based policy-making. Second, the estimation of unconditional effects can be done more accurately relative to the conditional effect. Third, if X solely contains pre-treatment variables, the definition of unconditional effects will be independent of the variables included in X. It is also worth noting that different sets of control variables X can also be used to examine the robustness of the results linked to the set of control variables (Frölich and Huber, 2018).

The treatment effect is thus estimated as:

$$\hat{\hat{\beta}} = \frac{\sum\left(\widehat{m}^{+}(X_i,z_0) - \widehat{m}^{-}(X_i,z_0)\right).K_h\left(\frac{Z_i - z_0}{h}\right)}{\sum K_h\left(\frac{Z_i - z_0}{h}\right)} \tag{7}$$

where $K_h(u)$is the kernel function with bandwidth *h*; $\widehat{m}^{+}(X_i, z_0) = \lim_{\varepsilon \to 0} E[Y|X, Z = z + \varepsilon]$ and $\widehat{m}^{-}(X_i, z_0) = \lim_{\varepsilon \to 0} E[Y|X, Z = z - \varepsilon]$. Frölich and Huber's (2018) estimator corrects for the covariate imbalances using a two-step procedure. *First*, estimates of $\widehat{m}^{+}$and $\widehat{m}^{-}$ are estimated by local linear regressions to identify values of intercepts at each side of the cutoff point (see Frölich and Huber's, 2018 for a detailed discussion of the objective function). *Second*, estimates of the treatment effects are obtained in a manner that removes the discontinuity in covariates, if any. This is achieved by either up- or down-weighting observations to the left or right of the cut-off. To illustrate; if a covariate *x* is found to be discontinuous at the threshold with an average value of a_1 to the left and a_2 to the right of the z_0 and if for instance $a_2 > a_1$, then in a symmetric neighbourhood around the threshold, the average value of *x* is given by $\bar{a} = (a_1 + a_2)/2$. The discontinuity in covariate *x* can be removed by re-weighting observations to the left by $a_1/\bar{a}$ and those on the right by $a_2/\bar{a}$. The estimator is implemented in R software.

Conventional confidence intervals typically based on the standard asymptotic distribution of the least squares estimator and robust standard often ignore the asymptotic bias of the non-parametric local polynomial estimator. This simplification can be justified only if the bandwidth in the estimator is selected to be small enough. In addition to including cross-validation and mean-squared error (MSE) minimization (Imbens and Kalyanaraman 2012), the common bandwidth selection methods tend to which generally lead to bandwidth choices that are invalid due to the large conventional confidence intervals used (Calonico, Cattaneo, and Titiunik, 2013; Keele and Titiunik, 2015). To mitigate for this, a process known as undersmoothing is employed. This involves selecting a smaller bandwidth and is used to specify the standard errors. Additionally, the asymptotic bias can be estimated and the standard errors can be corrected (Keele and Titiunik, 2015). Inferences to this approach can be made in two ways. First, by selecting a fixed bandwidth smaller than the MSE bandwidth, the use of conventional confidence intervals can be justified (Keele and Titiunik, 2015). Second, robust confidence intervals with MSE-optimal bandwidth are employed.

4. Data

The proposed study uses data drawn from two sources. The *first* source is the 2011 Census community profiles from Statistics South Africa. These community profiles contain detailed information on administrative boundaries with main place as the lowest geographical unit. Figure A.1 in the Appendix provides a detailed description of South Africa's geography: from the province level to the geo-referenced dwelling frame. Analysis in this study is based on main places which total 14 039 according to the 2011 geography. The community profiles contain information on demographic composition (e.g., age, gender, race and education); household infrastructure (e.g., access to water, electricity and sanitation); labour market outcomes (e.g., employment status), among others.

The *second* data source is cartographic data; the homeland map and a recent map based on the country's 2011 administrative boundaries. Geographic information system (GIS) techniques are used to identify former homeland areas, non-former homeland areas and main places straddling on former homeland boundaries. This information is also used to calculate distances of main places to/from former homeland area boundaries in ArcGIS. As alluded to earlier, distance is calculated in kilometers as the shortest distance from a main place (centroid) to the closest former homeland border. Community profiles and cartographic information are then merged based on a unique main place identifier to create a comprehensive cross-sectional dataset.

Measuring the key variable, poverty headcount, is a daunting task using the community profiles information due to the way the data is captured in Supercross as well as how the income measure is reported. *First*, household incomes are reported in bands. To calculate the poverty headcount measure, the actual point value of household income is required, the study then evaluates each household's annual income at the mid-point of the reported band. This poses a limitation for the analysis since it would be ideal to use actual point value household income. Unfortunately, given the structure of the dataset this is the best that can be done to approximate of household income. *Second*, Supercross allows for the cross tabulation of household income by household size. Using these cross tabulations for all main-places, poor households (i.e., households that have a per-capita income that falls below the upper bound poverty line of R779[3] per person per month) are identified. The number of poor people in a main place is then derived by counting the number of individuals within poor households. Aggregating the poor over households and dividing by the total number of people in the main place yields the poverty headcount measure.

The final analysis is based on data which excludes 371 main places with missing information on some of the key variables. Upon close scrutiny, it was found that these main places were mainly nature/game reserves. In addition, 1 549 main places straddling on the boundaries are dropped

[3] http://www.statssa.gov.za/publications/P03101/P031012018.pdf - page 4. Note in the calculations the poverty line is annualized (multiplied by 12) for consistency with the household incomes reported in the Census data that are annual values.

as they are neither inside nor outside – hence difficult to classify for the purposes of this study. As a result, the study is based on 12 119 main places. Table A.1 in the appendix provides a description of the study's key variables while Table 1 presents corresponding descriptive statistics for the final sample used.

Descriptive statistics indicate an average poverty rate of 70.2 percent in 2011. Admittedly, this is 24.7 percentage points higher than the 45.5 percent reported by StatsSA (StatsSA 2011). This discrepancy is potentially due to the use of mid-point rather than the actual point value of household income. Thus in the approximation, households income may be overestimated for households who have an income lower than the calculated mid-point and vice versa. As such, this over-estimation of poverty might be due to the under-estimation of incomes for households at the top-end of the reported income bands. Consequently, the analysis can be deemed as identifying the upper bound effects which to some is still informative as a first step.

Table 1: Descriptive statistics

	Overall		Inside 100%		Outside 100%	
Variable	Mean	SD.	Mean	SD.	Mean	SD.
Dependent variable						
Poverty headcount (%)	70.2	(17.5)	74.9	(11.4)	48.3	(23.5)
Independent variables						
% Aged 0 – 34	58.0	(8.4)	59.1	(7.5)	53.0	(10.5)
% Aged 35 – 49	22.9	(5.4)	21.9	(4.7)	27.9	(6.0)
% Aged 50 – 64	19.0	(6.6)	19.0	(6.2)	19.0	(8.2)
Black	93.7	(20.2)	99.6	(2.1)	65.4	(37.3)
Female	53.3	(12.5)	56.3	(9.7)	38.6	(13.9)
No schooling	18.5	(14.8)	25.7	(15.1)	11.3	(10.9)
Primary school	28.2	(12.8)	32.9	(12.6)	23.5	(11.2)
Secondary school	31.9	(10.2)	31.9	(10.3)	31.9	(9.7)
Matric	16.0	(9.9)	11.7	(8.9)	20.3	(10.6)
Degree /Certificate	5.4	(5.2)	3.4	(4.0)	7.5	(8.0)
Married	23.4	(8.7)	21.2	(5.3)	33.5	(13.5)
Unemployed	32.0	(22.1)	33.8	(22.8)	22.9	(16.0)
Urban area (dummy)	0.1	(0.3)	0.01	(0.1)	0.7	(0.5)
Electricity (Lighting)	64.4	(0.4)	61.4	(40.8)	78.8	(29.2)
Piped Water	64.2	(0.2)	58.9	(43.2)	89.7	(22.8)
N	12 119		10 038		2 079	

Statistics show higher poverty levels for former (74.9 percent) relative to non-former homeland areas (48 percent) – this yields a 27 percent point difference in poverty rates. These results are consistent with the notion of underdevelopment in former homeland areas owing to decades of repression as well as limited economic progress post-apartheid (Burger et. al., 2017; David et. al., 2018; Noble and Wright, 2013). This finding is also consistent with Shifa (2018) who found that the most disadvantaged and poor areas in the country are situated in the former homeland areas.

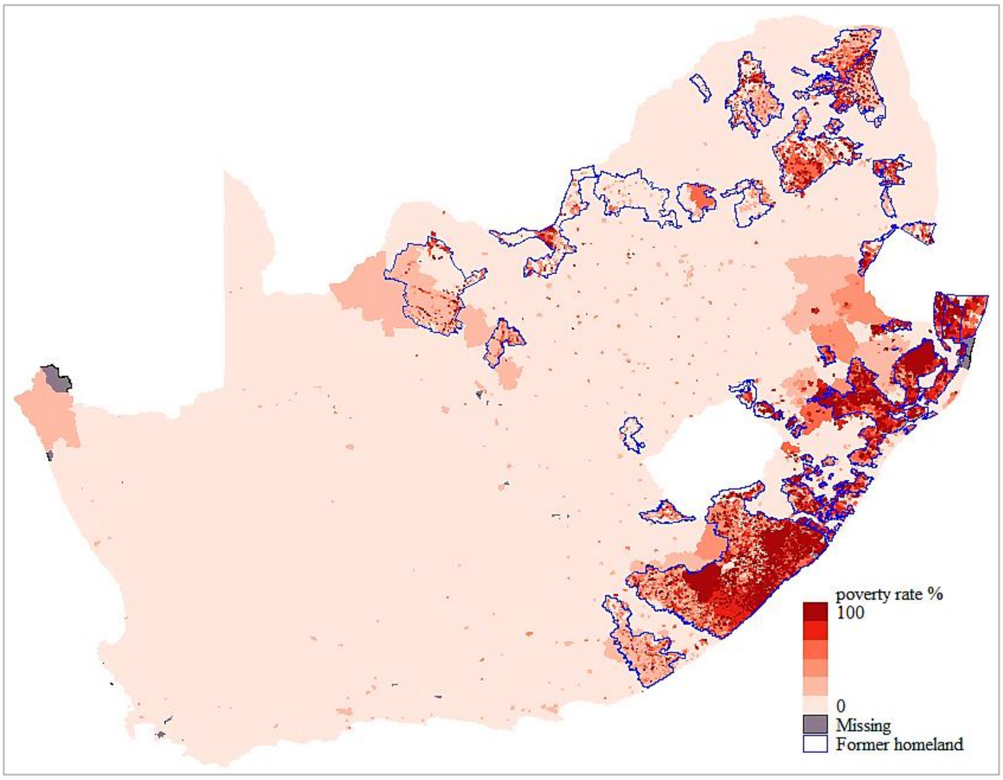
Turning to the distribution of other variables: it is observed that former homeland areas have a relatively high share of Black South Africans (99 percent) compared to the 65 percent for non-former homeland areas. The considerable share of Black South Africans in non-former homeland areas is consistent with migration patterns post-apartheid. Similarly, the distribution of gender across former (56.3 percent) and non-former homeland areas (38.6 percent) is also partly reflective of migration flows that were skewed in favour of men. These migrations potentially lead to damaged family relationships. As such, it is observed that former homeland areas have a lower share of married individuals compared to non-former homeland areas.

Pertaining to unemployment (as defined in Table A.1 in the appendix), former homeland areas exhibit higher levels of unemployment (33.8 percent) compared to non-former homeland areas (22.9 percent). This resonates well with the extant literature and the well-known fact of limited economic opportunities within former homeland areas (Turok, 1994; Terreblanche, 2002). The distribution of education is largely in favour of non-former homeland areas; it is observed that former (non-former) homeland areas have a larger share of individuals with no schooling and primary education (education above matric). Again, this is in line with disparities uncovered here that exists in terms of access to amenities; previous studies (e.g., Leibbrandt, Woolard and Woolard 2000; Makgetla, 2010; Turok, 2012) also point this out. Finally, it is notable that there are considerable differences in shares of urban main places across former and non-former homeland areas.

In order to get a clear picture of the spatial distribution of poverty across former and non-former homeland areas, Figure 2 maps the poverty headcount measure. The map clearly shows that poverty is concentrated in former homeland areas. These results are consistent with Table 1.

Although poverty is generally lower in non-former homeland areas spatial mapping shows noticeable pockets of high poverty outside former homeland areas.

Figure 2: Spatial distribution of poverty in South Africa, 2011



The map also shows that there are main places that are further removed from former homeland areas (e.g. Western Cape). To obtain a more accurate sense of differences in poverty between former and non-former homelands, Table 2 presents poverty headcount measures for main places truncated by distance from the former homeland boundary. This ensures a comparison of main places within the same vicinity to minimize confounded results arising from comparing main places that are too different.

Table 2: Poverty headcount rates (%)

	Inside			Outside			
	Mean	S.D.	N	Mean	S.D.	N	Raw-poverty rate diff.
Overall	74.9	(11.3)	10 038	48.3	(23.3)	2 079	26.6
Distance ± 5km	73.3	(11.7)	1 742	60.8	(22.9)	324	12.5
Distance ± 10km	73.8	(11.4)	3 920	58.9	(24.5)	556	14.8
Distance ± 15km	74.1	(11.4)	5 461	57.4	(25.1)	696	16.7

Distance ± 20km	74.0	(11.4)	6 405	57.4	(24.7)	798	16.6

Footnote: Distance is calculated as the shortest distance between a main place (centroid) and closest former homeland border (Overall table in Appendix, labelled Table A.2).

Table 2 shows that the raw poverty rate difference increases as distance increases. Within the 5km radius, the poverty differential is 12.5 percentage points in favour of non-former homeland areas. This differential increases to 14.8 and 16.6 percentage points within a 10 km and 20km radius. These results suggest that the standard of living along with access to opportunities increases the further out main places are from former homeland boundaries. Also noteworthy is the high variation in poverty for former homeland areas compared non-former homeland areas. The relatively low variation in former homeland areas is potentially reflective of exposure to a common repression mechanism. While it may be true that former homeland areas experience excessively high poverty levels due to this mechanism, non-former homeland areas were also vulnerable to its effects.

5. Empirical Results

This section presents results from the OLS as well as RDD estimations. Table 3 presents OLS estimates of λ as specified in equation 2 for five specifications using various distance restrictions. Table A.1 in the Appendix presents the full regression results. The preferred specification is IV as it controls for a full set of observable characteristics available in the data. Table 4 presents a test for covariate balance around the threshold while Table 5 presents the RDD estimates.

5.1 *Results from OLS model*

Results show that former homeland areas experience higher levels of poverty compared to non-former homeland areas. These associations are statistically significant at the 1 percent level across all four specifications. The raw poverty rate difference (specification I) indicate that there is a 26.6 percentage point difference in poverty rates. This decreases to 6.3 percentage points (specification II) when demographics, education and employment are controlled for. These results suggest that a substantial amount (76 percent) of the observed raw poverty differential is explained by differences in demographic composition of former/non-former homeland areas. Further adding controls of urban area, infrastructure and districts reduces the poverty rate difference by 1.4 percentage points. This leaves a 4.9 (specification IV) percentage point poverty differential which is still statistically significant at the 1 percent level.

When the estimations are restricted by distance around the threshold (i.e., 5 km, 10 km, 15 km and 20 km radius), results still point to a poverty rate difference that is statistically significant, albeit with varying magnitudes. An interesting pattern emerges; as radius distance increases, the difference in poverty rates also increases. This is consistent across all four specifications. Evaluations using the 5km radius uncover a poverty differential in the range 2-12.5 percentage points depending on specification while assessments based on a 20km radius yields a poverty differential in the range 4-17 percentage points.

It is important to note that observed characteristics explain a substantial amount of the variation – based on the specification that accounts for a full set of covariates (specification IV), the study uncovers relatively smaller (2-4 percentage points). Although now smaller, the effect remains statistically significant and is in favour of non-former homeland areas.

Turning to results for the other covariates (see Table A.3 in the Appendix), statistically significant results at the 1 percent level are observed for all variables with the exception of the electricity variable. Results show that an increase in the share of older individuals will reduce poverty levels compared to having a younger population. This is consistent with an increase with an increase in earnings across the life cycle. In line with previous studies, there is a positive correlation between the share of blacks and females with relation to poverty. This result is consistent across specifications. Pertaining to the education variable, poverty tends to be higher for lower levels of education relative to having a certificate/degree. Interestingly, the positive correlation weakens as education increases. Marital status is found to be negatively correlated to poverty – a result that extends to access to amenities. Finally, urban areas also exhibit low levels poverty. Results are generally consistent across all specifications for the distance truncation.

Table 3: OLS results by distance to/from former homeland border

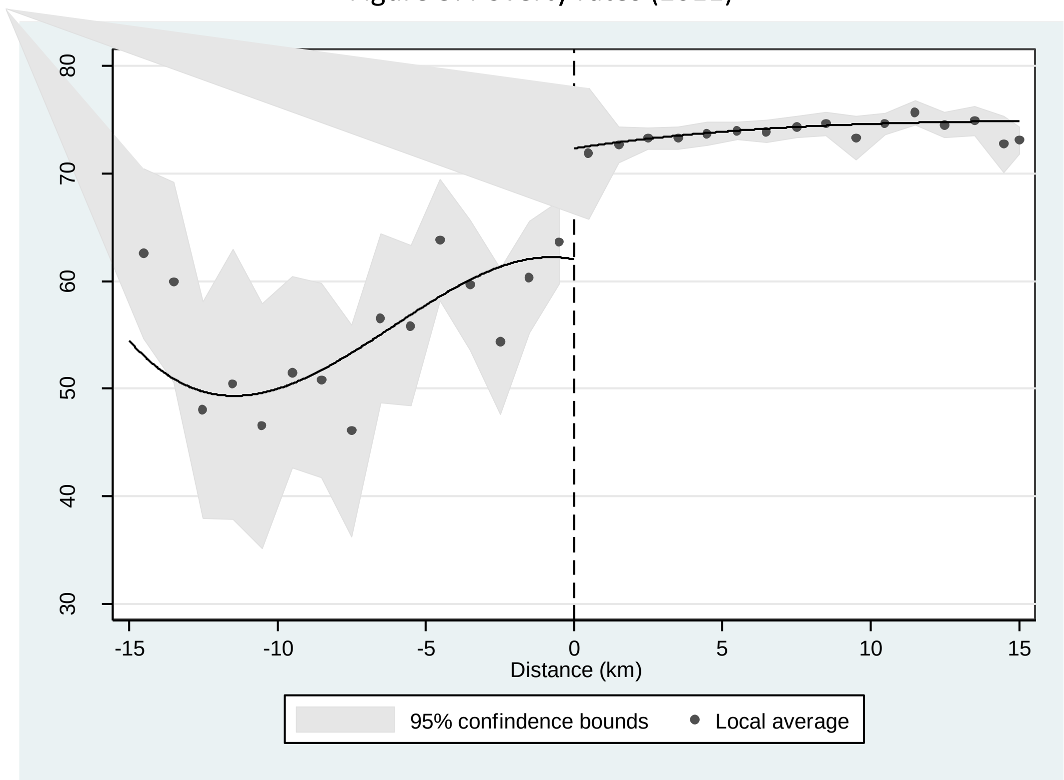
	I		II		III		IV	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
No distance restriction								
Former homeland (dummy)	26.65***	(2.34)	6.29***	(1.24)	6.06***	(0.98)	4.87***	(1.02)
Distance restriction from/to former homeland boundary								
Distance ± 5 km	12.5***	(2.51)	3.14***	(0.90)	2.79***	(0.79)	2.01**	(0.38)
Distance ± 10 km	14.8***	(0.62)	3.87***	(0.88)	3.79***	(0.88)	4.03***	(0.95)
Distance ± 15 km	16.7***	(0.55)	4.14***	(1.01)	4.39***	(0.88)	3.48***	(0.89)
Distance ± 20 km	16.6***	(0.51)	4.89***	(1.17)	5.03***	(0.97)	4.03***	(0.95)
Controls:								
Demographics	No		Yes		Yes		Yes	
Unemployment	No		Yes		Yes		Yes	
Education	No		Yes		Yes		Yes	
Infrastructure	No		No		Yes		Yes	
Urban	No		No		No		Yes	
District dummy	No		No		Yes		Yes	

5.2. Results from RD design estimates

Graphical results

Imbens and Lemieux (2008) argue that a graphical representation of the discontinuity should be a fundamental part of any RD design analysis. The authors further note that graphical illustrations are the most powerful methods to employ when assessing the credibility of any RDD strategy. Figure 3 plots average poverty rates by distance along with 95 percent confidence intervals. Results indicate that poverty levels tend to increase as the distance from former homeland border becomes smaller (from left hand side). There is a distinct sharp increase in poverty at the cut-off point, 0km, where main places become part of former homeland areas (right hand side). Poverty rates for former homeland areas (right hand side) remain consistently higher with little variation as distance from former homeland boundary increase. Interestingly, there is considerable variation in poverty rates for non-former homeland areas (left hand side) while relatively lower variation is exhibited in former homeland areas (right hand side); yet the sharp jump in poverty rates at the cut-off point is confirmed. By accounting for a 15km distance for the both former and non-former homelands, the figure aims to include a range of distances as is standard with an RDD graph (Jacob and Zhu, 2012). A closer bandwidth will be evaluated in the estimations.

Figure 3: Poverty rates (2011)



For the RDD approach to be valid, covariates (that are potentially correlated with poverty rates) should be balanced at the threshold. Covariate balance tests are conducted by running RDD estimations with the covariates as outcome variables using the 'rdd' package. Table 4 presents estimates and corresponding *p*-values as well as the sample mean. If covariates are properly balanced, the differences between former and non-former homeland areas at the threshold should be statistically insignificant. Results show statistically significant differences for the share of blacks, female, secondary school, matric, unemployed, electricity and piped water at the usual levels. There are more blacks and females in former homeland areas than non-former homeland areas. This is expected given that homelands were designated for black South Africans and out-migration patterns were skewed in favour of males. The differences in education i.e., more high school dropouts and fewer matriculants in former homeland areas is reflective of limited education opportunities in former homeland areas. There are also imbalances in access to amenities and share of main places classified as urban. The results for electricity and piped water access are difficult to explain. Admittedly, due to covariate imbalance RDD estimates of RDD are likely to be invalid. To account for the imbalance in covariates, this study applies a recent estimator proposed by Frölich and Huber (2018) as discussed in the methodology.

Table 4: Covariate sample means and balance tests at the threshold

Variable	Overall Sample mean	Diff.	P-value
% Aged 35 - 49	22.93	0.54	0.71
% Aged 50 - 64	19.00	0.71	0.11
Black	93.72	6.84	0.00
Female	53.28	5.41	0.00
Married	23.36	-2.01	0.24
No schooling	18.50	-1.12	0.55
Primary school	28.18	3.21	0.10
Secondary school	31.89	2.57	0.03
Matric	15.97	-3.66	0.01
Degree/certificate	5.40	-0.93	0.28
Unemployed	31.95	7.10	0.00
Electricity (Lighting)	64.37	21.06	0.00
Piped Water	64.18	6.12	0.09
Urban area (dummy)	0.06	0.29	0.00

Note: Optimal bandwidth is obtained using the Imbens-Kalyanaraman (2012) approach. The *p*-values used are based on standard errors bootstrapped 99 times.

Table 5 presents results from the standard RDD $\hat{\beta}$ and estimates from Frölich and Huber's (2018) estimator $\hat{\hat{\beta}}$. OLS results, $\hat{\lambda}$, are also presented in Table 5 for comparison. When covariates are ignored, standard RDD estimates show a poverty rate difference of 5 percentage points which is statistically significant at the 5 percent level for the optimal bandwidth of 5km. When distance is further reduced (increased), the difference increases to 7 percentage points (9 percentage points). Comparing OLS and the standard RDD estimates, OLS tends to over-estimate the difference.

Table 5: Effects of former homeland status on poverty

	OLS		RDD					
	without X	with X	without X			with X		
Bandwidth *h*	5km	5km	2.5km	5km	10km	2.5km	5km	10km
Treatment effect	12.5	2.01	6.74	5.12	8.8	2.47	2.12	2.58
S.E.	(1.33)	(0.38)	(2.48)	(2.13)	(1.76)	(0.93)	(0.95)	(0.90)
p-value	0.00	0.00	0.01	0.02	0.00	0.01	0.03	0.00

Notes: The bandwidths are picked using the Imbens-Kalyanaraman (2012) approach. The table presents several values half and double the identified optimal bandwidth value of 5km. Standard errors in parentheses are based on bootstrapping the estimate 99 times. X includes variables indicated in Table 4.

Turning to estimates that control for covariates, results indicate that former homeland areas experience high poverty levels compared to non-former homeland areas. The poverty rate difference of 2.1 percentage points is uncovered for the 5km radius. This is slightly higher than what is identified under OLS. Varying the distance around the cutoff points yields a 2.5 percentage point difference for poverty rates for the 2.5km radius which is of the same order as those obtained for the 10km radius. Thus, by using half of the optimal distance (2.5km), the model accounts for undersmoothing. Results obtained are still robust as with the optimal bandwidth. Differences in RD design estimates without covariates compared to those that include covariates emphasize the importance of differences in observed characteristics of former and non-former homeland areas in explaining the poverty rate difference. Notably, these characteristics do not fully explain the poverty rate difference suggesting the existence of some form of 'scarring effect' of the former homeland policy.

Discussion of results

Results of the study are largely consistent with the extant literature which examines poverty by location. Similarly, provinces or districts incorporating previously marginalised areas are characterised by relatively high levels of poverty (Adato et. al., 2006; Bhorat and Kanbur, 2006; Carter and May 1999; Noble and Wright, 2013; Leibbrandt, Woolard and Woolard, 2000; Woolard, 2012; Yamauchi, 2005; Hoogeveen and Özler, 2006). Unlike previous studies, the current analysis provides more insights on the poverty differential. Two important inferences can be drawn from these results. *First*, there are enormous differences in poverty levels between former and non-former homeland areas to the disadvantage of former homeland areas. However, raw differences are to a large extent underpinned by differences in observed characteristics or composition of former homeland areas. *Second*, there is evidence of 'scarring effects' of the former homeland policy. Using the descriptive statistics, the 2 percentage point difference translates to a 50 percent poverty rate for former homeland areas compared to 48 percent in non-former homeland areas. Thus, the role of former homeland policy in explaining poverty – scarring effect – can be considered to be considerable in terms of human welfare.

These results have important implications for policy. The existence of a 'scarring effect' implies that poverty differentials between former and non-former homeland areas may persist in the long run as they are partly structural. Fortunately, since a large share of the observed "raw" poverty differential is due to differences in observed characteristics (e.g. education, demographic composition and access to amenities) of former and non-former homeland areas, this leaves room for policy to redress this imbalance. If the differences in the distribution of aforementioned characteristics are narrowed, poverty differentials will become significantly smaller or converge to the 'scarring effect' values. This signifies the important role for policy aimed at improving observed characteristics of former homeland area residents. Government interventions can encompass improvement of human capital; employment prospects and support for women as they are disproportionately represented in former homeland areas.

5. Conclusion

The study sets out to investigate the effect of former homeland status on poverty in South Africa. This is achieved by employing OLS as well as RDD techniques using the 2011 Census community profiles from Statistics South Africa and cartographic data. At face value, substantial poverty rate differences are uncovered of 26.6 percentage points in favour of non-former homeland areas. However, when differences in observed characteristics (i.e., demographic composition, human capital, unemployment and access to basic amenities) are accounted for, the difference in poverty rates almost disappears leaving a 2 percentage point differential which is statistically significant. The remaining difference after netting out potential confounding factors is referred to as the 'scarring effect' of the former homeland policy on poverty.

Results of this study clearly highlight the importance of differences observed characteristics in explaining observed poverty patterns in South Africa. Accordingly, observed spatial disparities in poverty can be mitigated by government initiatives that seek to narrow differences in observable characteristics to the benefit of previously marginalised communities. While investing in interventions that improve observed characteristics of former homeland areas can go a long way in narrowing the poverty rate difference, it is unfortunate that part of the difference, although relatively small, is structural. This means that poverty rate differences across former and non-former homeland areas may persist to the extent to which poverty is underpinned by the homeland policy 'scarring effect' and the rate at which the 'scars' fades over-time.

While informative, this study has several limitations. *First*, poverty head count is measured using household income. Extant research argues that expenditure is a more reliable welfare indicator than income, especially in developing countries (Rao, 2006). Despite the debate regarding which approach to use when measuring poverty, the results obtained in the study give a first outlook on the prevalence of poverty in former homeland areas than previously possible. *Second*, the household income measure used is reported in bands and hence the use of mid-point values is inevitable for the analysis. This calculation may lead to the over (under) estimation of certain observations whose income level may have higher (lower) estimates than reported. Solving for this limitation would require datasets that collect detailed information on income at a large scale as the Census. *Third*, the study looks at poverty as a one-dimensional phenomenon. In the South African context, there is a case for analysing poverty in a multidimensional framework. Future studies can benefit from accounting for these issues as comprehensive datasets become available.

References

Abel, R., 2015. *Politics by other means: Law in the struggle against apartheid, 1980-1994*. Routledge.

Acemoglu, D., Johnson, S. and Robinson, J.A., 2001. The colonial origins of comparative development: An empirical investigation. *American economic review*, *91*(5), pp.1369-1401.

Adato, M., Carter, M.R. and May, J., 2006. Exploring poverty traps and social exclusion in South Africa using qualitative and quantitative data. *The Journal of Development Studies*, *42*(2), pp.226-247.

Alexander, M., 2018. South Africa Gateway: The nine provinces of South Africa *retrieved from* https://southafrica-info.com/land/nine-provinces-south-africa.

Aliber, M., 2003. Chronic poverty in South Africa: Incidence, causes and policies. *World Development*, *31*(3), pp.473-490.

Banerjee, A., S, Galiani, J, Levinsohn, Z, McLaren, I, Woolard (2008), "Why has Unemployment Risen in the New South Africa?" Economics of Transition, 16(4): pp. 715-740. Oxford: Blackwell Publishing.

Bastos, P. and Bottan, N., 2014. Overcoming the tyranny of history: Evidence from post- apartheid South Africa. *World Bank. Paper retrieved from the following: http://beta.udep.edu.pe/empresas/files/2014/07/3E_1_BastosBOTTAN_2.pdf*.

Bhorat, H. and Kanbur, S.R. eds., 2006. *Poverty and policy in post-apartheid South Africa*. HSRC press.

Bhorat, H. and van Der Westhuizen, C., 2013. Temporary Employment Services in South Africa: A Brief Note. *Washington, DC: The World Bank Group, World Development Report.*

Burger, R., Van der Berg, S., van der Walt, S. and Yu, D., 2017. The long walk: Considering the enduring spatial and racial dimensions of deprivation two decades after the fall of apartheid. *Social Indicators Research*, *130*(3), pp.1101-1123.

Carter, M.R. and May, J., 1999. Poverty, livelihood and class in rural South Africa. *World development*, *27*(1), pp.1-20.

Cutler, P., 1984. The measurement of poverty: A review of attempts to quantify the poor, with special reference to India. *World Development*, *12*(11-12), pp.1119-1130.

David, A., Guilbert, N., Hamaguchi, N., Higashi, Y., Hino, H., Leibbrandt, M. and Shifa, M., 2018. Spatial poverty and inequality in South Africa: A municipality level analysis.

Davids, Y.D., 2010. *Explaining poverty: a comparison between perceptions and conditions of poverty in South Africa* (Doctoral dissertation, Stellenbosch: University of Stellenbosch).

de Kadt, D. and Larreguy Arbesu, H., 2014. Agents of the Regime? Traditional leaders and electoral behavior in South Africa.

Dell, M., 2010. The persistent effects of Peru's mining mita. *Econometrica*, *78*(6), pp.1863-1903.

Foster, J., Greer, J. and Thorbecke, E., 1984. A class of decomposable poverty measures. *Econometrica: journal of the econometric society*, pp.761-766.

Frölich, M. and Huber, M., 2018. Including covariates in the regression discontinuity design. *Journal of Business and Economic Statistics*, pp.1-13.

Gwanya, T.T., 2010. South Africa position paper on rural development: A model for the Comprehensive Rural Development Programme. *Department of Rural Development and Land Reform*.

Hagenaars, A.J. and Van Praag, B.M., 1985. A synthesis of poverty line definitions. *Review of Income and Wealth*, *31*(2), pp.139-154.

Hanmer, L.C., Pyatt, G. and White, H., 1999. What do the World Bank's Poverty Assessments teach us about Poverty in Sub-Saharan Africa?. *Development and change*, *30*(4), pp.795-823.

Harris, B., Goudge, J., Ataguba, J.E., McIntyre, D., Nxumalo, N., Jikwana, S. and Chersich, M., 2011. Inequities in access to health care in South Africa. *Journal of public health policy*, *32*(1), pp.S102-S123.

Hoogeveen, J.G. and Özler, B., 2006. Poverty and inequality in post-apartheid South Africa: 1995-2000. *Poverty and policy in post-apartheid South Africa*, pp.59-94.

Hopkins, K., 2001. Assessing the World's Response to Apartheid: A Historical Account of International Law and Its Part in the South African Transformation. *Miami Int'l and Comp. L. Rev.*, *10*, p.241.

Imbens, G. and Kalyanaraman, K., 2009. Optimal bandwidth choice for the regression

Imbens, G. and Kalyanaraman, K., 2012. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, *79*(3), pp.933-959.

Imbens, G.W. and Lemieux, T., 2008. Regression discontinuity designs: A guide to practice. *Journal of econometrics*, *142*(2), pp.615-635.

Jacob, R., Zhu, P., Somers, M.A. and Bloom, H., 2012. A Practical Guide to Regression Discontinuity. *MDRC*.

Kanbur, S.R., 1987. Measurement and alleviation of poverty: With an application to the effects of macroeconomic adjustment. *Staff Papers*, *34*(1), pp.60-85.

Keele, L.J. and Titiunik, R., 2015. Geographic boundaries as regression discontinuities. *Political Analysis*, *23*(1), pp.127-155.

Lahiff, E. and Cousins, B., 2005. Smallholder agriculture and land reform in South Africa.

Lee, E.S., 1966. A theory of migration. *Demography*, *3*(1), pp.47-57.

Lee, D.S., 2008. Randomized experiments from non-random selection in US House elections. *Journal of Econometrics*, *142*(2), pp.675-697.

Lee, D.S. and Lemieux, T., 2010. Regression discontinuity designs in economics. *Journal of economic literature*, *48*(2), pp.281-355.

Leibbrandt, M., Woolard, C. and Woolard, I., 2000. The contribution of income components to income inequality in the rural former homelands of South Africa: a decomposable Gini analysis. *Journal of African Economies*, *9*(1), pp.79-99

Leibbrandt, M., Finn, A. and Woolard, I., 2012. Describing and decomposing post-apartheid income inequality in South Africa. *Development Southern Africa*, *29*(1), pp.19-34.

Lok-Dessallien, R., 2002. Review of Poverty Concepts and Indicators, United Nations Development Programme publications. *Retrieved March*, *2*, p.2005.

Makgetla, N., 2010. Rural development: Synthesis paper. *Development Bank of Southern Africa, Midrand*.

Mariotti, M. and Fourie, J., 2014. The economics of apartheid: An introduction.

Mlambo, V., 2018. An overview of rural-urban migration in South Africa: its causes and implications. *Archives of Business Research*, *6*(4).

National Development Plan 2030 – South African Government, 12 May 2012; *https://www.poa.gov.za/news/Documents/NPC%20National%20Development%20Plan%20 Vision%202030%20-lo-res.pdf*

Naudé, W., McGillivray, M. and Rossouw, S., 2009. Measuring the vulnerability of subnational regions in South Africa. *Oxford Development Studies*, *37*(3), pp.249- 276.

Noble, M. and Wright, G., 2013. Using indicators of multiple deprivation to demonstrate the spatial legacy of apartheid in South Africa. *Social Indicators Research*, *112*(1), pp.187 201.

Noble, M., Ratcliffe, A. and Wright, G., 2004. Conceptualizing, defining and measuring poverty in South Africa: An argument for a consensual approach. *SARPN.* www.sarpn.org.za.

Ntshidi, A.T., 2017. *Patterns of Rural-Urban migration in South Africa* (Doctoral dissertation, North-West University (South Africa), Mafikeng Campus).

Nunn, N., 2008. The long-term effects of Africa's slave trades. *The Quarterly Journal of Economics*, *123*(1), pp.139-176.

Özler, B. and Hoogeveen, J.G.M., 2005. Not separate, not equal: Poverty and inequality in post-apartheid South Africa.

Pienaar, L. and Von Fintel, Dieter, 2013. *Hunger in the former apartheid homelands: Determinants of converging food security 100 years after the 1913 Land Act* (No. 26/2013).

Porter, G. and Phillips-Howard, K., 1997. Agricultural issues in the former homelands of South Africa: The Transkei. *Review of African Political Economy*, *24*(72), pp.185-202.

Rao, D.P., 2006. Poverty measurement using expenditure approach. *School of Economics, University of Queensland, Brisbane, Australia*.

Ravallion, M., Datt, G. and Van De Walle, D., 1991. Quantifying absolute poverty in the developing world. *Review of Income and Wealth*, *37*(4), pp.345-361.

Rowntree, B.S., 1902. Poverty: A Study of Town Life Macmillan: London.

Sen, A.K., 1997. From income inequality to economic inequality. *Southern Economic Journal*, *64*(2), pp.384-401.

Terreblanche, S.J., 2002. *A history of inequality in South Africa, 1652-2002*. University of Kwazulu-Natal Press.

Townsend, P., 1979. *Poverty in the United Kingdom: a survey of household resources and standards of living*. Univ of California Press.

Turok, I. 1994. Urban planning in the transition from Apartheid: The legacy of social control. *Town Planning Review* 65, pp. 243–259.

Turok, I., 2012. Urbanisation and Development in South Africa: Economic Imperatives, Spatial Distortions and Strategic Responses. *International Institute for Environment and Development United Nations Population Fund.* Urbanisation and Emerging Population Series Working Paper No. 8

Van Praag, B.M., Hagenaars, A.J. and Van Weerden, H. 1982. "Poverty in Europe." *Review of Income and Wealth***,** 28(3): 345 – 359.

Von Fintel, D.P., 2018. Long-run spatial inequality in South Africa: early settlement patterns and separate development. *Studies in Economics and Econometrics*, *42*(2), pp.81-102.

Woolard, I., 2002. An overview of poverty and inequality in South Africa. *Unpublished briefing paper, HSRC, Pretoria*.

Yamauchi, F., 2005. Race, equity, and public schools in post-Apartheid South Africa: Equal opportunity for all kids. *Economics of Education Review*, *24*(2), pp.213-233

Appendix

Figure A.1: Nested hierarchy for the South African Census

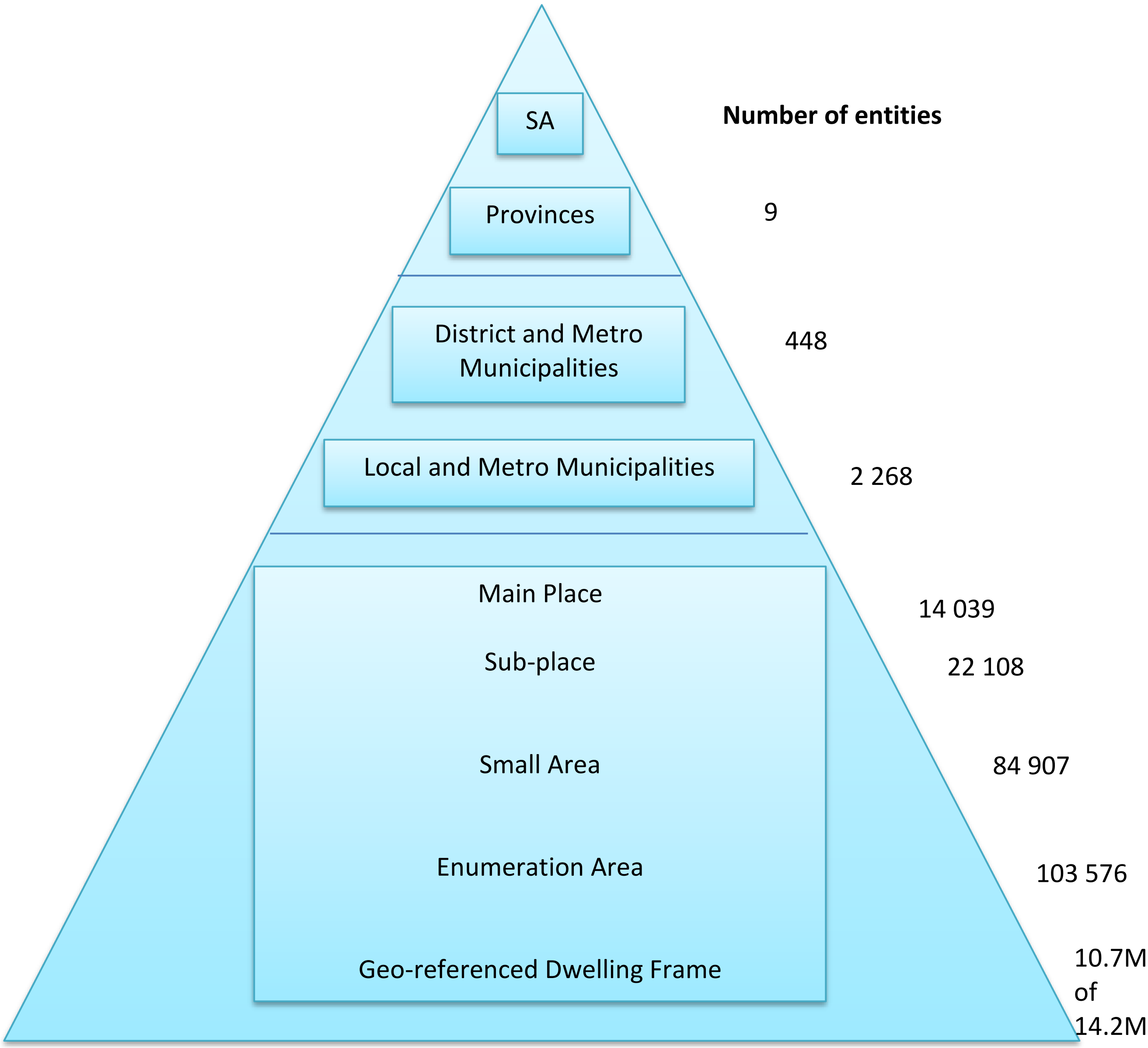


Table A.1: Description of Variables

Variable	Description
Inside	Main places located strictly inside former homeland areas (referred to as former homeland areas)
Poverty headcount (%)	Percentage of people living below the upper bound poverty line
Distance	Distance of main places to/from former homeland area boundaries calculated in kilometres
Aged 0 - 34	Percentage of people aged 0 - 34 years
Aged 35 - 49	Percentage of people aged 35 - 49 years
Aged 50 - 64	Percentage of people aged 50 - 64 years
Aged 65 and older	Percentage of people aged 65 years and older
Females	Percentage of female
Black	Percentage of Black
No schooling	Percentage of people aged 20+ who have no schooling
Primary school	Percentage of people aged 20+ who have a primary school schooling
Secondary school	Percentage of people aged 20+ who have a secondary school schooling
Matric certificate	Percentage of people aged 20+ who have a matric certificate
Certificate/ Degree	Percentage of people aged 20+ who have a certificate/degree
Married	Percentage of people who are married
Unemployed	Unemployment rate percentage of the adult population
Electricity	Percentage of people who can access electricity
Piped water	Percentage of people who can piped water
Urban	Main place classified as urban (dummy)

Table A.2: Effect of former homeland status on poverty - OLS estimates

	Overall							
	I		II		III		IV	
Variable	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Inside	**26.65**	**(2.34)**	**6.29**	**(1.24)**	**6.06**	**(0.98)**	**4.87**	**(1.02)**
% Aged 35 - 49			-0.42	(0.04)	-0.33	(0.04)	-0.33	(0.04)
% Aged 50 - 64			-0.35	(0.05)	-0.25	(0.05)	-0.26	(0.05)
Black			0.07	(0.02)	0.11	(0.02)	0.10	(0.02)
Female			0.24	(0.04)	0.19	(0.04)	0.19	(0.04)
No schooling			0.73	(0.05)	0.72	(0.05)	0.70	(0.05)
Primary school			0.71	(0.06)	0.71	(0.06)	0.69	(0.05)
Secondary school			0.62	(0.06)	0.63	(0.06)	0.62	(0.06)
Matric			0.40	(0.07)	0.40	(0.07)	0.40	(0.06)
Married			-0.23	(0.05)	-0.24	(0.05)	-0.24	(0.05)
Unemployed			0.05	(0.01)	0.05	(0.01)	0.05	(0.01)
Electricity					-0.002	(0.01)	-0.002	(0.005)
Piped water					-0.02	(0.004)	-0.02	(0.004)
Urban area							-3.24	(0.87)
District fixed effects					Yes		Yes	

Table A.3 : OLS results by distance to/from former homeland border

	5km radius							
	I		II		III		IV	
Variable	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Inside	**12.5**	**(2.51)**	**3.14**	**(0.90)**	**2.79**	**(0.79)**	**2.01**	**(0.83)**
% Aged 35 - 49			-0.42	(0.06)	-0.32	(0.04)	-0.31	(0.04)
% Aged 50 - 64			-0.43	(0.06)	-0.33	(0.07)	-0.35	(0.07)
Black			0.09	(0.07)	0.08	(0.07)	0.06	(0.07)
Female			0.27	(0.05)	0.19	(0.04)	0.19	(0.04)
No schooling			0.80	(0.07)	0.83	(0.06)	0.80	(0.05)
Primary school			0.73	(0.07)	0.78	(0.05)	0.75	(0.05)
Secondary school			0.70	(0.09)	0.75	(0.08)	0.72	(0.08)
Matric			0.44	(0.10)	0.49	(0.07)	0.48	(0.07)
Married			-0.24	(0.10)	-0.35	(0.08)	-0.33	(0.08)
Unemployed			0.05	(0.01)	0.06	(0.01)	0.06	(0.01)
Electricity					-0.01	(0.01)	-0.01	(0.01)
Piped water					-0.02	(0.01)	-0.02	(0.01)
Urban area+							-3.82	(1.37)
District fixed effects					Yes		Yes	

Table A.3 : OLS results by distance to/from former homeland border - continued

	10km radius							
	I		II		III		IV	
Variable	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Inside	**14.8**	**(0.62)**	**3.87**	**(0.88)**	**3.97**	**(0.74)**	**4,03**	**(0,95)**
% Aged 35 - 49			-0.48	(0.07)	-0.40	(0.08)	-0.39	(0.07)
% Aged 50 - 64			-0.48	(0.05)	-0.38	(0.07)	0.09	(0.05)
Black			0.11	(0.05)	0.11	(0.05)	0.19	(0.04)
Female			0.24	(0.05)	0.19	(0.04)	0.81	(0.05)
No schooling			0.83	(0.06)	0.83	(0.05)	0.76	(0.04)
Primary school			0.77	(0.06)	0.79	(0.04)	0.72	(0.07)
Secondary school			0.72	(0.07)	0.74	(0.07)	0.51	(0.06)
Matric			0.50	(0.08)	0.52	(0.06)	-0.26	(0.05)
Married			-0.20	(0.06)	-0.27	(0.06)	0.05	(0.01)
Unemployed			0.04	(0.01)	0.05	(0.01)	-0.01	(0.01)
Electricity					-0.006	(0.01)	-0.017	(0.01)
Piped water					-0.02	(0.01)	-3.90	(1.25)
Urban area+							0.00	(0.00)
District fixed effects					Yes		Yes	

Table A.3 : OLS results by distance to/from former homeland border - continued

	15km radius							
	I		II		III		IV	
Variable	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Inside	**16.7**	**(0.55)**	**4.14**	**(1.01)**	**4.39**	**(0.88)**	**3.48**	**(0.89)**
% Aged 35 - 49			-0.45	(0.06)	-0.36	(0.07)	-0.36	(0.07)
% Aged 50 - 64			-0.46	(0.04)	-0.34	(0.05)	-0.36	(0.05)
Black			0.13	(0.05)	0.12	(0.05)	0.10	(0.05)
Female			0.23	(0.04)	0.17	(0.04)	0.17	(0.04)
No schooling			0.82	(0.06)	0.83	(0.05)	0.81	(0.05)
Primary school			0.77	(0.06)	0.79	(0.04)	0.77	(0.04)
Secondary school			0.71	(0.07)	0.74	(0.06)	0.71	(0.06)
Matric			0.50	(0.08)	0.52	(0.06)	0.50	(0.06)
Married			-0.20	(0.06)	-0.27	(0.05)	-0.26	(0.05)
Unemployed			0.04	(0.01)	0.05	(0.01)	0.05	(0.01)
Electricity					-0.004	(0.01)	-0.004	(0.01)
Piped water					-0.02	(0.01)	-0.02	(0.01)
Urban area							-3.91	(1.05)
District fixed effects					Yes		Yes	

Table A.3 : OLS results by distance to/from former homeland border - continued

	20km radius							
	I		II		III		IV	
Variable	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Inside	**16.6**	**(0.51)**	**4.89**	**(1.17)**	**5.03**	**(0.97)**	**4.03**	**(0.95)**
% Aged 35 - 49			-0.41	(0.04)	-0.32	(0.04)	-0.32	(0.04)
% Aged 50 - 64			-0.46	(0.04)	-0.34	(0.04)	-0.36	(0.05)
Black			0.14	(0.05)	0.15	(0.04)	0.13	(0.04)
Female			0.19	(0.05)	0.14	(0.04)	0.14	(0.04)
No schooling			0.84	(0.05)	0.84	(0.04)	0.82	(0.04)
Primary school			0.82	(0.06)	0.83	(0.05)	0.81	(0.05)
Secondary school			0.76	(0.06)	0.78	(0.05)	0.75	(0.05)
Matric			0.53	(0.08)	0.55	(0.06)	0.53	(0.06)
Married			-0.18	(0.05)	-0.21	(0.05)	-0.21	(0.05)
Unemployed			0.03	(0.01)	0.03	(0.01)	0.03	(0.01)
Electricity					0.0002	(0.004)	0.0004	(0.00)
Piped water					-0.02	(0.00)	-0.02	(0.00)
Urban area+							-4.16	(1.04)
District fixed effects					Yes		Yes	