The persistence of apartheid regional wage disparities in South Africa

Gibson Mudiriza and Lawrence Edwards

ERSA working paper 816

March 2020

The persistence of apartheid regional wage disparities in South Africa*

Gibson Mudiriza[†] and Lawrence Edwards[‡]

April 2, 2020

Abstract

Despite the ending of apartheid, regional wage disparities remain prevalent in South Africa with the former homelands characterised by persistently low wages and incomes. In this paper, we use a new economic geography (NEG) model to estimate the extent to which the persistence in apartheid regional wage disparities are an outcome of economic forces such as access to markets. We estimate a structural wage equation derived directly from the NEG theory for 354 regions over the period 1996 to 2011. We find support for the NEG model in explaining regional wage disparities across regions in South Africa, although the market access effects are highly localised in view of high distance coefficients. We also find a wage deficit in homeland areas even after controlling for NEG and other region-specific characteristics. Average wages of workers in homeland areas were 11.8% lower than predicted in 1996, with this gap rising to 13.3% in 2011. This gap remains using alternative estimation approaches and the inclusion of controls for infrastructure, removal of incentives under the deconcentration policies and local amenities.

Keywords: Economic geography; Labour market, Wage differentials, Regional economic activity, Economic development

JEL Codes: F12, F16, J31, R11, 010

1 Introduction

Spatial variations in wages and income are a striking and persistent feature of the South African economy.[1] While spatial inequality is a common feature of

*We thank participants of the CSAE 2018 Conference for comments on an earlier draft. We are also grateful for the funding provided by African Economic Research Consortium (AERC), the South African National Research Foundation (CSUR 14080989186, Unique Grant No: 93648) and the Carnegie Foundation to carry out the research. All errors and omissions remain our own.

[†]Corresponding author: Post-doctoral Research Fellow, School of Economics, University of Cape Town, South Africa. Email: gmudiriza@gmail.com.

[‡]Professor, School of Economics, University of Cape Town, South Africa. Email: Lawrence.edwards@uct.ac.za.

[1]While poverty rates have declined post-apartheid, they remain very high in some municipalities (up to 63% in 2015) with very little change in the spatial pattern of poverty over

emerging economies (World Bank, 2009; Breinlich et al. 2014), South Africa is distinguished by the coincidence in the spatial distribution of low wages, incomes and high poverty and the location of apartheid homelands where black South Africans were forcefully relocated according to their ethnic groups Kingdon and Knight, 2006; Noble and Wright, 2013; Abel, 2019; von Fintel, 2018). These differences in economic outcomes have persisted despite the ending of apartheid with the advent of democracy in 1994 and the spatially targeted development initiatives and redistributive fiscal policies that followed (World Bank 2018a, b; Todes and Turok, 2018). Such uneven development has the potential to undermine social cohesion and foment political and economic tensions (Breinlich et al., 2014; Kanbur and Venables, 2005). These tensions may underpin the dramatic rise in municipal service delivery protests experienced in South Africa from 2008 (Municipal IQ, 2019).[2]

In this paper, we use a new economic geography (NEG) model to estimate the extent to which the persistence in apartheid regional wage disparities over the period 1996-2011 is influenced by economic forces such as access to markets. Central to NEG theory is the interaction of increasing returns to scale, transport costs and consumers' love of variety that together create pecuniary externalities in agent's location decisions. A key prediction of the theory, as formally shown by Fujita et al. (1999), is that wages are higher in regions with greater access to markets.[3]

Many of the homelands were designated in areas distant from the economic centres. Consequently, workers in these areas were penalised by their remoteness from markets in addition to the effects of the many other discriminatory apartheid policies affecting the labour market.[4] While the ending of these discriminatory policies could have been expected to have led to convergence in wages across regions, regional disparities may actually have been further entrenched by market forces unleashed with the liberalisation of the economy from the early 1990s.

Tests of the spatial wage structure implied by NEG have been conducted for several advanced economies, including Italy (Mion, 2004), Germany (Brakman et al. 2004; Kosfeld and Eckey, 2010), the United States (Hanson, 2005; Fallah et al., 2011) and Spain (Pires, 2006). More recently, empirical applications have extended to emerging economies such as Brazil (Fally et al., 2010), Indonesia

time (World Bank 2018a). Bosker and Krugell (2008) use spatial Markov chain techniques to reveal heavily diverging regional income distribution in South Africa from 1996 to 2004.

[2]According to Municipal IQ (2019) major municipal-level service delivery protests rose from 27 to 237 over the period 2008-2018.

[3]See Harris (1954) for an earlier version of the market-potential relationship where he specifies demand for goods produced in a location as a function of purchasing power in other locations, weighted by transport cost. Fujita et al. (1999) provide microeconomic foundations for this relationship using spatial general equilibrium models.

[4]These includes education policies under the Bantu Education Act (1953) that enforcing racially separated educational facilities with inadequate funding and training provided to African education institutions, job reservations for Whites and restrictions on the employment of Africans in skilled occupations (Native Building Workers Act, 1951) as well as laws prohibiting black Africans from forming unions and engaging in strike action (Native Labour Act, 1951).

(Amiti and Cameron, 2007), China (Hering and Poncet, 2009, 2010; Moreno-Monroy, 2011) and Chile (Paredes and Iturra, 2012; Paredes, 2015).[5] In general, market access is shown to have a positive effect on wages, but the benefits are localised given high distance coefficients (Chang et al., 2015). The strength of the effect also varies across regions within countries and in some cases, where resource dependency is high, e.g. Chile, the model performs poorly (Paredes and Iturra, 2012).[6] What is missing from this literature is a study on NEG and South Africa's wage disparities.[7]

The focus on South Africa provides a valuable contribution to the empirical literature for several reasons. Firstly, South Africa is one of the most unequal countries in the world. The income Gini coefficient is above 60 and has been rising since 1993 (Leibbrandt et al., 2010). The dominant contributors to this Gini coefficient and its rise are wage income and rising wage inequality (Wittenberg, 2017a, b).[8] However, the contribution of regional variation in wages towards overall inequality has not yet been comprehensively studied, despite evidence of vast differences in wage income across the rural-urban divide (Leibbrandt et al., 2010). Studies that do focus on regional wages include Kingdon and Knight (2006) and Magruder (2012), but their focus is not on explaining wage disparities across regions.[9]

Secondly, while several studies in South Africa have drawn on NEG theory to explain regional disparities in GDP per capita (Naudé and Krugell, 2006); output per worker (Krugell and Rankin, 2012) and export performance Naudé and Gries, 2009), these studies estimate reduced-form relationships and not the tight theoretically derived specification of the NEG model. A partial exception is von Fintel (2018), who grounds his decomposition of spatial earnings inequality in South Africa more strongly around NEG theory. He argues that while agglomeration forces explain large proportions of spatial inequality currently in

[5] For a review of the literature see Chang et al. (2015).

[6] While the signs are generally consistent with the theory, in many cases the sizes of the coefficients are implausible. For example, Hanson (2005) for the US and Brakman et al. (2004) for Germany specification estimate unrealistically high shares of expenditures on manufactures that are close to 1. In addition, in the case of Germany, the 'no black-hole' condition that determines whether transport costs affect spatial agglomeration of economic activity is not satisfied in several estimates (Brakman et al., 2004). The functional form and parameter constraints imposed by the theory are often too restrictive. For example, the simple Harris-based market potential function fits the data better than the strict theoretically derived Helpman-Hanson model in the case of Germany (Brakman et al., 2004; Niebuhr, 2006) but not the US (Hanson, 2005) and Spain (Pires, 2006).

[7] Studies on wage disparities within and between African countries are also rare. A notable exception is Bosker and Garretsen (2012) who find support for the wage equation across a group of African countries.

[8] Leibbrandt and Woolard (2001) found that wage income contributed 73.5% towards overall household income inequality in 1993. This share rose to 85% by 2008 (Leibrandt et al., 2012). Importantly in the case of South Africa where unemployment rates are very high, the lack of access to wage income is a considerable contributor towards this share.

[9] Magruder (2012) used magisterial districts as the unit of analysis to examine the effects of bargaining council on various labour market outcomes (employment, employment by firm size and wages by industry), while Kingdon and Knight (2006) used individual workers identified by their location (360 clusters, magisterial districts, districts councils and provinces) as the unit of analysis to examine the effects of local unemployment rate on individual wage.

South Africa, apartheid separate development policies exacerbated this inequality. Without these policies, the densely populated homeland areas would likely have developed into high paying local economies. Nevertheless, even von Fintel (2018) applies reduced form estimates of the NEG model using geographically weighted regressions explaining log earnings. These studies therefore only provide limited insights into the theoretically consistent mechanisms through which NEG forces drive regional wage disparities.

Thirdly, the application of the NEG model to South Africa allows us to assess the validity and robustness of NEG forces in driving regional wage disparities in resource-abundant emerging economies. South Africa is a resource-abundant economy, whose spatial development is closely associated with the exploitation of natural resources such as gold and diamonds (Wilson, 2001). Regional wages are thus likely to be influenced by a combination of first nature factors (mining endowments, climate, agricultural land), as well as second nature (NEG) factors that relate to the location of economic agents relative to one another in space (Venables, 2005). In Chile, an economy that is also resource-abundant, wages are poorly explained by NEG (Paredes, 2015). In this paper, we see whether this finding also holds for South Africa.

Fourthly, post-apartheid South Africa presents a context in which economic geography forces are anticipated to play an increasing role in the determination of wages and location of production. The ending of economic sanctions and the extensive liberalisation of trade helped integrate South Africa more strongly into the global trading environment (Edwards, 2005). The internal economy also opened up with the legal reintegration of the homeland areas into South Africa from 1994 (Clark and Worger, 2011) and the adoption of market-oriented policies including the deregulation of agricultural markets, telecommunications and finance (Lahiff et al., 2007).[10] Finally, mining as a share of the economy has declined, raising the relative importance of manufacturing (and services) in the economy, and consequently the importance of NEG forces of economies of scale, transport cost and manufacturing product varieties in driving the location of employment opportunities and local-level wages. Our study provides insight into how these developments are reflected in the changing importance of market access in determining regional wages.

Fifthly, our study will allow us to assess the extent to which the legacy of political institutions, such as the homeland system in the case of South Africa, persist over time. Adjustments to economic geography forces can be slow with large disparities sustained over many decades (Breinlich et al., 2014). In this paper, we estimate the contribution of market potential to the homeland wage deficit and consequently the relative importance of other non-economic geography factors in explaining the persistence in the wage deficit.

Finally, a central argument of the 2009 World Development Report (World Bank, 2009) was that while growth may be unbalanced, trade and migration

[10] The repeal of the remaining apartheid laws took place in 1990, but the effective end is regarded as the date of the democratic general elections in April 1994, following which the homelands were legally reintegrated into South Africa (Clark and Worger, 2011; Bastos and Bottan, 2016).

spread the gains to poorer regions and narrow the wealth gap. This is particularly relevant for South Africa, as income and production activities are concentrated in a few core regions (Bosker and Krugell, 2008). Using the estimated NEG model, we can see how growth in these economic centres filters down to outlying regions.

To analyse how NEG forces shape South African regional wage disparities, we estimate Brakman et al.'s (2004) extension of Hanson's (2005) NEG model that allows for real wage differences across regions. Hanson's (2005) model, which is based on Helpman (1998), assumes perfect mobility of labour within a country resulting in real wage equalisation across regions. Given South Africa's history of restrictions on the mobility of people, this assumption is unlikely to hold, explaining our preference for the Brakman et al. (2004) model.

For our empirical analysis, we augment the model to control for natural resource endowments and include indicators of the region's former homeland status to estimate the homeland wage gap after accounting for NEG forces. We apply the model to wage data for 354 regions (magisterial districts) that we obtain from the Population Censuses for 1996, 2001 and 2011.

We find support for the NEG model in explaining regional wage disparities across regions in South Africa, although the market access effects are localised in view of high distance coefficients. One implication of this finding is that growth in the economic centres has a weak impact on wages in other regions. We also find a wage deficit in homeland areas even after controlling for NEG and other region-specific characteristics. This deficit has persisted over time, despite a rise in the importance of market access in driving wage dispersion across regions in South Africa. Average wages of workers in homeland areas were 11.8% lower than predicted in 1996, with this gap rising to 13.3% in 2011.

We test for several potential explanations for the persistence of the homeland wage deficit. Our results suggest that wages in the former 'independent' homelands (Transkei, Bophuthatswana, Ciskei and Venda), are relatively high and rising compared to the other homelands, although wages in these 'independent' homelands have also not converged to what we would predict based on economic geography forces. We find that the wage penalty remains high and increases once we control for infrastructure, local amenities and proximity to the apartheid government regional industrial development programme zones. The results remain robust to several different estimation approaches and controls.

Our findings show that the ending of apartheid did not result in a reduction in the homeland wage gap despite the reintegration of homeland areas into South Africa and the implementation of regional policies since 1994. Overall the study shows that first and second geography factors together with the apartheid-era homeland policy are key factors in explaining the persistence of regional wage disparities in South Africa.

1.1 Spatial development in South Africa

South Africa's spatial economy is influenced by a combination of its history as a trading port, the discovery of minerals and the apartheid policy of separate development. During the 17^{th} and 18^{th} centuries, South Africa's spatial economy emerged in response to geographical differences in access to waterways, climates, and natural resources. For instance, the development of the port cities of Cape Town and Durban was driven by their close proximity to waterways that gave them an important role in the country as trading posts on the shipping route between Western Europe and Asia (Gelb, 2004; Bosker and Krugell, 2008). The discovery of minerals, mainly gold in the late 19^{th} century, led to the development of the mining industry that promoted rapid industrialisation, urbanization and migration of workers to support the growing industry in and around Gauteng (Turok, 2012; Todes and Turok, 2018). This allowed Gauteng to develop into the large urban agglomeration that still accounts for the bulk of the country's economic activities.

Overlaying these developments was a systematic policy of racial discrimination and spatial segregation. Although prior policies had already restricted access to land by black Africans, the primary policies that set the basis for spatial segregation along racial lines were the 1913 and 1936 Land Acts. These Acts restricted ownership by the African majority to 13% of the country's land area. Additional policies, such as the 1923 Native Urban Areas Act strengthened restrictions on access by black Africans to white areas, helping establish a large-scale, low-wage migrant labour system (Todes and Turok, 2018, World Bank, 2018b).

The new National Party government, elected in 1948, reinforced, extended and legalised racial and spatial segregation under its apartheid policy of separate development (Simkins, 2011; Clark and Worger, 2011). Under separate development, Africans were to be located in ethnically-distinct self-governed territories. A slew of legislation was imposed restricting ownership and settlement of people according to their racial classification (white, black, coloured and indian), as codified in the Population Registration Act (1950). One such policy was the Group Areas Act (1950) that systematized control over transactions and occupation of property by race group and gave rise to racially designated townships around all towns and cities in South Africa.

Other policies such as the Promotion of Bantu Self-Government Act (1959) transformed the native reserves into 'Bantu Homelands' or Bantustans (Transkei, Bophuthatswana, Ciskei, Venda, Gazankulu, KaNgwane, KwaNdebele, KwaZulu, Lebowa, and QwaQwa) that were ostensibly to become fully-fledged independent and self-governed states for the country's black African population (King and McCusker, 2007). The Black Homeland Citizenship Act (1970) gave force to this Act by denying black Africans their right to be citizens of South Africa, requiring them to become citizens of one of the self-governing territories.[11] Four of the homelands - Transkei, Bophuthatswana, Venda and Ciskei (called TBVC states) were declared constitutionally independent from South Africa (in 1976, 1977, 1979 and 1981 respectively), although their independence was not recognized outside of the country.

[11] Apart from being dispossessed of land, blacks were also dispossessed of other means of livelihoods such as ownership of livestock. The rights to own, rent or transfer property for blacks was limited to these ten homeland areas only.

The disparate spatial location of the homeland areas is illustrated in Figure 1 that shows the borders of the former homeland areas. Many of the homelands were fragmented comprising of non-contiguous areas. KwaZulu, Bophuthatswana and Lebowa were extreme examples of this.[12] In general, homelands were located remote from major markets, road networks, international airports and harbours, with the exception of Ciskei, parts of Kwazulu Natal and KwaNdebele that bordered the cities of East London, Durban and Pretoria, respectively. The location of these homelands enabled the apartheid government to implement policies of separate development along racial lines. Homelands were allocated limited resources leading to poorly developed public infrastructure (water, sanitation, electricity and transport) and inferior schooling facilities (teachers, classrooms, libraries, laboratories and sports fields).

A further feature of apartheid policies were the severe restrictions on the mobility of Africans. Employment in areas reserved for whites was controlled bypass laws that required all black Africans to carry a passbook stipulating where and for how long the person could remain in the white designated area. To further discourage migration and permit the homeland areas to become self-sustaining, the apartheid government implemented the Industrial Decentralization Strategy in 1961 and then the more extensive National Regional Development Plan in 1982. Under these policies manufacturing firms were provided generous employment subsidies, rebates and other incentives to locate in designated Growth Points/Industrial Development Points in and around homeland areas and Deconcentration Points located on the fringes of metropoles (Nel, 1994). The location of these zones is also illustrated in Figure 1. While these were effective in attracting industries to the designated areas, the costs were found to be very high, and an aspatial approach to incentives under the New Regional Industrial Development Programme was adopted from 1991 (Nel, 1994; Bell, 1997).

Following the advent of democracy in 1994, several redressal measures including black economic empowerment, provision of social grants, better housing and education were implemented, to uplift the social and economic outcomes of previously disadvantaged race groups (Jensen and Zenker, 2015). New administrative boundaries were demarked with the goal of dissolving the racial based spatial layout created by the apartheid system. This led to the creation of local authorities constitutionally responsible for the development of their areas Bosker and Krugell, 2008). Further, to level out regional imbalances several policies were adopted including the National Spatial Development Framework (NSDF) of 1995, the Spatial Development Initiatives (SDIs) of 1996, the Regional Industrial Development Strategy of 2006, and the National Spatial Development Perspective (NSDP, 2003; 2006). The spatial policies initially incorporated a strong spatial rebalancing objective, but the focus then shifted towards a space-neutral and place-based approach, although in a fragmented way (Todes and Turok, 2018).

Despite these efforts, former homeland areas remain underdeveloped with

[12] KwaZulu was made up of approximately 70 segments (Clark and Worger, 2011).

high levels of unemployment, poverty, deprivation, low wages and slow industrialisation Kingdon and Knight, 2006; Noble and Wright, 2013; Abel, 2019; von Fintel, 2018). For example, the 10 poorest municipalities in the country in 2016 fell within the former homeland areas (World Bank, 2018a). Using multiple deprivation indices for 2001, Noble and Wright (2013) found that 92% of the population in former homelands were income deprived compared to 64% for the rest of South Africa. According to the living environment domain index, 89% of the population were deprived, compared to 50% in the rest of South Africa. The implication is that despite two decades after the ending of apartheid, apartheid-era spatial inequalities (disparities) continue to be reproduced (Nel and Rogerson, 2009).

2 Model and estimation

The New Economic Geography (NEG) theory offers a theoretical framework to explain regional wage disparities. This theory consists of several models among them Krugman (1991), Krugman and Venables (1995), Venables (1996), Fujita et al. (1999), Helpman (1998) and Redding and Venables (2004). In these models, the interaction of transport costs and increasing returns to scale with either labour mobility (Krugman, 1991; Helpman, 1998) or intermediate inputs (Krugman and Venables, 1995) generates agglomeration and dispersion forces that determine the spatial distribution of economic activity, income and wages. While these models have different agglomeration and dispersion forces, they share a common feature in that they give rise to a spatial wage structure where regional wage levels are positively correlated with access to markets. The most prominent representation of this spatial wage structure is set out in .Fujita et al. (1999) as follows:

$$w_r = \left[\sum_{i=1}^{R} Y_i I_i^{\sigma-1} T_{ri}^{1-\sigma}\right]^{\frac{1}{\sigma}} \tag{1}$$

where w_r denotes the nominal average wage in region r, and the term in the square brackets denotes the region's market potential (market access) measured as the sum of income (Y_i) in other regions, weighted by transport costs (T_{ri}) and a composite manufacturing price index (I_i). According to equation (1), nominal wages in a region are rising in the income of surrounding regions, decreasing in the transport costs to these regions and increasing in the price of competing traded goods in these locations (Hanson, 2005). The theory, thus, predicts a spatial wage structure, where wages are systematically high in regions with greater access to markets.

Estimation issues

A major challenge to the direct estimation of equation (1) is the need for manufacturing price indices data I_i across regions, the availability of which is often lacking. The price index I_i is also endogenous and is itself a function of regional wage rates. Hanson (2005) draws on Helpman's (1998) model to resolve this problem by assuming real wage equalisation across regions and re-writing the manufacturing price index (I_i) in terms of income, wages and the fixed stock of housing in a region.[13]

Real wage equalisation assumes the economy has reached a long-run equilibrium where labour has no incentive to migrate (Brakman et al., 2004). This is highly restrictive in the case of South Africa given its history of movement restrictions on the black majority during much of the apartheid-era rule. Although there has been increased internal migration to urban areas in response to economic opportunities (World Bank, 2018b), population densities in the homeland areas continue to be much higher than other regions reflecting a persistency in the legacy of apartheid on the spatial distribution of the population (Reed, 2013). Further, the wage data for South Africa indicates regional variation in nominal wages that far exceed possible price differences, implying substantial and sustained real wage differences across regions.[14]

Our approach is therefore to follow Brakman et al. (2004) who derive an estimable wage equation that does not invoke the real wage equalisation assumption. They achieve this by replacing the manufacturing price index in equation (1) with a simplified local price index based on the average distance between regions. Defining the transport cost function by an exponential decay function ($T_{ri}= e^{\tau d_{ri}}$ where d_{ri} is the distance between regions r and i), taking logs of equation (1) and adding a disturbance error term ε_r, the wage equation to be estimated becomes:

$$\log\left(w_r\right) = k_0 + \sigma^{-1} log\left[\sum_{i=1}^{R} Y_i I_i^{\sigma-1} e^{-\tau(\sigma-1)d_{ri}}\right]^{\frac{1}{\sigma}} + \varepsilon_r \tag{2}$$

with

$$I_i = \left[\lambda_i\left(W_i e^{\tau D_{ii}}\right)^{1-\sigma} + \left(1-\lambda_i\right)\left(\bar{W}_i e^{\tau D_{i-centre}}\right)^{1-\sigma}\right)\Big]^{\frac{1}{\sigma-1}} \quad . \tag{3}$$

$\bar{W}_i$ is the average wage outside region i, D_{ii} is the internal distance of region i,

[13]Using the relationship that the market value of housing services supplied equals the share of income (1-δ) spent on housing services ($P_i H_i = (1-\delta)Y_i$ where P_i is the price of housing services) and assuming real wage equalisation across regions ($w_i/P_i^{1-\delta} I_i^{\delta} = \bar{\omega}$), the manufacturing price index can be rewritten as $I_i = \left[w_i Y_i^{\delta-1} H_i^{1-\delta} k\right]^{1/\delta}$ where k captures all terms that do not vary across regions. Substituting this modified price index into equation (1) yields the function estimated by Hanson (2005) that can be estimated given data on housing stocks, wages and incomes by region. An alternative approach is to asume that regions have a similar price index (I_i =1). This yields the simple Harris (1954) market potential index. Although popular in empirical circles, the Harris (1954) market potential index has been criticised for its lack of theoretical foundations. Alternatively, Redding and Venables (2004) extend the NEG model to specify wages as function of market and supplier access. This strategy is commonly used to explain cross-country differences in wages, but it is less appealing for within-country analysis as it requires inter-regional trade data within the country and assumes intra-country labour immobility.

[14]For example, in 2002, the average price of a medium sized house (140m^2 – 220m^2) in the most expensive province (Western Cape) exceeded that in the cheapest province (Northern Cape) by a factor of 1.9 (Luüs, 2005). The gap in average wages between these provinces in 2001 was 1.3.

$D_{i-centre}$ is the distance from region i to the economic centre (Johannesburg) and λ_i is a weighting parameter given by region i's share of employment.[15]

By estimating equation (2), we can obtain the values of σ - the elasticity of substitution between the different manufactured varieties and τ - the transport costs parameter. From the theoretical model we expect, $\sigma > 1$ and $\tau > 0$ and based on these parameters we can further derive the effect of market potential $(1/\sigma)$ and distance $(\tau(1-\sigma))$ on wages as well as the mark-up margin $(\sigma/(\sigma-1))$.

There are two further considerations in estimating the wage equation (2). Firstly, our results may be biased and inconsistent due to omitted variables. Apart from market potential, existing evidence stresses the importance of various first nature geography factors, as well as other location-specific factor endowments in explaining regional wage disparities (Combes et al., 2008; Fallah et al., 2011). Thus, to deal with omitted variable bias, we augment equation (2) with several controls that will be discussed in the data section.

Secondly, our results may be biased due to endogeneity problems between the explanatory variables and the error term. In equation (2), the dependent variable, wage, appear directly as an independent factor in the model. This is a potential source of bias as w_r and ε_r will be correlated. In addition, regional wages and incomes are simultaneously determined and this leads to possible correlation of income (other explanatory variables) and the error term, ε_r. Finally, the error term, ε_r, might also be correlated with explanatory variables (Y_i, I_i and w_i) in other regions that are all elements of the market potential of region r (Mion, 2004).

To address the endogeneity problem, we estimate equation (2) using generalised methods of moment (GMM) estimation, where we instrument market access using an index of terrain ruggedness obtained from Nunn and Puga (2012)[16] and lagged income, wage and population levels of the district as well as the closest neighbour by distance. Use of these instruments is consistent with other studies that have estimated NEG models (see Hanson, 2005; Moreno-Monroy, 2011). We show validity of these instrument's overidentification restrictions using the Hansen J statistic.

2.1 Data

The main data set for our empirical analysis is the population census for 1996, 2001 and 2011 collected by Statistics South Africa (Stats SA). The population censuses contain a rich set of information on demographics and labour market outcomes (including employment, industry, and income), among others, at a disaggregated spatial level.

Nevertheless, several challenges arose in using the census data. Changes

[15] The economic centre is defined as the region with the smallest average distance to other regions, calculated as $\bar{D}_r = \sum_i (Y_i / \sum Y_i) D_{ri}$.

[16] Nunn and Puga (2012) provide terrain ruggedness data on their website - https://diegopuga.org/data/rugged/ which allows us to derive the terrain ruggedness index. This index that was developed by Riley et al., (1999) allows us to quantify topographic heterogeneity by indicating how jagged or flat the terrain of a region is on average.

in the country's administrative boundaries over the period required us to use ArcGIS to overlay 2011 sub-place boundaries onto 1996/2001 magisterial district boundaries.[17] This provides us with a consistent longitudinal and cross-sectional dataset of 354 magisterial districts (hereon regions).

For wages (w_r), we use average income of workers, aged 15-64 years in each region. Three challenges in this regard are worth noting. Firstly, the census income covers all sources of income including basic salary, bonuses, allowances, income from grants, transfers, remittances and any other income source received by individuals. To test the possible distortion this may introduce in our analysis, we compare income per worker and wage per worker across 53 districts in South Africa based on the National Income Dynamics Study (NIDS) 2010/2011 household survey data. As discussed in more detail in Appendix B, we find that wages on average account for over 90% of total income of workers. Further, using simple regressions explaining regional average incomes of workers, we show that after accounting for average wages, no other district-specific factor has significant explanatory power of worker's incomes. This evidence highlights that average worker's income is a good proxy for average worker's wages in South Africa.

A second consideration is that the census income is provided in brackets with an open-ended top bracket. We address this challenge by assigning the midpoint of each income bracket to everyone in that bracket and setting the midpoint of the open-ended bracket to two times the lower bound of the highest bracket. This is an approach commonly followed by Statistics South Africa (Statistics South Africa, 2000: 9). Finally, between 2% (1996) and 13% (2011) workers in the census data have zero reported or missing information on incomes. Our approach is to drop these observations from the sample.[18] Background analysis of the missing income data reveal potential selection biases as the missing data are correlated with individual characteristics, e.g. black Africans, rural inhabitants and older workers are more likely to report income levels (see Appendix B). To control for the possibility of selection bias we test the sensitivity of our results to inclusion of region-level controls for these variables. For convenience, from hereon we refer to regional average worker's income as regional wage.

With respect to the other variables, we use the sum of personal income to proxy each region's total income (Y_r), calculate distance (d_{ri}) as the great-circle distance (in kilometres) between the geographical centres (centroids latitude and longitude) of region i and r, and, following Redding and Venables (2004), measure internal distance (d_{ii}) as $d_{ii} = 2/3\sqrt{area/\pi}$, where area is measured in square kilometres.

As controls, we include the regional shares of workers (in total working-age

[17] We assign the 2011 population values of each sub place unit to their corresponding 1996/2001 magisterial districts based on a ratio of the area common to both sub place and magisterial districts. We use the smaller sub-place units as opposed to municipality units to minimise the error associated with the areal-weighting interpolation technique that assumes homogeneously distributed population across space (Maantay et al., 2008).

[18] Unlike Bastos and Bottan (2016) who include the zero-income data and count missing values as zero in their calculations of average income by region.

population) with tertiary education to capture the effects of regional differences in skilled workers (human capital). Failure to control for human capital can bias the estimated market access coefficients as the wage premium for skilled workers in areas can be reinforced by human capital accumulation (Mion, 2004). We also separately estimate the wage equation for high skilled and low skilled workers. To capture the effects of local labour market inflexibility and the presence of a local wage curve, as found for South Africa in the early 1990s by Kingdon and Knight (2006), we include regional unemployment rate.

To account for the influence of mineral resource endowments, we include the regional shares of employment in mining. This measure enables us to control for the influence of a key first nature geography factor, mining activity, that has played a prominent role in South Africa's spatial development. As additional controls for first nature geography factors, we draw on the climate (rainfall and temperature) data produced by Harris et al. (2014) for the Climatic Research Unit (CRU) at the University of East Anglia that we map to each magisterial district for 1996, 2001 and 2011.

Finally, to account for the effects of the apartheid-era homeland policy, we incorporate a homeland status indicator derived from the share of each region's area that falls in former homeland areas.[19] We select a cut-off point of 30% but test the robustness of the results to a more stringent 70% cut-off point. We interpret the coefficient on this variable as a measure of the conditional homeland wage penalty.

Summary statistics

Table 1 presents summary statistics of the homeland areas compared to the rest of South Africa for the years 1996 and 2011. Homelands perform poorly in key dimensions in both years. They are characterised by lower labour force participation (35-36% vs. 55-56%), and of those in the labour market, a higher proportion are unemployed (45-60% vs. 27.5%). The consequence is a substantially lower share of the working-age population with employment (15-20% compared to 41-42%). Using the share of households that use electricity for lighting as a proxy for economic infrastructure, we also see substantially lower levels in homelands, although this share has risen strongly from 1996 with the mass electrification drive (from 26% in 1996 to 75% in 2011). The provision of municipal/town services, as proxied by the share of households with refuse collection, has increased on average for all regions, but remains substantially lower in homelands at 10-15% compared to other regions (62-69%). The population density is far higher in homelands and has risen, but at a slower pace than the rest of the country as migration to urban areas has taken take place (World Bank, 2018b).

Homeland areas are also characterised by low average distance from the economic centre (Johannesburg) than the rest of the country. Despite their

[19] To derive the indicator, we use ArcGIS to overlay magisterial district boundaries to former homeland boundaries. From the resulting mapping, we use areal-weighting interpolation technique to derive a ratio based on the area of each magisterial district that fall in a given homeland area. The ratio ranges between 0 and 1, with 1 indicating that a district falls completely in a homeland area, while 0 indicates that a district falls in a non-homeland area.

proximity to the economic centre relative to the rest of the country, homeland areas remain underdeveloped. In addition, homeland areas tend to receive more rainfall compared to the rest of the country. However, these areas have low agricultural productivity.

Looking at average regional wages, average wages in homeland areas were 14% lower in 1996 than in other regions of South Africa. While real average wages in homeland areas rose by 61% over the 1996-2011 period, growth was nevertheless slower than in the rest of South Africa, raising the homeland wage penalty to 19% by 2011. The wage penalty is rising, yet the homeland areas have high proportions of skilled workers (13.9 and 18.4% vs. 8.8 and 14.0%) than the rest of the country. The high proportions of skilled workers in homeland areas might be explained by the high proportions of public sector workers in these areas (49.4 and 45.7% vs. 30.8 and 35.9%).

Figures 2 and 3 displays the spatial distribution of average wages across regions in South Africa in 1996 and 2011, respectively. The darker colours are regions with high average wages and lighter colours low average wages. Wages vary significantly across regions in South Africa, but there has been limited changes in the spatial pattern over time. For example, the simple pairwise correlation coefficient for regional wages in 1996 and 2011 is a high 0.83. The World Bank (2018a) similarly finds wide spatial variation in poverty that has remained stable over time. The figures further indicates relatively low wages in the former homelands (areas with blue colour boundaries) in both periods. Finally, we see some evidence of clustering of wages. Regions with high (low) wages, tend to be surrounded by other regions with high (low) wages although more so in 1996 than in 2011. This suggests the presence of NEG forces giving rise to spatial autocorrelation in wages across regions.

To further assess the presence of spatial autocorrelation in wages, Figure 4 presents the Moran scatter plot of wages for 2011. This figure plots the relationship between standardized regional wages and the average standardized wage of surrounding regions within a 100km distance.[20] The slope of the linear regression (with constant = 0) is equivalent to the Moran I statistic, that is commonly used to measure the extent to which regions are interdependent. Positive values reflect the presence of positive spatial autocorrelation of wages across regions, as opposed to randomly distributed (no slope), or negatively autocorrelated wages across space.

Figure 4 highlights several features about the spatial distribution of wages in South Africa. The positive slope for 2011 (coefficient = 0.21) provides evidence that wages in a region are positively correlated with those in surrounding areas. High (low) wage regions are generally surrounded by other high (low) wage regions. Wages are therefore spatially clustered as opposed to randomly allocated providing support for the presence of NEG forces.

Not all observations lie in the top right and lower left quadrants with several low (high) wage regions located in the vicinity of high (low) wage regions. These

[20] For simplicity purposes, we use uniform weights in the spatial weight matrix rather than an exponential or power functional types of weights where the weight decays with distance.

spatial discontinuities suggest the presence of other local-specific factors driving regional wages. Homeland status is one such factor. As shown by the red cross symbols, homelands, in general, have lower wages and are surrounded by other low wage regions.[21] This is consistent with the idea that homelands were located remote from markets. Their low wages, in part, reflect their 'poor' location relative to markets.

However, we also find many homeland areas situated close to regions with higher than average wages (see top left quadrant). These regions are primarily found within the former homelands of Bophuthatswana, KwaNdebele, and Lebowa that were located close to regions in the economic centre, now Gauteng. Wages in these areas are disproportionately low relative to their neighbours suggesting a larger than expected wage deficit given their geographical location close to large markets.

Finally, the figure presents the linear regression line for 1996. As can be seen, the spatial autocorrelation of wages appears to have diminished between 1996 and 2011.[22] Separate tests of the data indicate that some of this decline arises from a diminished contribution of homelands towards spatial autocorrelation over the period 1996 to 2011. This suggests that homeland areas are becoming more isolated from economic geography forces. Further insight on this will be obtained from the estimates of the theoretically derived market access wage equation.

3 Empirical results

Our aim is to first establish whether the spatial distribution of wages in South Africa are explained by the Brakman et al. (2004) extension of the NEG model that allows for real wage differences across regions. To do so, we present several estimation results commencing with a baseline model (2) specification that includes a limited range of controls covering skill composition, share employment in mining, the unemployment rate and a dummy for homeland status. We then sequentially extend this to include additional controls or restrictions. Estimates are performed separately for 1996, 2001 and 2011 as well as using the pooled data.

Baseline results of the Helpman-Hanson model

The results of several GMM estimations are presented in Table 2. Before presenting the results, we observe that according to the Hansen J-stat test of overidentifying restrictions we are unable to reject the null-hypothesis that the instruments are uncorrelated with the errors for all years. This gives us some confidence that our instrument set is valid.

[21] Districts in the former homelands of Transkei, Ciskei and Kwazulu make up the bulk of these regions.

[22] Tests show a statistically significant decline in the slope coefficients from 0.31 in 1996 to 0.21 in 2011. The slope in 2001 (0.26) is not significantly different from the slopes in 1996 and 2011. We also find no significant correlation between wages and those in surrounding areas for former homelands.

The two key parameters establishing the presence of a wage structure implied by NEG are the elasticity of substitution (σ) and transport costs (τ). Looking first at the baseline results for 1996 presented in column (1), the model presents a poor fit to the regional wage data. The elasticity of substitution, σ is very high (69.3) in 1996, although it imprecisely estimated. While τ, the measure of transport costs, is significant, the effect of distance on wages depends on the value of τ(1-σ) that is large, but not significantly different from zero. The implication is that we find no association between market access and wages in 1996, as we would expect under the NEG theory.

In contrast, the results presented in columns (2) and (3) support the presence of a spatial wage structure for 2001 and 2011, although the relevant coefficients are in several cases only significant at the 10% level. The estimated value of σ is greater than 1, as required by the NEG model and ranges from 5.4 (2001) to 4.4 (2011). The value of σ suggests that firms across regions in South Africa are operating under increasing returns to scale, with price-cost margins (given by $\sigma/(\sigma - 1)$) of 1.23 (2001) and 1.3 (2011). Given σ, the estimated coefficients on market potential ($1/\sigma$) are positive and imply that a 10% increase in market potential is associated with a 1.85% to 2.29% increase in wages in that region over the period 2001 to 2011. Looking within the market potential measure, the transport cost parameter tau is positive, meaning that as a region's access (in terms of distance) to markets decreases, its average wage levels also decrease.

The results are broadly comparable to those of other studies. The estimated price-cost margins correspond closely with estimates by Aghion et al. (2008) of between 1.21 and 1.23 for the manufacturing sector in South Africa over the period 1976 to 2000. The estimates of σ for South Africa also fall within the range estimated for other countries. For example, they straddle the median estimate of 4.93 for the 262 NEG studies covering within and across country wage differences surveyed by Bosker and Garretson (2010). Compared to similar NEG studies looking at wage dispersion within countries, our 2001 and 2011 results exceeds those of the US (GMM estimates of Hanson (2005, p.19), Spain (Pires, 2006, p.103), Italy (Mion, 2004) and China (Moreno-Monroy, 2011, p.19) that range from 1.7 to 3.03, but fall within the range estimated for Germany (Brakman et al., 2004) and Java in Indonesia (Amiti and Cameron, 2007, p. 23 Table 2) and are substantially lower than the 41.1 to 46.1 reported for Chile (Paredes, 2015, p.73).

Cross-study comparisons of trade costs as measured by τ are more difficult as estimates (including of σ) are sensitive to the measure of distance used and the choice of trade cost specification (Bosker and Garretsen, 2010). In comparison to similar studies that use the exponential distance function to model transport costs, our coefficient estimate for τ of 0.102 per km in 2011 is large relative to countries such as Spain (Pires, 2006), US (Hanson, 2005) and Java in Indonesia (Amiti and Cameron, 2007) (all less than 0.02), but similar or lower than Italy (Mion, 2004) and China (Moreno-Monroy, 2011). The net effect of distance on wages, as given by τ(1-σ), however, appears relatively high in South Africa, at least compared to advanced economies, given the combination of relatively high trade costs and a high elasticity of substitution. The implication, which will be explored in more detail later, is that the effects of changes in market potential on wages are highly localised.

Looking at the control variables, the skilled worker employment share is generally positive, signifying, the importance of controlling for human capital accumulation (Mion, 2004). The unemployment rate is negative and weakly significant (at 10% level) for 2011, pointing to the possible presence of local wage curve, as was found by Kingdon and Knight (2006) for South Africa in the early 1990s. Surprisingly, local wages are not explained by the share of mining employment in any period. One explanation is that given South Africa's historical pattern of development around mining, the mining intensive regions coincide with areas with high market potential. The aggregation of data to regions may also obscure some of the local level wage effects of mining, as is found for South Africa by Bastos and Bottan (2016).

Homeland wage penalty

Of central interest to us is the homeland wage penalty. The coefficient on the homeland dummy variable is negative and significant (at 10% or below) in all years suggesting that regions in homeland areas have low wages compared to other regions, even after controlling for their location (differences in first and second nature geography factors). Further, the wage penalty has remained remarkably resilient despite the abolishment of apartheid policies, the reintegration of homeland areas into South Africa in 1994 and the implementation of various regional policies. The deficit of 11.8% estimated for 1996 rises to 24.3% in 2001 and while it drops to 13.3% in 2011, it remains significantly larger than zero. The legacy of apartheid-era homeland policy continues to be reflected in their workers' disproportionately low wages.

In columns (4) to (6), we extend the basic model to allow for spill-over effects of homeland onto bordering regions by including a dummy variable (=1, 0 otherwise) for regions within 50km of the former homeland areas.[23] In general, we obtain more precise coefficient estimates of the elasticity of substitution and distance for 2001 and 2011. The elasticity of substitution also falls slightly to 5.27 and 3.09, respectively. The estimated homeland wage penalty does not change much for 1996 and 2001 but now rises to 22.3% for 2011, indicating a widening penalty over time.

Looking at the dummy variable for regions bordering homelands, a significant negative coefficient is estimated that increases from -0.06 in 1996 to -0.12 in 2011. The wage penalty in bordering regions is, however, not as large as that of the homeland areas. Nevertheless, the rising penalty for these regions suggest that the lagging behind of regions in terms of wages is more pervasive than just the homeland areas.

So far, we have treated the former homelands as homogenous units, whereas regions within and between the homelands vary enormously in terms of average wages. For example, the average monthly regional wage in KwaZulu (2225 Rands) in 2011 exceeded that of KwaNdebele (1913 Rands). Within KwaZulu,

[23] Distance is measured as the number of kilometers between the centroids of the region and its nearest homeland boundary.

regional average wages also varied from a minimum of 1467 Rands to a maximum of 5049 Rands in 2011.

Therefore, in columns (7) to (9) of Table 2, we extend the benchmark model to include a dummy variable equal to one (zero otherwise) if a region falls within one of the 'independent' homeland states (Transkei, Bophuthatswana, Venda and Ciskei). The coefficient on the TBVC dummy measures the marginal effect on wages of these regions relative to other homeland areas. The wage penalty for TBVC states is found to be no different from the other homeland areas in 1996 and 2001, but a 10.9% wage gap (only statistically significant at 10% level) between TBVC and other homelands emerges over the period 2001 to 2011. By 2011, the wage penalty for TBVC states equalled 16.6% (= 0.275 – 0.109) compared to a high 27.5% for other homeland areas. The rising wage penalty for the homeland areas found earlier can therefore primarily be attributed to regions outside of the TBVC states – their wage penalty nearly doubled from 13.9% in 1996 to 27.5% in 2011. However, even for TBVC states, the wage penalty in 2011 (16.6%) exceeds that of 1996 (13.5%). Wages in all homeland areas, therefore, continued to lag those of other regions.

Several potential explanations underpin the high and rising wage penalty faced by homeland areas. One explanation is that homeland areas remain isolated from the influence of stronger demand linkages between regions, as is reflected in the rising coefficient on market potential ($1/\sigma$) and the declining coefficient on distance. To test for this, we estimate the wage equation using pooled data where we impose constant coefficients on the NEG variables, but include year fixed effects to control for year-specific wage shocks common to all regions as well as interactions between homeland status and the year dummy variable. The final column of Table 2 presents the coefficient estimates.

The coefficients on the NEG variables (σ and τ) and homelands are precisely estimated with their sizes falling within the range of the 1996 to 2011 estimates. Consistent with the earlier results, wages in regions bordering homeland areas are 8.6% lower than expected based on their economic geography. Looking at the results for homelands, the wage penalty for non-TBVC homeland areas is estimated at 20.6% and the penalty for TBVC state significantly smaller at 14.6% (= -0.206+0.060). The coefficients on the interaction of the homeland and year dummy variables are insignificant, implying no significant change in the homeland wage penalty from 1996 to 2001 and 2011.

The pooled results imply that the apparent rise in the homeland wage penalty from 1996 to 2011 shown in columns (1) to (10) of Table 2 can largely be explained by the rising influence of market potential in driving wage dispersion across regions in South Africa. Once common market potential effects are imposed on each year, the homeland wage penalty remains stable over time. The implication is that wages in the homeland areas have risen more slowly than would be predicted on the basis of their changing market potential. Homeland areas, therefore, remain isolated from the rising demand linkages between regions. This will be looked at further when we look at changes in the spatial decay of shocks in market potential over the period.

Other omitted factors may also play a role in explaining the rising homeland wage penalty. Firstly, under the apartheid decentralisation policy, firms were incentivised to locate in designated Growth Points with RIDP zones, the bulk of which were situated within and around the former homelands. The removal of these in 1991 will have disproportionately affected homelands relative to other regions. Secondly, as shown in Table 1, homelands are characterised by poor economic infrastructure as proxied by the share of households with electricity access and piped water and poor municipal services infrastructure as proxied by the share households with refuse removal. Poor infrastructure services raise production costs and diminish the wage firms can pay in order to remain competitive and therefore may explain both the level and trend in the homeland wage penalty. Finally, while the prior estimates control for mining in the region, other first nature factors such as agricultural potential and exogenous amenities such as climate may also influence wages and spatial agglomeration. If homeland areas have poor agricultural potential and disadvantageous amenities, this would contribute towards the estimated wage penalty.

To see how these potentially confounding factors might influence our homeland wage penalty, we extend the pooled specification to include variables that control for their influence. If these channels contribute to the homeland wage penalty, we would expect a diminished coefficient on the homeland dummy variable.

Table 3 presents the coefficient estimates, with column (1) comprising the benchmark regression (similar to column (10) of Table 2, but without the TBVC dummy variable) and the other columns containing results for estimates that control for the different channels. The heading of each column denotes the control variable included in the estimates. Looking first at column (2) that test for the influence of the RIDP zones. The coefficient on the RIDP zones dummy variable[24] is positive and weakly significant, signifying the location of these zones in low wage regions. The coefficient on the homeland variable rises slightly (from -0.191 in column (1) to -0.213) reflecting the prevalence of the RIDP zones around the homeland areas. However, the coefficients on the interactions between the homeland indicator and the dummy variables for 2001 and 2011, remain insignificant. Controlling for the RIDP zones, therefore, has no influence on the estimated trend in the wage penalty.

Columns (3) controls for household access to infrastructure services that is calculated as the principal component of the share households with electricity, piped water, flush toilet and refuse collection that we obtain from the Census data. Columns (4) to (7) present results using the individual infrastructure measures. The coefficient on the principal component measure of infrastructure is insignificant and there is no measurable change from the benchmark coefficients for the homeland variables. Similarly, the share of households with piped water has no additional influence on regional wages.

In contrast, the share of households with electricity and piped water in a region are negatively associated with regional wages. This is surprising as in-

[24] The RIDP zones were in about 40 magisterial districts across the country, which represents about 11.3% of the districts considered in the study. Thus, the RIDP zone dummy is equal (=1, 0 otherwise) for these districts.

frastructure is expected to complement wages. There are several explanations for this outcome. Firstly, the effects of infrastructure are already captured by the market potential indicator. Regions with high market potential are characterised by better and more extensive infrastructure. The infrastructure variable, therefore, contributes little additional positive explanatory power once market potential is controlled for.[25] Secondly, the post-apartheid government implemented a roll-out of infrastructural services to rural households. This mass roll-out of electricity infrastructure has been shown to have raised employment levels of females, increased work hours of both males and females, while reducing female wages (through the increase in female participation in the labour force) and raising male wages in these areas (Dinkelman, 2011). What the results in Table 3 appear to capture are the selection effects associated with the targeting of electrification and other social infrastructure in low wage areas.[26] Once the targeting of infrastructure roll-out is controlled for, the homeland penalty actually rises, although, as with the prior results, there is no significant change in this over time.

The final columns (7) and (8) present results that control for agricultural potential using share employment in agriculture and amenities using the log of rainfall in the region. The results show no effect of the share employment in agriculture on wages, but the log rainfall has a strong positive association with regional wages. Interestingly, the inclusion of rainfall raises the homeland penalty to 25.4%, despite these areas having higher average rainfall than other areas. One explanation for this result is the very low agricultural productivity in the homeland areas associated with communal subsistence farming and land scarcity given their very high population densities (Percival and Homer-Dixon, 1998; Feynes and Meyer, 2003; Aliber and Hart, 2009).

Additional robustness checks

We estimate several alternative specifications to test the sensitivity of the regression results. To control for outliers, we exclude regions with income levels three deviations from the mean. The results for these regressions are reported in columns (1) to (3) of Table A1 in Appendix A. The coefficients change little, although the influence of scale economies (σ) and market access (as measured by $1/\sigma$) become marginally stronger. The homeland wage penalty remains significant and rising.

To test the sensitivity of the results to the share of the region's area falling within a former homeland for it to be classified as a homeland in the estimation, we raise the minimum area threshold from 30% to 70%. As shown in columns (4) to (6) of Table A1, the NEG coefficients are of similar size, but the estimated homeland wage penalty is lower in each year, while those of the bordering re-

[25] A simple regression of wages on the infrastructure variables using the pooled data reveals a significant positive association in each case.

[26] To test this, we regress the change in the infrastructure indicator from 1996 to 2011 on log wages in 1996 and the homeland dummy variable. With the exception of share households with refuse collection, in all the other cases (the infrastructure index and the individual infrastructure measures - the share households with electricity, piped water and flush toilet), we find significantly stronger improvements in infrastructure in low wage areas with even stronger increases in homeland areas.

gions rise such that they are no longer significantly different from the homeland areas in 2011. The border coefficient is most likely picking up the direct effects associated with the overlap in the area of the border regions and the former homelands, as well as indirect effects associated with their proximity to the former homelands. For example, with the 70% cut-off, the former homelands cover on average 15% of the area in the border regions.

Although we have included several regional controls, including for the skill composition of employment, and have instrumented to deal with some of the endogeneity problems, unobserved region-specific effects correlated with market potential may still bias our estimated coefficients. Consequently, we follow Hansen (2005) and estimate time-differenced versions of the wage equation for the periods 1996-2001 and 2001-2011. Table A2 in Appendix A presents the results using non-linear least squares, as we were unable to obtain estimates using GMM that satisfied the overidentifying restrictions. The results are consistent with those estimated in levels, with market potential coefficients of 0.27 to 0.32, and a 14.2% rise in the homeland deficit from 1996 to 2001 with no further change from 2001 to 2011.

Finally, we estimate regressions for skilled and unskilled labour separately, as in Hering and Poncet (2010) for China. They find greater wage sensitivity to market access for highly skilled workers and attribute this to the greater internal migration of low skilled workers. The results are reported in Table A3 in Appendix A. Like Hering and Poncet (2010), we find that skilled wages are more sensitive to market access. The homeland wage penalty is also substantially larger for skilled (27.2 to 36.8%) than unskilled (15.7 to 17.9%). This is unexpected, as unlike China, skilled labour tends to be more mobile in South Africa (Statistics South Africa, 2015).[27]

There are several potential explanations for this finding. Firstly, regional wage dispersion of unskilled workers is truncated by minimum wages set by the centralised sector Bargaining Councils or Sectoral Determinations.[28] In the former case, firms and unions negotiate wages that can be extended to non-participants through ministerial agreements, while in the latter case, the Ministry of Labour sets minimum wages for vulnerable sectors and occupations not covered by collective bargaining. These wage-setting processes have led to wage increases that are more binding for small firms (Magruder, 2012) and low skilled low wage workers (Bhorat et al., 2014; Dinkelman and Ranchhod, 2012). While wage gaps are permitted between regions (and industries and occupations), these tend to be based on quite crude rural-urban demarcations.[29]

[27] Using logit regressions and the 2011 Population Census data, Statistics South Africa (2015) find that a one level increase in education attained on a 28-point ordinal scale is associated with an increase of 2,9% in the odds of a person having migrated between 2006 and 2011.

[28] Wage setting and the conditions of service are determined by the 1995 Labour Relations Act and the Basic Conditions of Employment Act (BCEA) of 1997.

[29] For example, the new minimum wages in agriculture implemented in 2003, prescribed two separate wage levels for full-time farm workers: (R800) for those working within urbanized municipal areas classified as Area A (defined as areas where 1996 incomes were greater than Rands 24,000 per annum), and a lower wage (R650) in rural areas classified as Area B (areas

In some cases, e.g. the extension and enforcement of minimum wages agreed upon by the National Bargaining Council for the Clothing Manufacturing Industry to rural areas compressed regional wage variation for low skilled clothing workers, leading to substantial job losses in some towns such as Newcastle, a decentralised Growth Point under apartheid (Nattrass and Seekings, 2014). The effect of these labour regulations is a diminished impact of NEG forces on spatial wage dispersion of unskilled workers in South Africa.

A second reason is the very high unemployment rate amongst unskilled workers in South Africa. South Africa has one of the highest unemployment rates in the world, with rates in the first half of the 2000s that exceeded 30% for workers without a matric, compared to less than 8% for workers with a college degree (Banerjee et al., 2008). In this environment, labour demand shocks associated with changes in market potential are more likely to be reflected in changes in employment for low skilled workers, whereas for skilled workers that are scarce, the adjustment will take the form of wage changes.

Overall, our results show that the legacy of apartheid-era homeland policy continues to negatively influence wage levels for workers in former homelands as these areas remain isolated from the economy and the forces of economic geography despite 16 years of change since the ending of apartheid. These results support findings of a growing body of research that shows that distinctive historical events have long-lasting effects and continue to influence economic development to the present day (Banerjee and Iyer, 2005; Nunn, 2009; Dell, 2010; Angeles and Elizalde, 2017).

Spatial decay

The declining elasticity of substitution combined with the fall in the absolute value of the distance coefficient from 1996 to 2011 suggests rising importance of economies of scale and demand linkages in other locations in driving the spatial wage relationship in South Africa. To illustrate these changes in the spatial wage relationship, we follow Hanson (2005) and calculate the predicted change in wages in a region associated with a 10% increase in the region's market potential assuming the increase is concentrated at a specific point in space. We then plot how the local wage effect diminishes as the distance to the location of this shock increases. This involves plotting out the relationship $\Delta \ln(\hat{w}) = \frac{1}{\sigma}\ln(1 + 0.1 * e^{D\tau(1-\sigma)})$, where D is the distance from the region to the location of the shock.

Part (a) of Figure 55 plots out this relationship using the coefficient estimates taken from our strongest set of results in columns (7) to (9) of Table 2. Part (b) compares the decay relationship in South Africa with other countries for which similar studies have been conducted.

As illustrated in part (a) of Figure 5, the strength of local demand linkages has increased over the period. The 10% increase in market potential in the immediate vicinity of the region increased wages by 0.2% in 1996, 1.8% in 2001 and 3.1% in 2011. The role of distance can be seen in the decay of the wage

with 1996 incomes lower than R24, 000 per annum). In 2009, the Area A schedule was applied nationally (Bhorat et al., 2014). More recently (from January 2019), national minimum wages, with no variation across regions, have been implemented.

effect as distance from the location of the shock increases. In all years, the effect of the shock diminishes sharply as distance increases, but less so for 2011 than prior years, as is reflected by the rightward shift of the curve. However, even in the case of 2011, the effect of a shock to market potential is highly localized. By 10km, over 75% of the effect of the shock has dissipated resulting in only a 0.46% increase in local wages.

The highly localized effects of an increase in market potential are not unusual in the literature (Chang et al., 2015). In part (b) of Figure 5 we present the spatial decay of predicted wage changes from a 10% increase in market potential for other countries using NEG coefficients from those countries. We observe a more gradual spatial decay of predicted wage changes for the US and Spain, as well as Java in Indonesia, although the wage change from a shock in the immediate vicinity is much lower in this case. As in South Africa for 2011, Italy and China show a very sharp decay in predicted wages, with the curve for South Africa falling between these two. The spatial decay curve for Chile is low and flat, revealing the poor fit of the NEG model to the country, as argued by Paredes and Iturra (2012). South Africa is therefore not an outlier in the localisation of wage changes arising from shocks in market potential.

The very low spillovers in South Africa limit the extent to which the economic centres can drive wage growth in the outlying homeland areas. To illustrate this, we use the coefficients from columns (4) to (6) in Table 2 to simulate the wage effects arising from a 10% boost in income in the dominant regions of the economic centres of Gauteng, Cape Town, Bloemfontein, Port Elizabeth and Durban.[30] Together, the regions in these centres make up over half of South Africa's household income. The average effect across regions in the homelands and rest of South Africa are presented in Table A4 in Appendix A.

The impacts are very low, although they are stronger in 2011 than in 1996. For example, the 10% shock to income in the centres raises wages by on average 0.385% across regions outside of the homelands in 2011, but only 0.058% within the homelands. The effects in 1996 are a fraction of these outcomes. We also see a differential effect across homelands, with larger wage increases (but still extremely small) in homelands close to the economic centres (Bophuthatswana, KwaNdebele, Kwazulu) and zero impacts in other regions. The outlying homeland regions are delinked from the rest of the economy.

4 Conclusion

Despite the ending of apartheid, regional wage disparities remain prevalent in South Africa with the former homelands characterised by persistently low wages

[30] Following Moreno-Monroy (2011), the elasticity of wages in location j to income shocks in location i is calculated as $\frac{\partial lnW}{\partial lnY_i} = \frac{1}{\varepsilon} \frac{1}{\sum_s Y_s I_s^{\varepsilon-1} e^{\tau(\varepsilon-1)D_{rs}}} Y_i I_i^{\varepsilon-1} e^{\tau(\varepsilon-1)D_{rs}}$

$= \left[\frac{1}{\varepsilon} \frac{MA_{ir}}{MA_r}\right]$, i.e. it is proportional to the contribution of location i to location rs market access. We use this to calculate the combined effect on wages arising from of a 10% increase in incomes in the economic centres.

and incomes. In this paper, we use a new economic geography (NEG) model to estimate the extent to which the persistence in apartheid regional wage disparities are an outcome of economic forces such as access to markets. We estimate a structural wage equation derived directly from the NEG theory for 354 regions over the period 1996 to 2011.

We find support for the NEG model in explaining regional wage disparities across regions in South Africa. While demand linkages have become stronger over time, the effects of increased market potential remain highly localised in view of high distance coefficients. We also find a wage deficit in homeland areas even after controlling for NEG and other region-specific characteristics. Average wages of workers in homeland areas were 11.8% lower than predicted in 1996, with this gap rising to over 13.3% by 2011. One reason is that the homeland areas have remained isolated from the rising demand linkages that are influencing the spatial distribution of wages in South Africa. Our results are robust to several sensitivity tests carried out to check for potential bias due to the way we measure homeland status as well as due to omitted variables.

Our findings show that the reintegration of homeland areas into South Africa and the implementation of regional policies since the end of apartheid have not been enough to reduce the homeland wage penalty. These will continue to persist without regional policy initiatives aimed at improving the underlying conditions within these regions as well as their lack of connectivity to the urban centres. A key target of such policies is the very poor infrastructure within these regions and the weak transport linkages of some to the urban centres. Wage policies that sharply narrow the urban-rural wage gap are also detrimental to the entry and survival of businesses in outlying regions and can prevent homeland areas from attracting new businesses based on their lower wages. Labour policies that facilitate the migration of workers from the homeland areas to the higher wage urban areas can assist in overcoming the homeland wage penalty.

4.1 References

References

[1] Abel, M. (2019). Long-run effects of forced resettlement: evidence from Apartheid South Africa. *The Journal of Economic History*, *79*(4), 915–953.

[2] Aghion, P., Braun, M., & Fedderke, J. (2008). Competition and productivity growth in South Africa. *Economics of Transition*, *16*(4), 741–768.

[3] Aliber, M., & Hart, T. G. (2009). Should subsistence agriculture be supported as a strategy to address rural food insecurity? *Agrekon*, *48*(4), 434–458.

[4] Allison, P. D. (2001). *Missing Data* (Vol. 136). SAGE Publications.

[5] Amiti, M., & Cameron, L. (2007). Economic geography and wages. *The Review of Economics and Statistics*, *89*(1), 15–29.

[6] Angeles, L., & Elizalde, A. (2017). Pre-colonial institutions and socioeconomic development: The case of Latin America. *Journal of Development Economics*, *124*, 22–40.

[7] Ardington, C., Lam, D., Leibbrandt, M., & Welch, M. (2006). The sensitivity to key data imputations of recent estimates of income poverty and inequality in South Africa. *Economic Modelling*, *23*(5), 822–835.

[8] Banerjee, A., Galiani, S., Levinsohn, J., McLaren, Z., & Woolard, I. (2008). Why has unemployment risen in the new South Africa? 1. *Economics of Transition*, *16*(4), 715–740.

[9] Banerjee, A., & Iyer, L. (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in India. *American Economic Review*, *95*(4), 1190–1213.

[10] Bastos, P., & Bottan, N. L. (2016). Resource Rents, Coercion, and Local Development: Evidence from Post-Apartheid South Africa. *World Bank Policy Research Working Paper*, (7572), The World Bank, Washington, DC.

[11] Bell, T. (1997). South African regional industrial development policy: critical issues. *Transformation*, (32).

[12] Bhorat, H., Kanbur, R., & Stanwix, B. (2014). Estimating the impact of minimum wages on employment, wages, and non-wage benefits: the case of agriculture in South Africa. *American Journal of Agricultural Economics*, *96*(5), 1402–1419.

[13] Bosker, M, & Garretsen, H. (2012). Economic geography and economic development in Sub-Saharan Africa. *The World Bank Economic Review*, *26*(3), 443–485.

[14] Bosker, M, & Krugell, W. (2008). Regional income evolution in South Africa after apartheid. *Journal of Regional Science*, *48*(3), 493–523.

[15] Bosker, Maarten, & Garretsen, H. (2010). Trade costs in empirical New Economic Geography. *Papers in Regional Science*, *89*(3), 485–511.

[16] Brakman, S., Garretsen, H., & Schramm, M. (2004). The spatial distribution of wages: estimating the Helpman-Hanson model for Germany. *Journal of Regional Science*, *44*(3), 437–466.

[17] Breinlich, H., Ottaviano, G., & Temple, J. (2014). Regional Growth and Regional Decline. *Elsevier*, *2*, 683–779.

[18] Chang, H.-H., van Marrewijk, C., & Schramm, M. (2015). 19 Empirical studies in geographical economics. In *Charlie Karlsson, Martin Andersson and Therese Norman (eds.) Handbook of Research Methods and Applications in Economic Geography, Edward-Elgar Publishing.*

[19] Clark, N. L., & Worger, W. H. (2011). *South Africa: The rise and fall of apartheid, 2nd Edition.* Routledge.

[20] Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, *6*(4), 330.

[21] Combes, P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, *63*(2), 723–742.

[22] De Bruyne, K. (2010). Explaining the Location of Economic Activity. Is there a Spatial Employment Structure in Belgium? *International Journal of Economic Issues*, *3*(2), 199–222.

[23] Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica*, *78*(6), 1863–1903.

[24] Dinkelman, T, & Ranchhod, V. (2012). Evidence on the impact of minimum wage laws in an informal sector: Domestic workers in South Africa. *Journal of Development Economics*, *99*(1), 27–45.

[25] Dinkelman, T. (2011). The Effects of Rural Electrification on Employment: New Evidence from South Africa. *American Economic Review*, *101*(7), 3078–3108.

[26] Edwards, L. (2005). Has South Africa liberalised its trade? *South African Journal of Economics*, *73*(4), 754–775.

[27] Fallah, B. N., Partridge, M. D., & Olfert, M. R. (2011). New economic geography and US metropolitan wage inequality. *Journal of Economic Geography*, *11*(5), 865–895.

[28] Fally, T., Paillacar, R., & Terra, C. (2010). Economic geography and wages in Brazil: Evidence from micro-data. *Journal of Development Economics*, *91*(1), 155–168.

[29] Fenyes, T., & Meyer, N. (2003). Structure and production in South African agriculture. *In: L. Nieuwoudt and J. Groenewald (Eds.) the Challenges of Change: Agriculture, Land and the South African Economy.* Pietermaritzburg, South Africa: The University of Natal Press.

[30] Fujita, M., Krugman, P., & Venables, A. (1999). *The spatial economy: cities, regions and international trade.* Cambridge, Massachusetts London, England: MIT Press.

[31] Gelb, S. (2004). An overview of the South African economy. *State of the Nation: South Africa*, *2005*, 367–400.

[32] Hanson, G. (2005). Market potential, increasing returns and geographic concentration. *Journal of International Economics*, *67*(1), 1–24.

[33] Hanson, G. H., & Xiang, C. (2004). The Home-Market Effect and Bilateral Trade Patterns. *American Economic Review*, *94*(4), 1108–1129.

[34] Harris, C. C. D. (1954). The Market as a Factor in the Localization of Industry in the United States. *Annals of the Association of American Geographers*, *44*(4), 315–348.

[35] Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. *International Journal of Climatology*, *34*(3), 623–642.

[36] Head, K., & Ries, J. (2001). Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of U.S.–Canada Trade. *American Economic Review*, *91*(4), 858–876.

[37] Helpman, E. (1998). The size of regions. *Topics in Public Economics: Theoretical and Applied*, 33–54.

[38] Hering, L., & Poncet, S. (2009). The impact of economic geography on wages: Disentangling the channels of influence. *China Economic Review*, *20*(1), 1–14.

[39] Hering, L., & Poncet, S. (2010). Market access and individual wages: Evidence from China. *The Review of Economics and Statistics*, *92*(1), 145–159.

[40] Kanbur, R., & Venables, A. J. (2005). *Rising spatial disparities and development: Why do they matter?. Oxford: Oxford University Press.*

[41] King, B. H., & McCusker, B. (2007). Environment and development in the former South African bantustans. *The Geographical Journal*, *173*(1), 6–12.

[42] Kingdon, G., & Knight, J. (2006). How flexible are wages in response to local unemployment in South Africa? *Industrial and Labour Relations Review*, *59*(3), 471–495.

[43] Kosfeld, R., & Eckey, H. (2010). Market access, regional price level and wage disparities: the German case. *Jahrbuch Für Regionalwissenschaft*, *30*(2), 105–128.

[44] Krugell, W., & Rankin, N. (2012). Agglomeration and Firm-Level Efficiency in South Africa. *Urban Forum*, *23*(3), 299–318.

[45] Krugman, P. (1991). Increasing Returns and Economic Geography. *The Journal of Political Economy*, *99*(3), 483–499.

[46] Krugman, P., & Venables, A. J. (1995). Globalization and the Inequality of Nations. *The Quarterly Journal of Economics*, *110*(4), 857–880.

[47] Lahiff, E., Borras, S. M., & Kay, C. (2007). Market-led agrarian reform: policies, performance and prospects. *Third World Quarterly*, *28*(8), 1417–1436.

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

[49] Leibbrandt, M., & Woolard, I. (2001). The labour market and household income inequality in South Africa: existing evidence and new panel data. *Journal of International Development*, *13*(6), 671–689.

[50] Leibbrandt, M., Woolard, I., Finn, A., & Argent, J. (2010). Trends in South African Income Distribution and Poverty since the Fall of Apartheid. *OECD Social, Employment, and Migration Working Papers, No. 101, OECD Publishing.*

[51] Luüs, C. (2005). The Absa residential property market database for South Africa-key data trends and implications. *Bank for International Settlements*, *21*, 149–170.

[52] Maantay, J. A., Maroko, A. R., & Porter-Morgan, H. (2008). Research Note—A New Method for Mapping Population and Understanding the Spatial Dynamics of Disease in Urban Areas: Asthma in the Bronx, New York. *Urban Geography*, *29*(7), 724–738.

[53] Magruder, J. (2012). High Unemployment Yet Few Small Firms: The Role of Centralized Bargaining in South Africa. *American Economic Journal: Applied Economics*, *4*(3), 138–166.

[54] Mion, G. (2004). Spatial externalities and empirical analysis: the case of Italy. *Journal of Urban Economics*, *56*(1), 97–118.

[55] Moreno-Monroy, A. (2011). Market access and the heterogeneous effect of shocks on wages: evidence from Chinese cities. *Papers in Regional Science*, *90*(1), 9–25.

[56] Municipal IQ. (2019). *3 key trends from 2018's all-time service delivery protest record – for immediate release. Report published online, 16th January 2019 https://www.municipaliq.co.za/index.php [Accessed 15 November 2019].*

[57] Nattrass, N., & Seekings, J. (2014). Job destruction in Newcastle: minimum wage-setting and low-wage employment in the South African clothing industry. *Transformation: Critical Perspectives on Southern Africa*, *84*(1), 1–30.

[58] Naudé, W, & Gries, T. (2009). Explaining regional export performance in a developing country: The role of geography and relative factor endowments. *Regional Studies*, *43*(7), 967–979.

[59] Naudé, WA, & Krugell, W. (2006). Sub-national growth rate differentials in South Africa: an econometric analysis. *Papers in Regional Science*, *85*(3), 443–457.

[60] Nel, E, & Rogerson, C. (2009). Re-thinking spatial inequalities in South Africa: Lessons from international experience. *Urban Forum*, *20*(2), 141–155.

[61] Nel, Etienne. (1994). Regional Development in South Africa: From Apartheid Planning to the Reform Era. *Geography Research Forum*, *14*, 13–29.

[62] Niebuhr, A. (2006). Market access and regional disparities. *The Annals of Regional Science*, *40*(2), 313–334.

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

[64] Nunn, N. (2009). The importance of history for economic development. *Annual Review of Economics*, *1*(1), 65–92.

[65] Nunn, N., & Puga, D. (2012). Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics*, *94*(1), 20–36.

[66] Paredes, D. (2015). Can NEG explain the spatial distribution of wages of Chile? *Tijdschrift Voor Economische En Sociale Geografie*, *106*(1), 65–77.

[67] Paredes, D., & Iturra, V. (2012). Market access and wages: A spatially heterogeneous approach. *Economics Letters*, *116*(3), 349–353.

[68] Percival, V., & Homer-Dixon, T. (1998). Environmental scarcity and violent conflict: the case of South Africa. *Journal of Peace Research*, *35*(3), 279–298.

[69] Pires, A. (2006). Estimating Krugman's economic geography model for the Spanish regions. *Spanish Economic Review*, *6*(2), 83–112.

[70] Redding, S. J. (2013). Economic Geography: A review of the theoretical and empirical literature. *In Palgrave Handbook of International Trade. Palgrave Macmillan UK*, 497–531.

[71] Redding, S., & Venables, A. J. (2004). Economic geography and international inequality. *Journal of International Economics*, *62*(1), 53–82.

[72] Reed, H. E. (2013). Moving across boundaries: migration in South Africa, 1950–2000. *Demography*, *50*(1), 71–95.

[73] Riley, S. J., DeGloria, S. D., & Elliot, R. (1999). Index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences*, *5*(1–4), 23–27.

[74] Roth, P. L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology*, *47*(3), 537–560.

[75] Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys. *New York: Wiley.*

[76] Simkins, C. (2011). *"The Evolution of the South African Population in the Twentieth Century." In The Cambridge History of South Africa. Volume 2: 1885–1994, edited by Robert Ross, Anne Kelk Mager, and Bill Nasson, 492–517. Cambridge: Cambridge University Press.*

[77] Statistics South Africa. (2000). *Measuring poverty in South Africa. Pretoria: Statistics South Africa.*

[78] Statistics South Africa. (2015). *Census 2011: Migration dynamics in South Africa, Report No. 03-01-79. Pretoria: Statistics South Africa.*

[79] Todes, A., & Turok, I. (2018). Spatial inequalities and policies in South Africa: Place-based or people-centred? *Progress in Planning*, *123*, 1–31.

[80] Turok, I. (2012). *Urbanisation and Development in South Africa: Economic Imperatives, Spatial Distortions and Strategic Responses. Urbanisation and Emerging Population Issues. Working Paper 8, IIED, London.*

[81] Venables, A. (1996). Equilibrium locations of vertically linked industries. *International Economic Review*, *37*(2), 341–359.

[82] Venables, A. (2005). Spatial disparities in developing countries: cities, regions, and international trade. *Journal of Economic Geography*, *5*(1), 3–21.

[83] von Fintel, D. (2007). Dealing with Earnings Bracket Responses in Household Surveys–How Sharp are Midpoint Imputations? *South African Journal of Economics*, *75*(2), 293-312.

[84] von Fintel, D. (2018). Long-run spatial inequality in South Africa: early settlement patterns and separate development. *Studies in Economics and Econometrics*, *42*(2), 81–102.

[85] Wilson, F. (2001). Minerals and migrants: how the mining industry has shaped South Africa. *Daedalus*, *130*(1), 99–121.

[86] Wittenberg, M. (2017a). Wages and Wage Inequality in South Africa 1994-2011: Part 2 - Inequality Measurement and Trends. *South African Journal of Economics*, *85*(2), 298–318.

[87] Wittenberg, M. (2017b). Wages and Wage Inequality in South Africa 1994–2011: Part 1–Wage Measurement and Trends. *South African Journal of Economics*, *85*(2), 279–297.

[88] World Bank. (2009). Reshaping Economic Geography: World development report. *The World Bank, Washington DC.*

[89] World Bank. (2018a). *Overcoming poverty and inequality in South Africa: An assessment of drivers, constraints and opportunities.* The World Bank, Washington DC.

[90] World Bank. (2018b). *Republic of South Africa-Systematic country diagnostic: An incomplete transition-Overcoming the legacy of exclusion in South Africa.* The World Bank, Washington DC.

[91] Zenker, O., & Jensen, S. (2015). Homelands as Frontiers: Apartheid's Loose Ends–An Introduction. *Journal of Southern African Studies*, *41*(5), 937–9

Table 1: Summary statistics of homeland areas and the rest of the country.

	Average wages Rand/month	LFPR %	UR %	ER %	SW %	PD Pop/km^2	Electricity %	Refuse %	Distance km^2	Rainfall mm
1996										
Rest of country	1613.8	56.4%	27.5%	41.7%	8.8	20.6	64.0%	61.7%	630.2	54.6
Homeland	1385.3	36.3%	60.0%	15.3%	13.9	94.4	26.4%	10.3%	602.8	75.3
2011										
Rest of country	2772.5	55.2%	27.4%	40.6%	14.0	25.9	85.2%	68.6%	630.2	49.7
Homeland	2234.0	35.0%	44.7%	19.6%	18.4	98.4	74.5%	15.7%	602.8	64.1

Source: Own calculations using South African Population Census data for 1996 and 2011.
Notes: Homeland is defined as a magisterial district with at least 30% of its area covered by a former homeland.
LFPR is labour force participation rate, which equals the number of participants in the labour force divided by the working-age population. ER is employment rate calculated as the number of employed divided by the working-age population. UR is unemployment rate calculated as the number of unemployed divided by the labour force. SW is the share of workers shares (in total working-age population) with tertiary education. PD is population density calculated as the number of people per km^2. The exponent of the mean of the log population density is presented given skewed distribution across regions. Electricity is the share of households with electricity for lighting and Refuse is the share of households with refuse collection. Rainfall is the average annual rainfall measured in millimetres. Distance is the average distance to the economic centre (Johannesburg). Average wages is average real wages per month and we proxy regional average wages with average incomes per employed individuals aged 15-64 years. Wages are converted to 1996 value equivalents using the national consumer price index (CPI) provided by Statistics South Africa.

Table 2: Estimates of the market potential function

	Baseline			+ Border			+ TBVC			Pooled
	1996	2001	2011	1996	2001	2011	1996	2001	2011	1996-2011
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Constant	-0.168	-1.095*	-1.284*	-0.129	-1.109*	-1.852*	-0.138	-1.022*	-1.754*	-1.026**
	(0.353)	(0.557)	(0.557)	(0.350)	(0.464)	(0.842)	(0.331)	(0.457)	(0.782)	(0.288)
Sigma	63.109	5.407+	4.370*	115.978	5.266*	3.087*	88.06	5.712*	3.271*	5.612**
	(262.110)	(2.790)	(1.874)	(879.347)	(2.261)	(1.413)	(486.671)	(2.613)	(1.472)	(1.603)
tau	0.046*	0.136	0.102+	0.034+	0.091*	0.118	0.033+	0.088*	0.107+	0.070**
	(0.020)	(0.093)	(0.052)	(0.017)	(0.044)	(0.074)	(0.018)	(0.043)	(0.062)	(0.017)
Homeland	-0.118*	-0.243+	-0.133*	-0.125**	-0.249**	-0.223*	-0.139**	-0.250**	-0.275**	-0.206**
	(0.055)	(0.125)	(0.058)	(0.046)	(0.088)	(0.088)	(0.050)	(0.089)	(0.097)	(0.053)
Skill share	0.632*	0.602+	0.543	0.748*	0.713*	0.851*	0.682+	0.634*	0.799*	0.646**
	(0.288)	(0.310)	(0.338)	(0.329)	(0.309)	(0.404)	(0.373)	(0.295)	(0.389)	(0.162)
Unemployment rate	0.057	-0.02	-0.297+	0.038	0.038	-0.076	0.038	0.031	-0.054	-0.054
	(0.064)	(0.092)	(0.155)	(0.052)	(0.123)	(0.273)	(0.054)	(0.121)	(0.248)	(0.048)
Share mining	0.032	-0.025	-0.187	0.054	-0.028	-0.248	0.043	-0.021	-0.22	0.022
	(0.050)	(0.206)	(0.271)	(0.061)	(0.081)	(0.153)	(0.055)	(0.080)	(0.145)	(0.046)
Border region				-0.062*	-0.090*	-0.123*	-0.061*	-0.087*	-0.115*	-0.086**
				(0.030)	(0.041)	(0.057)	(0.031)	(0.040)	(0.055)	(0.022)
TBVC							0.04	0.037	0.109+	0.060*
							(0.039)	(0.057)	(0.063)	(0.029)
D2001										-0.007
										(0.015)
D2011										-0.042**
										(0.015)
D2001 x Homeland										-0.031
										(0.044)
D2011 x Homeland										0.007
										(0.046)
Market access	0.016	0.185+	0.229*	0.009	0.190*	0.324*	0.011	0.175	0.306	0.178**
	(0.066)	(0.095)	(0.098)	(0.065)	(0.082)	(0.128)	(0.063)	(0.080)	(0.138)	(0.051)
Distance	-2.857	-0.599*	-0.344	-3.907	-0.388*	-0.246+	-2.873	-0.415	-0.243	-0.323**
	(11.313)	(0.287)	(0.219)	(29.037)	(0.185)	(0.128)	(15.352)	(0.206)	(0.119)	(0.075)
Markup	1.016**	1.227**	1.297**	1.009**	1.234**	1.479**	1.011	1.212	1.440	1.217**
	(0.068)	(0.144)	(0.165)	(0.067)	(0.124)	(0.325)	(0.064)	(0.118)	(0.286)	(0.075)
Observations	352	354	354	352	354	354	352	354	354	1062
Hansen J-stat	0.266	0.476	0.096	0.13	2.304	0.44	0.327	3.013	0.758	7.792
Hansen p-value	0.601	0.7883	0.953	0.719	0.316	0.803	0.567	0.222	0.685	0.051

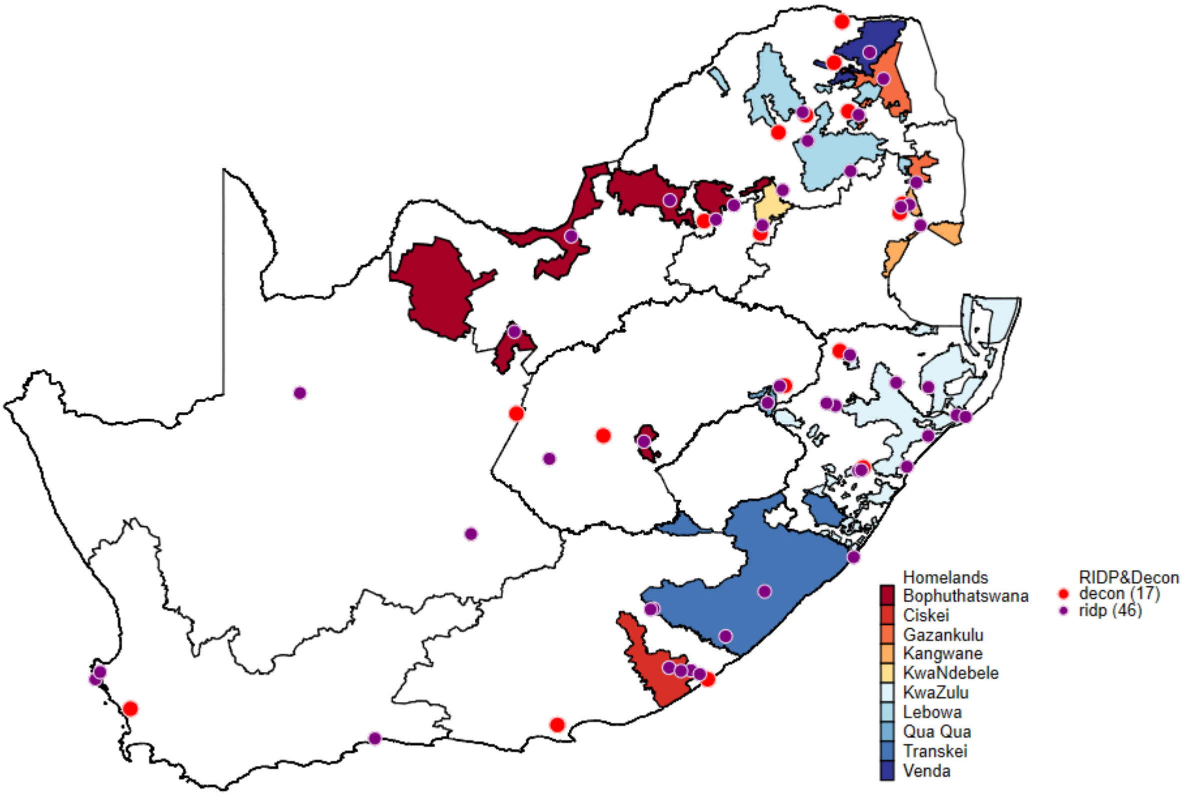
Notes: Heteroskedasticity-consistent standard errors are in parentheses: ** p<0.01, * p<0.05, + p<0.1. The 2001 and 2011 estimates use as instruments 1996 income and wages in the own region, 1996 income of the closest bordering region and a terrain ruggedness index obtained from Nunn & Puga (2012), 1996 income and population in own region and wages and 1996 income and wages in closest border as instruments. In addition to the terrain ruggedness index, the 1996 estimates use as instruments 1991 population and 1995 income per worker. Alternative sets of instruments are used, but these tend not to satisfy the Hansen J-test. The Hansen J statistic (p-value) is for a test of overidentifying restrictions on the instruments.

Table 3: Sensitivity of homeland effect to controls for decentralisation growth points, infrastructure and amenities

	Base	RIDP	Infr	Elect	Refuse	Toilet	Water	Public	Agriculture	Manufact	Tempera	Rainfall
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	-1.117**	-1.241**	-0.989**	-0.899**	-1.114**	-1.207**	-0.956**	-1.360**	-0.975**	-1.519*	-0.135	-1.931**
	(0.302)	(0.376)	(0.327)	(0.293)	(0.364)	(0.420)	(0.307)	(0.451)	(0.266)	(0.672)	(0.233)	(0.590)
Sigma	5.167**	4.666**	5.802**	5.952**	5.222**	4.854**	5.627**	4.092**	6.068**	3.790*	7.644**	3.650**
	(1.426)	(1.445)	(1.919)	(1.777)	(1.697)	(1.691)	(1.653)	(1.388)	(1.829)	(1.692)	(2.406)	(1.151)
tau	0.075**	0.072**	0.071**	0.082**	0.070**	0.074**	0.078**	0.103*	0.071**	0.103+	0.071**	0.128*
	(0.020)	(0.018)	(0.018)	(0.022)	(0.019)	(0.023)	(0.020)	(0.048)	(0.017)	(0.058)	(0.015)	(0.063)
Homeland	-0.191**	-0.213**	-0.190**	-0.212**	-0.174**	-0.179**	-0.209**	-0.218**	-0.174**	-0.235**	-0.127**	-0.254**
	(0.054)	(0.059)	(0.050)	(0.059)	(0.049)	(0.054)	(0.055)	(0.077)	(0.048)	(0.091)	(0.045)	(0.092)
Skill share	0.712**	0.778**	0.725**	0.643**	0.721**	0.701**	0.668**	0.821**	0.725**	0.708**	0.521**	0.662**
	(0.168)	(0.177)	(0.163)	(0.161)	(0.174)	(0.183)	(0.157)	(0.202)	(0.176)	(0.175)	(0.146)	(0.177)
Unemployment rate	-0.053	-0.057	-0.077	-0.076	-0.062	-0.045	-0.096	0.014	-0.029	-0.063	-0.044	-0.193**
	(0.049)	(0.049)	(0.057)	(0.056)	(0.049)	(0.058)	(0.059)	(0.076)	(0.059)	(0.052)	(0.050)	(0.066)
Share mining	0.022	0.035	0.009	-0.007	0.02	0.02	-0.003	-0.045	0.042	0.036	0.046	-0.085
	(0.048)	(0.048)	(0.050)	(0.056)	(0.047)	(0.048)	(0.053)	(0.070)	(0.049)	(0.059)	(0.046)	(0.082)
Border region	-0.090**	-0.098**	-0.094**	-0.091**	-0.089**	-0.088**	-0.098**	-0.101**	-0.085**	-0.100**	-0.051**	-0.149**
	(0.022)	(0.023)	(0.022)	(0.021)	(0.022)	(0.022)	(0.022)	(0.028)	(0.020)	(0.028)	(0.016)	(0.041)
D2001	-0.006	-0.008	-0.003	0.005	-0.007	-0.009	-0.002	-0.01	-0.01	-0.001	-0.005	-0.005
	(0.016)	(0.016)	(0.016)	(0.017)	(0.016)	(0.016)	(0.015)	(0.017)	(0.015)	(0.020)	(0.014)	(0.018)
D2011	-0.045**	-0.049**	-0.038*	-0.022	-0.047**	-0.052**	-0.035*	-0.041*	-0.040*	-0.045**	-0.028*	-0.042*
	(0.015)	(0.016)	(0.017)	(0.018)	(0.015)	(0.019)	(0.015)	(0.017)	(0.020)	(0.017)	(0.014)	(0.018)
D2001xHomeland	-0.035	-0.037	-0.038	-0.023	-0.037	-0.035	-0.035	-0.036	-0.035	-0.037	-0.016	-0.039
	(0.046)	(0.043)	(0.044)	(0.049)	(0.044)	(0.045)	(0.048)	(0.056)	(0.045)	(0.055)	(0.046)	(0.066)
D2011xHomeland	0.007	0.009	-0.002	0.021	0.004	0.011	-0.004	0.009	0.000	0.019	0.007	0.005
	(0.049)	(0.046)	(0.048)	(0.052)	(0.047)	(0.052)	(0.052)	(0.060)	(0.046)	(0.061)	(0.048)	(0.070)
Control		0.045+	-0.006	-0.093*	0.016	0.034	-0.073*	-0.206	0.038	0.153	-0.222**	0.110**
		(0.025)	(0.007)	(0.042)	(0.036)	(0.041)	(0.037)	(0.132)	(0.055)	(0.172)	(0.052)	(0.036)
Market access	0.194**	0.214**	0.172**	0.168**	0.191**	0.206**	0.178	0.244**	0.165**	0.264*	0.131**	0.274**
	(0.053)	(0.066)	(0.057)	(0.050)	(0.062)	(0.072)	(0.052)	(0.083)	(0.050)	(0.118)	(0.041)	(0.086)
Distance	-0.315**	-0.265**	-0.341**	-0.406**	-0.296**	-0.285**	-0.361	-0.318**	-0.360**	-0.287**	-0.472**	-0.339**
	(0.072)	(0.061)	(0.086)	(0.120)	(0.068)	(0.063)	(0.091)	(0.079)	(0.095)	(0.062)	(0.145)	(0.090)
Markup	1.240**	1.273**	1.208**	1.202**	1.237**	1.259**	1.216	1.323**	1.197**	1.358**	1.151**	1.377**
	(0.082)	(0.107)	(0.083)	(0.072)	(0.095)	(0.113)	(0.077)	(0.145)	(0.071)	(0.217)	(0.054)	(0.164)
Observations	1062	1062	1062	1062	1062	1062	1062	1062	1062	1062	1062	1062
Hansen J-stat	5.051	4.519	4.896	3.147	6.628	6.397	3.59	3.344	5.748	3.278	8.649	0.148
Hansen p-value	0.168	0.211	0.180	0.370	0.085	0.094	0.309	0.342	0.125	0.351	0.034	0.985

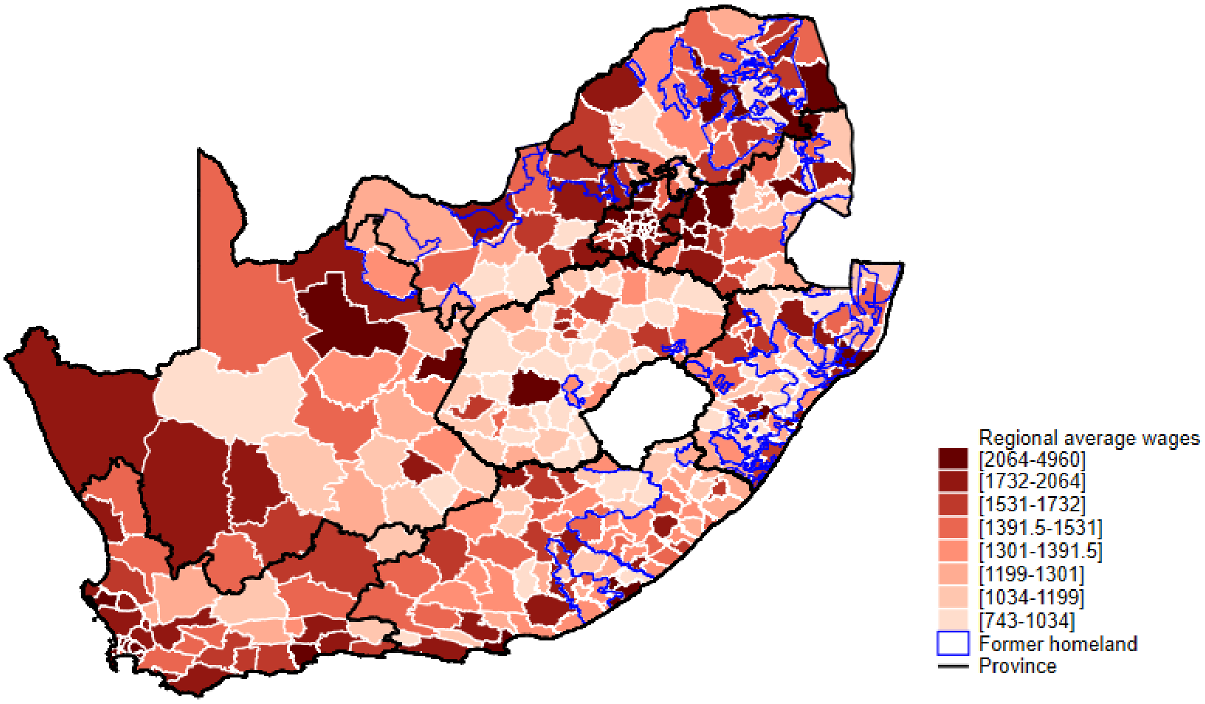
Notes: Heteroskedasticity-consistent standard errors are in parentheses: ** $p<0.01$, * $p<0.05$, + $p<0.1$. RIDP is a dummy variable for regions that contains a designated RIDP zone. Infr is an infrastructure index based on the principal component of the variables share household with access to electricity, piped water, flush toilet and refuse removal. Elect, Refuse, Toilet and Water refer to the share of households in a region that have access to electricity, piped water, flush toilet and refuse removal facilities, Public, Agriculture and Manufact denote the share of total employment accounted for by Public, Agriculture and Manufacturing sectors, Tempera measures the average annual temperature in the region and Rainfall measures the average annual rainfall in the region. Based on tests of overidentifying restrictions, we reject the hypothesis that the instruments are uncorrelated with the errors in columns (6), (7) and (11).

Figure 1: Location of former homelands areas.



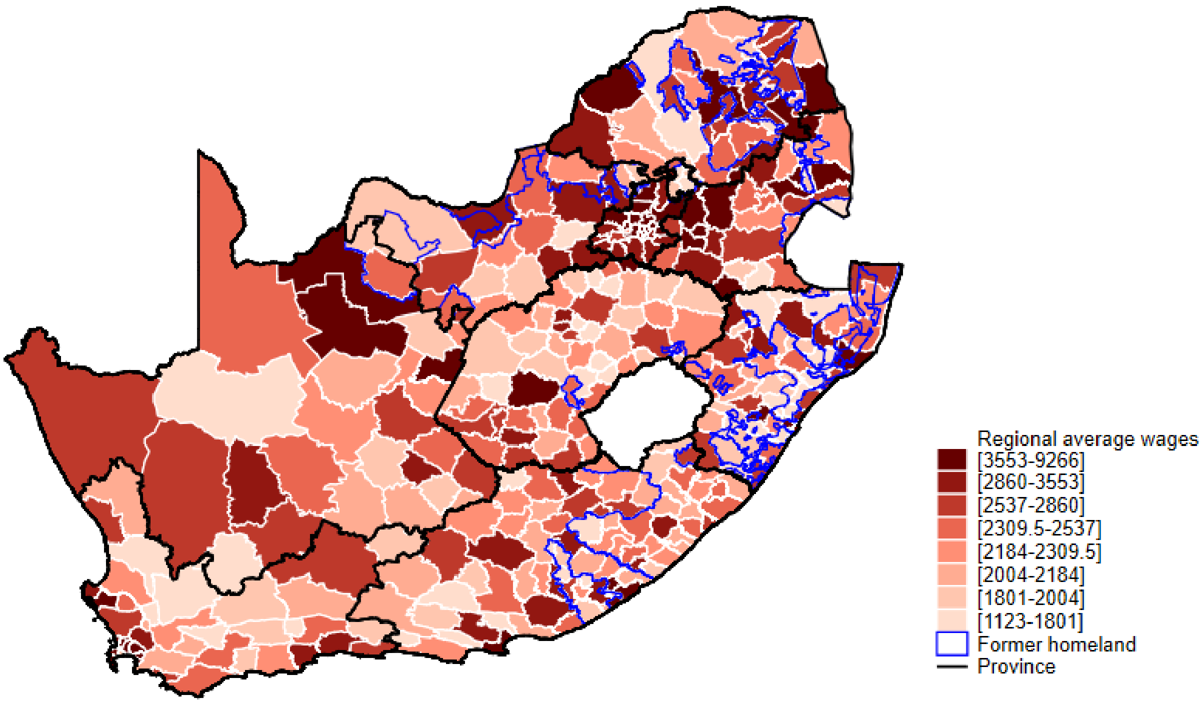
Notes: Authors' construction using GIS shapefiles obtained from Statistics South Africa and data on industrial zones from Kerby (2016). The legend ridp denotes Growth Points/Industrial Development Points and decon denotes Deconcentration Point.

Figure 2: Spatial distribution of monthly average wages across regions in 1996.



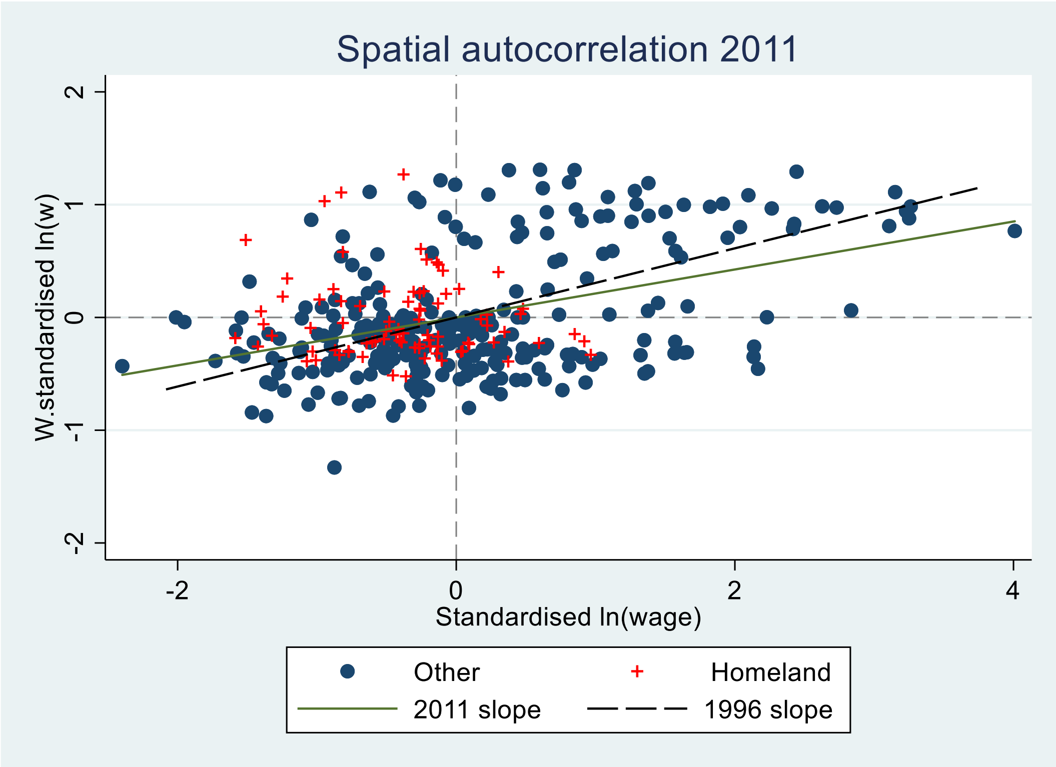
Notes: Author's calculations based on 1996 census data aggregated to 354 magisterial districts. We proxy regional wages with average incomes per employed individuals aged 15-64 years.

Figure 3: Spatial distribution of monthly average wages across regions in 2011.



Notes: Author's calculations based on 2011 census data aggregated to 354 magisterial districts. We proxy regional wages with average incomes per employed individuals aged 15-64 years. Wages are converted to 1996 value equivalents using the national consumer price index (CPI) provided by Statistics South Africa.

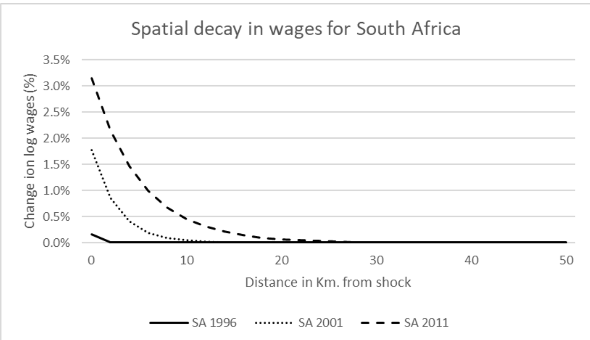
Figure 4: Moran scatter plot of regional wages, 2011



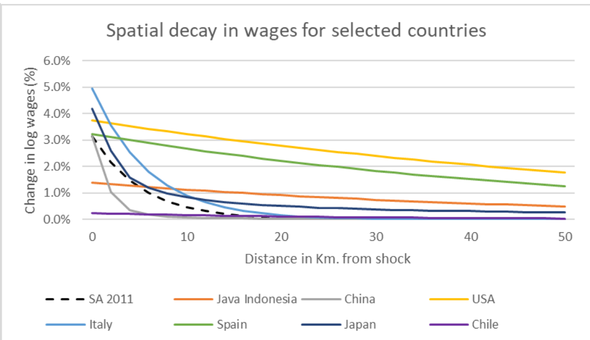
Notes: Own calculations using Population Census data for 2011 and 1996. W.standardised ln(w) reflects the simple average standardized wage in surrounding regions within a 100km distance.

Figure 5: Spatial decay of wage changes

(a) South Africa



(b) Cross-country comparison



Notes: Coefficient estimates for the comparator countries are obtained from the following sources: Table 3, column 6 from Amiti and Cameron (2007) for Java in Indonesia, Table 1, column 3 from Moreno-Monroy (2011) for China, Table 4, 1980-90 from Hanson (2005) for USA, Table 3, column 1 from Mion (2004) for Italy, Table 11, 1995-1988 from Pires (2006) for Spain and Table 5 from Paredes (2015) for Chile. Trade costs are estimated using a power distance function for Japan. All other studies use exponential distance functions.

Appendix A: Additional Tables

Table A1: Sensitivity of results to exclusion of outliers and more restrictive homeland status (70% area threshold)

	Excluding outliers			70% homeland share	
	1996	2001	2011	2001	2011
	(1)	(2)	(3)	(4)	(5)
Constant	-0.169	-1.080*	-1.916*	-0.857+	-1.748*
	(0.384)	(0.456)	(0.819)	(0.438)	(0.845)
Sigma	61.049	5.381*	3.033*	6.353*	3.136*
	(279.399)	(2.326)	(1.325)	(3.121)	(1.479)
tau	0.032+	0.086*	0.097+	0.081**	0.125
	(0.018)	(0.038)	(0.058)	(0.032)	(0.085)
Homeland	-0.138**	-0.244**	-0.288**	-0.167*	-0.230*
	(0.049)	(0.087)	(0.096)	(0.077)	(0.097)
Skill share	0.682+	0.603*	1.082*	0.504	0.637
	(0.414)	(0.290)	(0.473)	(0.359)	(0.531)
Unemployment rate	0.041	0.011	0.023	-0.011	-0.11
	(0.048)	(0.117)	(0.234)	(0.121)	(0.332)
Share mining	0.047	-0.019	-0.268+	-0.165*	-0.428+
	(0.050)	(0.072)	(0.152)	(0.080)	(0.219)
TBVC	0.039	0.03	0.097	0.063	0.147*
	(0.038)	(0.055)	(0.059)	(0.051)	(0.072)
Border region	-0.062*	-0.096*	-0.133*	-0.141**	-0.161*
	(0.031)	(0.042)	(0.059)	(0.048)	(0.071)
Market access	0.016	0.186*	0.330*	0.157*	0.319*
	(0.075)	(0.080)	(0.144)	(0.077)	(0.150)
Distance	-1.921	-0.377*	-0.197*	-0.434+	-0.267
	(8.119)	(0.171)	(0.082)	(0.250)	(0.178)
Markup	1.017**	1.228**	1.492**	1.187**	1.468**
	(0.077)	(0.121)	(0.320)	(0.109)	(0.324)
Observations	348	351	348	354	354
Hansen J	0.178	2.862	0.77	2.297	0.561
Hansen p-value	0.673	0.239	0.680	0.317	0.755

Notes: Heteroskedasticity-consistent standard errors are in parentheses: ** p<0.01, * p<0.05, + p<0.1. Regions with log wages more than 3 standard deviations from the mean are excluded in the regressions for columns (1) to (3). No results are presented for 1996 when using the 70% homeland share threshold (columns 4 and 5) as the GMM estimations struggled to converge.

Table A2: Nonlinear least-squares (NLS) Estimates for Time-difference model

	Time Diff 2001 - 1996	Time Diff 2011 - 2001
Constant	0.005	0.034
	(0.019)	(0.027)
Sigma	3.127**	3.748**
	(0.439)	(0.653)
tau	0.063**	0.096*
	(0.016)	(0.046)
Homeland	0.142**	0.026
	(0.027)	(0.030)
Skill share	2.807**	1.765**
	(0.356)	(0.322)
Unemployment rate	-0.287+	-0.027
	(0.150)	(0.182)
Share mining	-0.096+	0.791**
	(0.055)	(0.147)
Border region	0.064*	0.046
	(0.026)	(0.032)
TBVC	-0.038	0.094**
	(0.030)	(0.034)
Market access	0.320**	0.267**
	(0.045)	(0.046)
Distance	-0.134**	-0.265+
	(0.044)	(0.141)
Markup	1.470**	1.364**
	(0.097)	(0.086)
Obs	354	354
Log-likelihood	164.536	100.396
R-Squared	0.359	0.307

Notes: Heteroskedasticity-consistent standard errors are in parentheses: ** p<0.01, * p<0.05, + p<0.1.

Table A3: GMM Estimates for skilled and unskilled workers

	Skilled workers			Unskilled workers	
	1996	2001	2011	2001	2011
Constant	-0.452	-1.493+	-1.750*	-0.071	-0.385
	(2.699)	(0.825)	(0.892)	(0.330)	(0.354)
Sigma	14.830	4.244+	3.659+	16.783+	10.395*
	(103.912)	(2.430)	(1.899)	(9.704)	(4.710)
tau	0.020	0.069	0.051	0.043**	0.042**
	(0.132)	(0.075)	(0.043)	(0.008)	(0.014)
Homeland	-0.255	-0.368**	-0.272**	-0.157**	-0.179**
	(0.280)	(0.096)	(0.074)	(0.048)	(0.050)
Control	0.828*	1.629**	1.391**	-0.365+	-0.234
	(0.387)	(0.439)	(0.406)	(0.215)	(0.227)
Unemployment rate	-0.034	-0.129	-0.119	0.020	-0.073
	(0.362)	(0.167)	(0.166)	(0.082)	(0.160)
Share mining	-0.023	-0.158	-0.307**	0.021	-0.077
	(0.302)	(0.111)	(0.109)	(0.051)	(0.072)
TBVC	0.101*	0.089	0.116*	0.043	0.090*
	(0.048)	(0.070)	(0.048)	(0.036)	(0.041)
Border region	-0.047*	-0.051	-0.079**	-0.077*	-0.069*
	(0.020)	(0.031)	(0.028)	(0.030)	(0.029)
Market access	0.067	0.236+	0.273**	0.060+	0.096*
	(0.472)	(0.135)	(0.142)	(0.034)	(0.044)
Distance	-0.277	-0.225	-0.137*	-0.686+	-0.398*
	(0.309)	(0.137)	(0.035)	(0.371)	(0.192)
Markup	1.072*	1.308**	1.376**	1.063**	1.106**
	(0.543)	(0.231)	(0.268)	(0.039)	(0.053)
Obs	352	354	354	354	354
Hansen J	0.223	2.298	0.163	3.154	0.755
Hansen p-value	0.637	0.317	0.922	0.207	0.685

Notes: Heteroskedasticity-consistent standard errors are in parentheses: ** p<0.01, * p<0.05, + p<0.1. Column (1) – (3) presents estimates for Skilled workers, defined as the share of workers with any tertiary education – more than 12 years of education and column (4) and (5) presents for Unskilled workers defined as the share of workers with matric or lower than matric level of education - 12 or less years of education. No results are presented for 1996 for Unskilled workers as the gmm estimations struggled to converge. In columns (1) – (3) Control is the skill share variable, while in columns (4) and (5) Control is the unskill share variable.

Table A4: Impact on average wages associated with 10% increase in income of economic centres (%)

	1996	2001	2011
Rest of SA	0.029	0.236	0.385
Homeland	0.003	0.034	0.058
Bophuthatswana	0.026	0.180	0.327
Ciskei	0.000	0.000	0.000
Gazankulu	0.000	0.000	0.000
Kangwane	0.000	0.000	0.000
KwaNdebele	0.038	0.473	0.438
KwaZulu	0.014	0.101	0.152
Lebowa	0.000	0.000	0.000
Qua Qua	0.000	0.000	0.000
Transkei	0.000	0.000	0.000
Venda	0.000	0.000	0.000

Notes: Based on estimation excluding outliers and not including TBVC.

Appendix B: Addressing the challenges with the census data

Inconsistent geographical units across the censuses over time

An advantage of the censuses is the availability of information at different geographical levels (including municipalities, main places (cities/towns) and sub-places (villages/suburbs). However, a major challenge with these geographical units is their inconsistencies across the censuses over time. To address this challenge, we use ArcGIS to overlay 2011 sub-place boundaries onto 1996/2001 magisterial district boundaries. Based on the overlaying results, we use areal-weighting interpolation technique to assign 2011 sub-place population values to their corresponding 1996/2001 magisterial district population values based on a ratio of the area common to both sub place and magisterial districts. This leads to a consistent longitudinal and cross-sectional dataset containing a total of 354 magisterial districts that we use as our unit of analysis.

Challenges with the census income data

An additional challenge with the census is the unavailability of wage (labour income) data which is critical in a study of regional wage disparities. To overcome this challenge, we use regional income per worker to proxy for regional wage per worker. Regional income per worker is calculated as the simple average income of all workers, aged 15-64 years, with a positive income in a region. Three challenges with the census income data are worth noting. Firstly, the census income covers all sources of income including basic salary, bonuses, allowances, income from grants, transfers, remittances and any other income source received by individuals. Our measure of wages is thus likely to exaggerate worker's wages. We test the possible distortion this may introduce in our analysis by comparing regional income per worker and regional wage per worker across 53 districts in South Africa based on National Income Dynamic Study (NIDS) 2010/2011 household survey data.

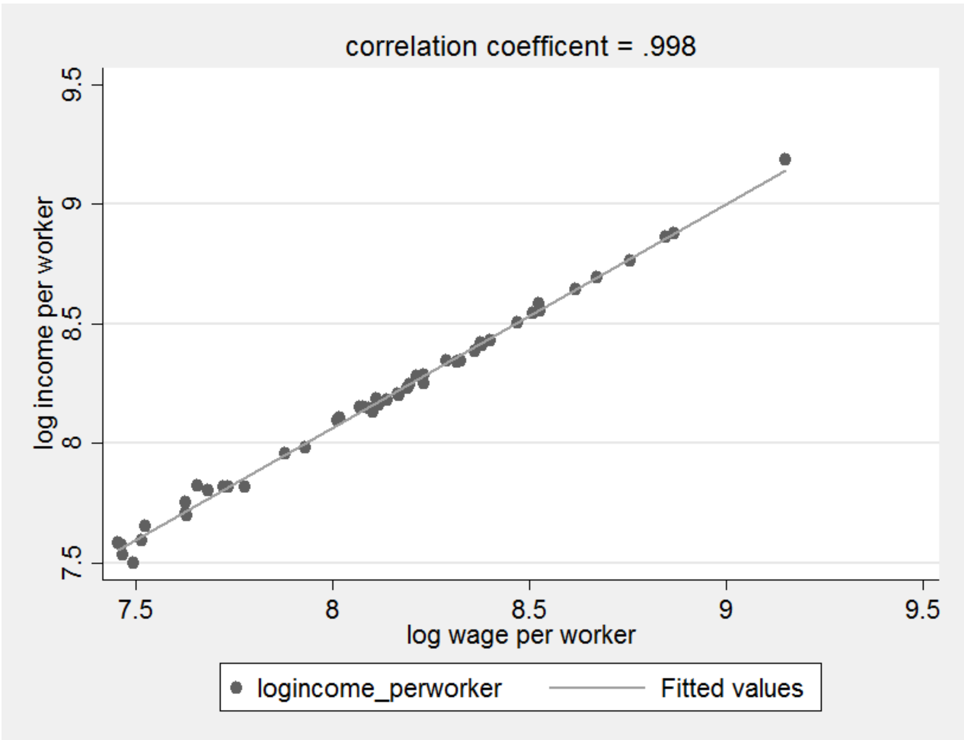
The NIDS survey collects data on various sources of individual income such as labour (wage) income, government grant income, other government income, investment income,

income of a capital nature and remittance income. Of these sources, labour (wage) income is the sum of income from main and secondary job wages, casual wages, self-employment income, 13th cheque, other bonus, profit share, income from helping a friend and extra piece-rate income. From these different sources, we derive three income variables: individual income, income from employed individuals and labour (wage) income. We further derive district income per worker, given as total income from employed individuals divided by total employed individuals in each district, as well as district wage per worker, given as total labour (wage) income divided by total employed individuals in each district.

Our analysis shows that, on average, income from employed individuals contributes about 72.5% to total income across all districts, a figure close to 73.3 percent we find from the analysis of 2011 census data. On the other hand, on average, labour income accounts for about 69% of total income across districts, a figure consistent with 70% normally found in most individual and household studies in post-apartheid South Africa (see Leibbrandt et al., 2010). We also find that on average, labour income (wages) accounts for 94.7% of total income for employed individuals across districts. This shows that the bulk of the income from employed individuals comes from labour income, with roughly 5.3% coming from other sources. Accordingly, we expect regional income per worker to overstate regional wage per worker by a small negligible proportion.

Figure B1 shows a close relationship between income per worker and wage per worker across districts with a correlation coefficient of 0.998. This suggests that, on average, income per worker is a good predictor of wage per worker and income per worker can be said to be a good proxy of wage per worker.

Figure B1: Association between income and wage per worker across districts



Notes: Calculation using income and wage data from NIDS Wave 2 for a sample of 53 district councils.

To support our argument, Table B1 presents simple OLS regression results where we assessed factors explaining district income per worker. In Table B1 district-specific factors (wage per worker, human capital, rural population, participation rate, unemployment rate, share employment in manufacturing, agriculture and mining) are controlled for in a stepwise manner with column (9) incorporating variables that are significant in the stepwise regressions only. Our results in column (9) show that after controlling for wage per worker, no other district-specific factor has significant explanatory power of district income per worker. This evidence highlights that income per worker is a good proxy for wage per worker in South Africa.

Table B1: Association between income per worker and wage per worker across districts in South Africa

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log wage per worker	0.934*								0.935*
	(0.009)								(0.012)
% Human capital		20.746*							-0.284
		(3.942)							(0.398)
% Rural Population			0.347*						0.001
			(0.093)						(0.006)
% Participation rate				-1.259+					-0.017
				(0.529)					(0.033)
% Unemployment rate					-1.309*				0.012
					(0.421)				(0.030)
% Share manufacturing						0.188			
						(1.225)			
% Share agriculture							-0.494		
							(0.548)		
% Share mining								1.264	
								(0.902)	
Constant	0.596*	7.951*	7.992*	9.145*	8.641*	8.201*	8.256*	8.180*	0.594*
	(0.075)	(0.073)	(0.067)	(0.405)	(0.130)	(0.101)	(0.079)	(0.065)	(0.102)
Observations	53	53	52	53	53	53	53	53	52
R-squared	0.997	0.321	0.214	0.151	0.162	0.000	0.015	0.020	0.996
F-test	10586	27.70	14.08	5.663	9.661	0.0236	0.811	1.963	2202

Notes: Standard errors in parentheses ** p<0.01, * p<0.05, + p<0.1. Using income per worker as the dependent variable, column (1) reports the estimates of the association between regional income and wage per worker, while column (2) adds regional specific factors. Human capital is the share of each region's population with at least a tertiary degree.

Secondly, the census income data is provided in brackets with an open-ended top bracket (see Table B2). To construct a continuous income measure, we assign the midpoint of each bracket to everyone in that bracket[1] and we set the midpoint of the open-ended bracket to two times the lower bound of the highest bracket, a rule employed by Stats SA (Statistics South Africa, 2000). While the midpoint approach has been found to exaggerate income inequality (Wittenberg, 2017a) and reduce income variability, it has been found to lead to similar conclusions as other complicated techniques such as the reweighting approach, hot deck approach, mean imputation, and multiple imputation (Ardington et al. 2006; von Fintel, 2007).

Table B2: Monthly personal income brackets in the various censuses

1996 census	2001/2011 census
No income	No income
R1 - R200	R1 - R400
R201 - R500	R401 - R800
R501 - R1000	R801 - R1600
R1001 - R1500	R1601 - R3200
R1501 - R2500	R3201 - R6400
R2501 - R3500	R6401 - R12 800
R3501 - R4500	R12 801 - R25 600
R4501 - R6000	R25 601 - R51 200
R6001 - R8000	R51 201 - R102 400
R8001 - R11000	R102 401 - R204 800
R11001 - R16000	R204 801+
R16001- R30000	
R30001+	

Notes: Brackets for income in the 1996, 2001 and 2011 censuses.

Finally, the census income data is characterised by high rates of reported zero and missing income. For example, 70.7%, 68.2% and 49.1% of the respondents in 1996, 2001 and 2011, respectively, had zero or missing income (see Table B3). Narrowing down to employed individuals who are the focus of this study significantly reduces the proportion of individuals with zero or missing income to 5%, 2.2% and 13.3% in 1996, 2001 and 2011, respectively. Given employed individuals with missing and zero income, the question is how to deal with these individuals. The most common technique which we use in this study is to drop these individuals. However, dropping these individuals can introduce potential bias in parameter estimation (Rubin, 1987). [2]

The extent of this bias is negligible when data is missing completely at randomly (MCAR) and to some extent when data is missing at random (MAR) but not negligible when data is missing not at random (MNAR)[3]. Thus, it is important to check whether income

[1] For example, a band of 1 to 400 rand will take the value of 200.5 Rands. Alternative approaches include weighting, multiple imputation and non-parametric techniques (see Wittenberg, 2017a).

[2] We assume that all employed individuals who reported zero income have missing income because any employed individual is highly likely to receive a positive income. Thus, we drop employed individuals with missing income information.

[3] Data is MCAR if the probability of missingness does not depend on any variable, either observed or unobserved, while data is MAR if the probability of missingness depends only on observed variables and not unobserved or missing information. MCAR is a special case of MAR. Finally, data is MNAR if the probability of missingness depend on unobserved factors which are not measured by the researcher.

information is MCAR, MAR or MNAR. While it is difficult to check whether data is MAR and MNAR, we can easily check whether data is MCAR. We achieve this by running a logistic regression predicting missingness (0 = not missing, 1 = missing) from specific observed variables. Our results presented in Table B4 show highly significant coefficients for the bulk of the factors in all columns, suggesting a violation of MCAR. These results suggest that rather than missing completely at random, income information is missing systematically driven by age, education, race, gender, and location[4].

Table B3: Proportion of individuals with missing and zero income

	All Individuals			Employed Individuals		
Year	Missing	Zero	Missing & Zero	Missing	Zero	Missing & Zero
1996	10.1%	60.6%	70.7%	3.8%	1.2%	5%
2001	0%	68.2%	68.2%	0%	2.2%	2.2%
2011	7.9%	41.2%	49.1%	4.7%	8.5%	13.3%

Note: All Individuals includes all the people interviewed in the censuses.

Given these results, excluding employed individuals with missing income information from our analysis might bias our results[5]. Since our analysis is at the regional level, if those workers with missing income data are concentrated at the bottom of the distribution, then the level of income per worker of a given region will be overestimated. Alternatively, if those workers with missing income information disproportionally fall at the top of the distribution, then the level of income per worker of a given region will be underestimated.

While acknowledging this potential bias, we argue that dropping workers with missing income information will not change our overall conclusions given the small sample size of workers with missing income information in the census (5%, 2% and 13% in 1996, 2001 and 2011). Although there is no established cut-off from the literature regarding an acceptable percentage of missing information for valid statistical inferences, our claim finds support from Roth (1994) who argued that the choice of a missing data estimation technique can have substantial implications for the parameter estimates as the portion of missing data reaches 15% to 20%.

[4] As a robustness check, we also estimated a logistic regression model with regional specific factors like income per worker, market potential, share of workers with higher education and unemployment rate. Our results continued to reveal highly significant coefficients for these factors for all the years, indicating a violation of MCAR.

[5] Interestingly, research suggests that violation of the MCAR does not seriously bias parameter estimates (Collins et al., 2001), especially after controlling for those factors highly correlated with the variable of interest (see Allison, 2001). Accordingly, our empirical analysis controlled for a number of regional specific factors highly correlated with income per worker, to reduce the potential bias due to omitted workers with missing income information.

Table B4: Logistic regression model predicting missingness

VARIABLES	1996 Missing dummy	2001 Missing dummy	2011 Missing dummy
Age	-0.005***	-0.025***	-0.032***
	(0.001)	(0.001)	(0.000)
Education	-0.002	-0.053***	-0.085***
	(0.002)	(0.002)	(0.001)
Gender: Female	0.107***	0.251***	0.380***
	(0.011)	(0.015)	(0.006)
Race: Coloured	0.565***	-0.214***	0.127***
	(0.021)	(0.031)	(0.011)
Indian	0.435***	-0.293***	0.193***
	(0.030)	(0.046)	(0.015)
White	1.146***	0.174***	0.378***
	(0.015)	(0.023)	(0.009)
Location: Urban	-0.013	-0.211***	-0.157***
	(0.016)	(0.020)	(0.008)
Province: Eastern Cape	0.114***	0.327***	0.272***
	(0.024)	(0.034)	(0.013)
Northern Cape	-0.372***	-0.017	-0.261***
	(0.045)	(0.056)	(0.023)
Free State	-0.257***	0.014	-0.143***
	(0.031)	(0.040)	(0.016)
Kwazulu-Natal	0.217***	0.260***	0.203***
	(0.023)	(0.032)	(0.012)
North West	-0.021	-0.197***	-0.089***
	(0.029)	(0.040)	(0.015)
Gauteng	0.176***	0.026	0.076***
	(0.019)	(0.029)	(0.010)
Mpumalanga	0.331***	-0.190***	-0.118***
	(0.028)	(0.041)	(0.015)
Limpopo	0.327***	0.010	-0.075***
	(0.031)	(0.040)	(0.015)
Constant	-3.425***	-2.397***	-0.030*
	(0.033)	(0.045)	(0.018)
Observations	722,718	767,180	1,073,587

Note: Missing dummy is an indicator variable taking a value of 1 if a worker has missing income information and 0 otherwise. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1