Long-run spatial inequality in South Africa: early settlement patterns and separate development

DIETER VON FINTEL

UNIVERSITEIT
STELLENBOSCH
UNIVERSITY

BER
BUREAU FOR ECONOMIC RESEARCH

Long-run spatial inequality in South Africa: early settlement patterns and separate development[1]

Dieter von Fintel*

Abstract

New economic geography theories predict that historically densely settled areas also become more industrialised. Industrial agglomeration has therefore cultivated spatial inequalities in all parts of the world. South Africa presents an interesting case study, where institutional failures interrupted the 'usual' agglomeration process. On the one hand, current day metropolitan regions are located in historically densely populated areas. On the other hand, apartheid-era homelands also had highly concentrated populations, but did not industrialise to the same extent as other parts of South Africa. Much earlier in history, following the *mfecane*, these locations attracted migrants in search of favourable agricultural conditions and physical security in the face of conflict (they were high rainfall, rugged areas). The benefit of settling in these areas, however, only remained prior to imposed restrictions on land ownership (1913 Land Act) and movement of people (during apartheid). This paper decomposes modern spatial inequality, and establishes that agglomerations and historical institutional failures explain large proportions of spatial inequality. Furthermore, the homelands wage penalty reverses once these controls are introduced into various models: had agglomeration taken its course without institutional constraints, the homelands would likely have developed into high paying local economies. While new economic geography theories hold in the urban core, the densely populated former homelands did not follow this trajectory. Spatial inequality is therefore more severe than it would have been had institutional failures not prevented the former homelands from industrialising at the same pace as other historically densely populated areas.

---

1   Introduction

African inequality has not abated significantly from the high levels that arose at the time of colonisation (Bigsten, 2016); in particular, *inter-regional* inequality has remained persistently high, despite predictions of the famous Kuznets (1955) hypothesis in the opposite direction. The foundation of his theory rests on narrowing *spatial* inequalities as countries pass a critical developmental threshold (Kanbur, 2017). Alternative explanations are therefore required to understand the slow movement in overall and spatial inequality. Institutions – whether political or informal – can explain the persistence of *overall* inequality; democratisation provided the context for reductions in inequality in many currently developed countries (Acemoglu & Robinson, 2002; Lakner & Milanovic, 2016; Lindert, 2000) and in other contexts elites continue to protect the status quo to maintain their economic advantages (Sokoloff & Engerman, 2000; Fourie & von Fintel, 2012; Acemoglu & Robinson, 2012). However, none of these "broadly defined" institutions explain why spatial inequality might persist.

New economic geographers provide a lens through which to understand why *spatial* inequality can persist or propagate (Krugman, 1991; Krugell, 2014). The primary reason is the presence of agglomeration forces. If transport costs are high enough (to discourage long distance movement of goods from production plants to consumers), and if economies-of-scale are present in an industry, firms tend to concentrate in regions close to potential markets (or where populations are already densely settled). New firms perpetuate and intensify this pattern, as they attempt to benefit from external economies-of-scale: that is, they prefer to locate close to existing firms to benefit from knowledge spill-overs and to break into existing markets (instead of creating new ones). Existing geographic distributions of industry continue. These factors explain why spatial inequality persists and even grows. 'Initial' pre-industrial settlement patterns therefore matter for the geographic distribution of modern economic activity, but only because it makes sense for new firms to follow their predecessors – even if the contexts and reasons for choosing their locations are no longer relevant in later periods.

This paper considers the case of South Africa, where spatial inequalities are severe and historically entrenched. I show that a large part of spatial earnings inequality can be explained by current labour market conditions (labour demand and worker bargaining power), which are steeped in agglomeration around historically densely populated areas. Labour market conditions are most favourable in current-day metropoles: these regions were also more densely populated than most of the country before industrialisation, and arose around port cities and areas with valuable mineral deposits. The extraction of minerals is one of the factors that has influenced overall inequality in South Africa (Wilson, 2011). Despite South Africa's declining reliance on the primary sector, secondary and tertiary industries continued to develop in these urban pockets to form the industrial core of South Africa. Up to this point, it appears as if South Africa's trajectory conforms to the theories of new economic geography and may therefore suffice to understand modern spatial inequality.

However, separate development also segmented labour and product markets across space. As was the case with current-day metropolitan areas, the former apartheid homelands were also historically densely populated prior to industrialisation. Populations have remained densely concentrated in these regions as a result of limitations on black African migration; however, these same regions have experienced extreme poverty and slow industrialisation, defying the predictions of new economic geography. Apartheid government-led attempts at decentralised industrialisation close to the borders of these regions failed. Historically these areas were densely settled by black Africans, and European colonisers did not sufficiently invest in infrastructure to connect these areas to the industrial core. For instance, the massive expansion of Cape Colony railways between about 1900 and 1906 continued to bypass what would later become the homelands (Herranz-Loncán & Fourie, 2018). Under-investment in homelands infrastructure continued into the apartheid era. Coupled with political uncertainty, this limited the success of decentralised border industries (Lowenberg, 1997). In essence, agglomeration did not occur in these densely populated areas, because of institutional failures that were propagated by the apartheid regime and its predecessors.

I argue that geographically defined separate development prevented agglomeration, which would otherwise have taken root in the former apartheid homelands. Lagging industrialisation in these regions persists despite high historical and current population density. Firms still do not primarily

locate close to these areas, despite potentially high market demand from large settlements. Instead, formal sector firms continue to locate primarily in regions where they benefit from the physical and knowledge infrastructure that supports the existing urban core of the economy. This particular anomaly to new economic geography theories can only be explained by the institutions that were determined by a geographically targeted policy – namely separate development.

Empirical estimates show that historical initial conditions play a similar role to current labour market conditions in explaining regional earnings imbalances. Hence, both agglomeration that emerged in metropoles, *and* historical patterns and institutions that *limited* agglomeration in homelands, explain the extent of spatial inequality in South Africa. I hypothesise that spatial inequality would not have been absent if separate development had not occurred – agglomeration forces would have created a core and periphery as in most countries. Spatial inequality would, however, have been lower had separate development not taken its course: welfare differences between former homelands and current metropolitan areas would have been narrower, as the latter would have attracted more firms. This paper therefore argues that while usual market interactions can lead to spatial inequality, institutional failures can potentially exacerbate the channels proposed by new economic geographers.

This paper decomposes South African spatial earnings inequality in the first decade of the 2000s. The analysis identifies the relative contributions of modern and historical conditions on inter-regional spatial inequality. In particular, I focus on early 20th century settlement patterns in relation to geography and historical climate shocks; subsequently, segregation reinforced the spatial distribution of people on the periphery of the economy, while the structure and spatial pattern of the economy changed to benefit an urban core. Kuznetsian thought would predict that the commencement of freedom of movement in the 1980s should have progressively alleviated large spatial welfare differences. However, apartheid-era spatial differences continue to be replicated (von Fintel & Moses, 2017; Burger *et al.,* 2017). New agglomeration did not occur close to the densely populated homelands in the democratic era. This is because external economies of scale that result from locating close to existing firms remain absent in accordance with the historical pattern.

I use geographically weighted regressions (Fotheringham, Brunsdon, & Charlton, 2002) to find 'local' estimates of mean log earnings, conditional on the factors which I intend to isolate in the

various decompositions. Results show that homelands regions still experience a large wage penalty relative to the rest of the country. This phenomenon can, however, be explained by high local unemployment and unionisation rates, as well as historical patterns of settlement and climate shocks. Had these proxies for agglomeration (or lack thereof) been equalised across space, the apartheid homelands would likely have been high-paying regions today. The variance decomposition allows me to estimate reductions in spatial earnings inequality that would likely occur should the constraints of the past be removed in a counterfactual world. I argue that one part of the homeland penalty results from settlement patterns that occurred before the unification of South Africa in 1910: while these conditions were favourable for welfare at that time, this is no longer the case in the 2000s. Migration limitations imposed by the apartheid government prevented people from moving at the time that these same conditions became less favourable for economic success. In the context of structural evolution, whereby the economy moved from primary to secondary and tertiary production, the existing spatial population pattern locked people into rural areas that were becoming increasingly isolated from core economic activity. The persistence of this effect is also associated to low labour demand that has remained statically low, even after democratisation and the liberalisation of internal migration. This concords with new economic geography theory: new firms locate where other firms are already established. Though former political institutions are abandoned, their legacy remains manifest through permanent changes in local labour demand. In turn, spatial inequality continues to follow historical patterns.

2   Spatial inequality in South Africa

Spatial inequality is one of the defining features of overall inequality in South Africa, and remains strongly correlated with race (Burger *et al.,* 2017). A string of discriminatory legislation that followed the unification of South Africa under British rule in 1910 laid the foundations for these strong and persistent regional differences. Most famously, the Land Act of 1913 limited acquisition of land by the black African majority to an area that comprises only 7% of South Africa's surface; this was later expanded to 13.5% by the 1936 Land Act. These areas were then converted to apartheid-era homelands – some with self-governing status – which were to form the pillars of so-called "separate development". Black Africans were not regarded as citizens of South Africa, but of homelands. Between the 1960s and 1980s, approximately 3.5 million individuals were forcibly relocated to live in these densely populated regions (Abel, 2016). In 1950, the Group Areas Act

reinforced this segregation. Black African individuals were not allowed free access to live and work in (white) urban areas, and influx controls were introduced. Predominantly men followed a circular migration pattern between permanent households in the homelands and temporary homes in urban areas where they found jobs.

Former homelands regions continue to face the highest levels of unemployment in the country (von Fintel, 2017). Standard theory predicts that under conditions of free movement, migration should take place and partially equalise wage and employment imbalances across regions (Harris & Todaro, 1970). Unemployment also remains high in urban centres, however, yet lower than in former homelands, so that the probability of obtaining a job after migration remains poor. Therefore, despite large migration flows since the abolition of influx controls in 1986, unemployment has also not shifted to urban areas to equalise across space (von Fintel & Moses, 2017). Mudiriza (2018) provides first estimates of regional wage convergence for South Africa: using post-apartheid census data, he estimates that it would take 19 years to halve the regional wage gap in South Africa. Convergence across regions therefore remains slow, and local labour markets are likely to remain segregated (into low paying, high unemployment rural areas and higher paying, lower unemployment urban areas) in line with existing patterns of spatial inequality. Once controlling for structural hindrances such as local human capital, local unemployment and homeland status, however, this figure drops to only 5 years. Consequently, the former homelands is one notable anomaly that consistently lags behind the rest of the country and slows down convergence across space.

The apartheid homelands are rooted in historical political institutions and continue to have a strong bearing on maintaining spatial inequality. In the early 2000s the former homelands regions – which continue to be dominantly settled by black Africans – still lagged behind the rest of South Africa in terms of other basic welfare indicators, such as self-reported hunger. The expansion of cash transfers towards these regions has, however, gradually narrowed the spatial gap in food adequacy (Pienaar & von Fintel, 2014). A substantial homelands earnings penalty exists (Kingdon & Knight, 2006). Accounting for high local unemployment rates overturns this result, yielding what would have been a wage premium had labour demand been more robust. Kingdon and Knight (2006) argue that institutionalised wage setting, together with limited labour demand in the former homelands, has raised wages above what would otherwise be a restrictively low market clearing

level – even if wages remain far lower than in the rest of the country where labour demand is higher. The apartheid geographic split continues to operate in the labour market. While people can now move relatively easily into other labour markets with higher demand for their work, firms do not necessarily relocate to areas where labour surpluses exist. Despite the abolition of restrictions on human movement, separate development still manifests through a low labour demand and poor agglomeration in the former homelands regions. This follows historically low infrastructure investment and the failure of decentralised industrialisation (Lowenberg, 1997). Political institutions can therefore linger beyond their expiration if they result in structural economic differences that persist. A 'break' in agglomeration during colonial times and the apartheid period now continues as a dispersion force in the former homelands.

While separate development has contributed to large spatial inequality in current-day South Africa, historical evidence suggests that relative prosperity existed in the regions that were to become homelands. In a counterfactual world, therefore, had institutional hindrances not taken their course, it is likely that agglomeration may have emerged in these regions, leading to their industrialisation. For instance, black African peasant farmers successfully participated in commercial activities and exported produce beyond colonial borders (Bundy, 1979). Prior to 20th century restrictions on land ownership, black African agriculture provided a viable livelihood for this population group. Before the imposition of the 1913 Land Act, black African populations were already concentrated in the areas that would later formally become the homelands (see Figure 1, bottom left panel). Therefore, the spatial distribution of human settlement was determined by decisions and migrations that occurred *before* cross-cutting segregation laws were formalised by the Union government, and also before forced resettlement occurred (Abel, 2016). Restrictions on ownership and movement that followed, however, did not allow populations to move in response to *later* regional economic changes. The result is, therefore, that the spatial distribution of people may have matched favourable agricultural conditions for black Africans in the early 20th century; this spatial pattern was 'fixed' by limitations on ownership and movement. As the economy diversified away from agriculture, and colonial authorities forced African participation in wage labour through hut and poll taxes (Redding, 2000), living in locations away from emerging centres placed black Africans at a severe economic disadvantage. These areas could – in all probability would – have followed agglomeration trajectories; however, institutional limitations prevented this process.

An exploration of early settlement patterns is therefore necessary to understand forward linkages to later outcomes. Hannaford *et al.* (2014) critically evaluate a large literature on state formation and the movement of Bantu populations in southern Africa. A popular argument posits that climate variability was pivotally associated with social and political upheaval; in particular, the *mfecane* is a collection of conflicts that resulted in the dispersion of the black African Bantu-speaking people groups across the region up until 1890. While Hannaford *et al.* (2014) dispute the veracity of the causal links between climate and peoples' settlement choices during this period, it nevertheless remains a compelling argument that drought was a determining factor. Modern studies that explore the link between climate and conflict enjoy the benefit of widely available data sources. The causal relationship has been clearly established by multiple authors (Burke, Hsiang & Miguel, 2015), lending credence to earlier arguments that the *mfecane* may have been causally attributed to climate instability.

The combination of conflict and climate shocks critically have bearing for locational decisions. Figure 1 shows that – apart from large urban centres – the 1911 population was concentrated in areas where rainfall (the average between 1900 and 1913) was particularly high and the terrain was rugged. On the first count, it suggests that populations settled in areas which supported successful peasant agriculture (Bundy, 1979). Terrain ruggedness, however, is known to be detrimental for economic prosperity in a global perspective. African populations, on the other hand, were shielded in the long run by this 'poor' geography, particularly because of its link with the slave trade; slave raids were prevented in rough terrain, so that exploitation was limited in these regions (Nunn & Puga, 2012). Similarly, one might argue, rugged areas were beneficial for the escape from conflict associated with the *mfecane* in southern Africa (von Fintel & Fourie, 2018). The added benefit of high rainfall made this an appropriate place of settlement for the promotion of agriculture in safety. While these conditions were favourable in the beginning of the 20th century, the structural transformation away from agriculture, the cessation of tribal conflict, the advent of the migrant labour system and influx controls meant that an initially favourable spatial equilibrium converted into large and enduring spatial inequalities. This is particularly true because people were forced to remain in these regions, but industry did not follow the same pattern.

The rest of this paper will weigh up the respective contributions of limiting residence to regions that became increasingly isolated from the core economy (the direct effect) and the long-run impact on local labour demand (the indirect effect) on modern day spatial earnings inequality.

3   Methods and data

This paper decomposes one measure of labour market earnings inequality – the variance of log earnings – into various effects, including historical and geographic factors (climate shocks, initial settlement patterns and ruggedness) and modern local labour market conditions (which may also be influenced by past institutions). Labour market earnings constitute the bulk of total income inequality in South Africa (Leibbrandt, Woolard & Woolard, 2009). If individual earnings are determined by:

$$\log(earnings_{irt}) = \beta_0 + \alpha' x_{irt} + \mu_r + u_{irt} \tag{1}$$

where i indexes individuals, t indexes time, and r represents various regions; x_{irt} is a vector of 'standard' covariates that determines earnings in the modern labour market (such as education, age, race and gender), and μ_r is a regional fixed effect that can be further decomposed as:

$$\mu_r = \phi_r + \gamma'_r h_r + \delta'_r m_r + \varepsilon_r \tag{2}$$

where h_r is a vector of variables that determine historical settlement patterns *before* the formalisation of many pivotal separate development laws, and m_r represents a vector of modern, region-specific, labour market outcomes that are directly influenced by agglomeration and dispersion effects (that are in turn influenced by segregation policies). I first estimate the Mincerian earnings equation (1) with regional fixed effects. I then extract $\hat{\mu}_r$ from this regression to model the second equation (2). Coefficients of this equation, ϕ_r, γ_r and δ_r, vary by region: I implement geographically weighted regressions (Fotheringham, Brunsdon & Charlton, 2002), allowing the estimated relationships to be heterogeneous over space. For instance, separate development has led to clear differences in wages across homeland borders. These can potentially be attributed to

geographically heterogeneous effects of historical settlement patterns that are no longer favourable in an evolving economy.
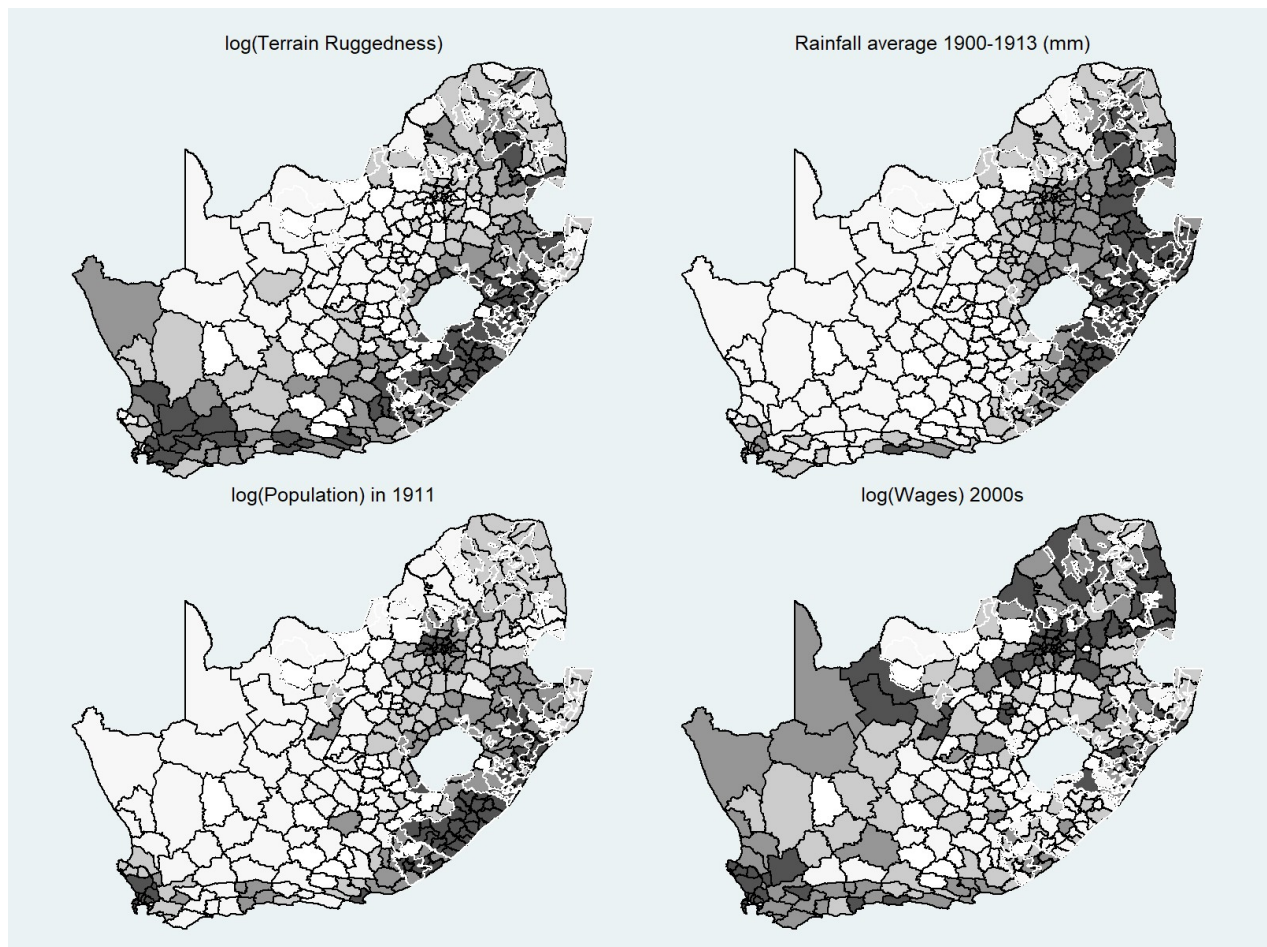
Figure 1: Spatial Distribution of terrain ruggedness, pre-1913 rainfall, 1911 population and wages in early 2000s

Typically a spatial regression discontinuity design is suitable to model distinct border effects which we may expect to find for the former homelands (Dell, 2010; Magruder, 2012). Alternatively, one could match and compare spatially contiguous regions on either side of the border. These tools, however, entail working with small regions close to the homelands only, and do not allow for a decomposition of country-wide spatial inequality. Instead, geographically weighted regression produces separate coefficient estimates and residuals for *all* districts and are also adept at detecting spatial breaks (Fotheringham, Brunsdon & Charlton, 2002). They depend on a set of local models that cover the entire country, whether they are near the homelands or not. The approach also accounts for spatial autocorrelation. I can therefore quantify the effect that occurs when rich (poor)

regions interact with other proximate rich (poor) regions to re-inforce inequality. The same functional form for μ_r is estimated on a set of surrounding districts that fall within an optimal distance bandwidth. Hence, the intercept ϕ_r in each region functions as a fixed effect, and provides a depiction of how average wages vary across space once conditioning on other covariates. The bandwidth for the spatially local regressions is chosen by leave-one-out cross validation. Controls are sequentially introduced to monitor cross-district variation in wages that remains after stripping out the effects of h_r and m_r respectively. At each stage a decomposition is done, making use of the following identity:

$$
\begin{aligned}
Spatial\ inequality =\ & Var(\log(regional\ earnings)) \\
=\ & Var(\hat{\mu}_r) \\
=\ & Var(\hat{\phi}_r + \hat{\gamma}'_r h_r + \hat{\delta}'_r m_r) + Var(\hat{\varepsilon}_r) \qquad (3)
\end{aligned}
$$

Reductions in spatial inequality attributable to control variables is represented as the ratio $\frac{Var(\hat{\mu}_r) - Var(\hat{\varepsilon}_r)}{Var(\hat{\mu}_r)}$, which is essentially the R^2 of that regression. I also place this in broader context, by comparing changes in relation to *total* (a-spatial) inequality - $Var(\log(earnings_{irt}))$ - and inequality that strips out individual-specific covariates.

To conduct the proposed analysis, and especially to recover the regional fixed effects for the second stage, microdata over time is required. The reason for this is that some of the elements in x_{irt} are measured as regional aggregates. A further necessary criteria is that a region with the same geographic definition should be followed throughout the cross sections. Labour Force Surveys from September 2000 to March 2004 meet these criteria (Magruder, 2012). They are enumerated twice annually, and contain indicators of magisterial districts, an administrative division that is associated with South Africa's local courts (see Figure 1, where they are sketched). Labour market earnings are enumerated in brackets, and I take the midpoint. I also construct local unemployment and unionisation rates from this data to construct the vector m_r. The set of controls x_{irt}, includes standard Mincerian variables (quadratics in age and education), race, gender and regional skills and sector composition of jobs.

I construct h_r from various historical and geographic sources. Historical population densities were transcribed from the district-level records of the 1911 census (Union of South Africa, 1912). Terrain

ruggedness is sourced from the data appendix of Nunn and Puga (2012). Grid cells are aggregated to match the magisterial districts. Rainfall grids – both for the periods corresponding to the Labour Force Survey, and the average of 1900-1913 records – are obtained from the reconstructions by Willmott and Matsuura (2015). In particular, I analyse rainfall shocks rather than levels. That is, the variables are expressed as a deviation from their long-run trend in millimetres.

4   Results

4.1   Global regressions

Tables 1 and 2 show estimates from Ordinary Least Squares models, with samples differentiated by former homeland status. Estimates rely on the assumption of parameter homogeneity across space within the specified sample. The dependent variable is constructed from the regional fixed effects of Mincerian equation (1)[2]. These results illustrate the 'average' effects of various modern and historical factors on regional wages.

Table 1 Global model

Dependent: $\mu_r = \log(district\ wage)$	1	2	3	4
Sample is in former homelands	N	Y	N	Y
Average education in district	0,079	0,136	0,065	0,083
	(0,011)***	(0,023)***	(0,011)***	(0,022)***
Proportion of district female	-1,779	-0,607	-0,769	0,163
	(0,306)***	(0,573)	(0,350)**	(0,511)
log(2000s unemployment rate)			-0,345	-0,069
			(0,223)	(0,236)
2000s Unionisation rate			0,921	2,319
			(0,153)***	(0,389)***
Constant	3,496	2,410	3,814	4,070
	(0,191)***	(0,431)***	(0,171)***	(0,459)***
Difference in intercepts				
(Homelands disadvantage)		108,6%		-25.6%
N	224	91	224	91
R^2	0,339	0,389	0,436	0,607

NOTES: *p<0,1 **p<0,05 ***p<0,01

---

[2] These results are available on request.

Baseline estimates in columns 1 and 2 of Table 1 show that significant gender discrimination exists outside the former homeland borders. Returns to education are substantially higher within the former homelands. Of greatest interest, however, is that the intercept remains lower in the former homelands regions, even after introducing these controls. Differencing the intercepts across specifications indicates that average wages are more than double outside the former homelands than inside them. This penalty represents a large contribution to spatial inequality that requires explanation.

Controlling for early 21st century local labour market conditions in columns 3 and 4 reverses this pattern. In a counterfactual world where local unemployment and unionisation rates would be equal across the region types, wages would be approximately 25% higher in the homelands compared to the rest of the country. These two variables represent the benefits of agglomeration in urbanised regions (or the lack of agglomeration in the former homelands): higher labour demand and better organisation of workers. The analysis confirms that low labour demand in the homelands drives wages down (Kingdon & Knight, 2006). If labour demand were not as depressed in the former homelands as it currently is, prevailing conditions would allow wages to grow significantly. The reversal represented by these results indicate a missed opportunity: the homelands could have been high-paying regions had 'natural' agglomeration forces not been quenched by separate development and historically poor infrastructure development.

In specifications 5 and 6 in Table 2, I control only for historical and geo-physical variables. Foremost, initial population density (in 1911) is only positively significantly correlated with wages outside the former homelands; to the contrary, this is not the case inside these borders. This supports the notion that agglomeration effects only followed early settlement patterns in the more developed part of the economy; separate development, however, prevented the link between early market potential on the one hand, and subsequent wage and productivity growth on the other, in the marginalised homelands.

Table 2 Global model

Dependent: $\mu_r = \log(district\ wage)$	5	6	7	8
Sample is in former homelands	N	Y	N	Y
Average education in district	0,058	0,112	0,036	0,067
	(0,013)***	(0,022)***	(0,013)***	(0,020)***
Proportion of district female	-1,913	-0,830	-0,806	-0,034
	(0,280)***	(0,521)	(0,300)***	(0,461)
log(2000s unemployment rate)			-0,417	-0,107
			(0,226)*	(0,219)
2000s Unionisation rate			1,046	2,127
			(0,144)***	(0,344)***
log(Terrain ruggedness)	-0,003	-0,090	0,004	-0,085
	(0,020)	(0,032)***	(0,019)	(0,028)***
Current rainfall deviation (mm)	-0,250	-0,382	-0,233	-0,288
	(0,085)***	(0,112)***	(0,077)***	(0,089)***
1900-1913 rainfall deviation (mm)	0,582	0,085	0,903	0,482
	(0,342)*	(0,489)	(0,290)***	(0,346)
log(1911 population)	0,027	-0,002	0,030	0,008
	(0,015)*	(0,022)	(0,014)**	(0,019)
Constant	3,630	2,747	4,028	4,164
	(0,186)***	(0,394)***	(0,160)***	(0,427)***
Difference in intercepts (Homelands disadvantage)	88,3%		-13,6%	
N	224	91	224	91
R^2	0,406	0,536	0,528	0,712

NOTES: *p<0,1 **p<0,05 ***p<0,01

Similarly, positive rainfall shocks prior to 1913 (which also proxy for early settlement patterns) remain correlated with wages outside the former homelands; the effect is zero inside the homelands. Keeping all else constant, wages are lower in regions that experience positive *contemporaneous* rainfall shocks. The contrast between the effects of rainfall in various periods on modern wages emphasises structural change in the economy: initial reliance on agriculture meant that populations flocked towards regions with high rainfall; after industrialisation this settlement pattern persisted, even if today's settlement patterns or industrial choices are not primarily driven by agro-climatic considerations. While I hypothesise that early settlement flocked *towards* rugged regions for the

purposes of escaping conflict, wages today are lower in rugged homeland regions, in line with more standard views on the effects of terrain. These coefficients shed light on different agglomeration patterns, but columns 5 and 6 cannot fully explain why wages in homelands regions are lower than in the rest of the country. Columns 7 and 8, which control for both current labour market conditions and historical variables, again yield a small homelands premium. In summary, these results show that South Africa experienced a process of agglomeration that followed initial settlement patterns – except for the homelands regions. Local labour market conditions that resulted from this imbalanced development pattern increased spatial inequality. I hypothesise that had separate development not occurred, the densely settled former homelands may also have developed into industrialised regions, reducing spatial inequality somewhat.

4.2   Local regressions

'Global' results mask local heterogeneity and spatial autocorrelation. Figures 2 therefore disaggregates these results by location, using local models with similar specifications to those in Tables 1 and 2. I limit myself to analysing mean wages (or the intercept), conditional on various covariates.[3] The top left panel of Figure 2 reflects the patterns found in the global model. Workers in urban centres (Gauteng in the North and Cape Town in the South) earn the highest wages, while lower wages are concentrated in outlying areas – especially in the former homelands. Most notably, the former Transkei along the south east coast is the poorest earning local labour market in the early 21st century.

The top right panel conditions on contemporary local unemployment and unionisation rates. Would the pressure of low labour demand be removed, the Transkei would be classified as one of the *highest* earning regions in the country. This reflects the estimates of Kingdon and Knight (2006), who conclude that the homelands are high wage regions once controlling for their poor labour demand.   A similar pattern emerges if I only control for historical and geographical conditions (as in specifications 5 and 6 of Table 2). Historical patterns therefore exert similar influence on the spatial wage structure as do modern local labour market conditions. I conjecture that they represent similar effects. Both sets of variables capture patterns of agglomeration:

---

[3] All intercepts have sufficiently large T-values to be considered statistically significant. Results for other coefficients follow similar signs to those in the global regressions, but are only significant in the most and least populated regions in 1911.

historical and geographic variables capture 'initial' conditions; local labour market variables represent the resulting agglomeration that followed these patterns, together with the dispersion effects caused by institutional failure in the apartheid homelands. Results from local regressions therefore support patterns identified in the global regressions; had separate development not prevented agglomeration, the former homelands could (all else being constant) have been high paying labour markets.
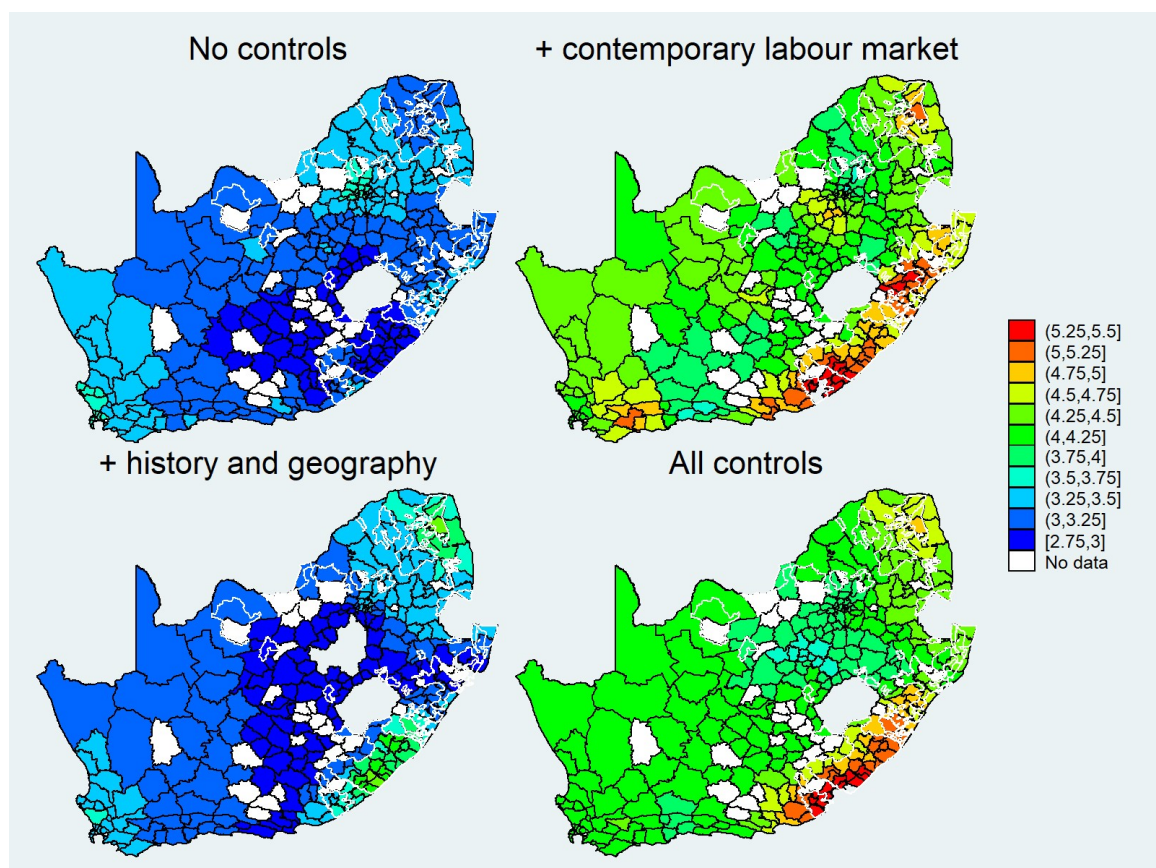

Figure 2: Conditional mean wages from various geographically weighted models

4.3   Decomposition of spatial inequality

I now turn to decomposing the contributors that explain (spatial) earnings inequality. Table 3 presents the proportion of the variance of log earnings (or spatial inequality) that is explained by various modelling strategies. These figures are presented relative to various benchmarks. Firstly, we analyse the contributors to individual inequality noticed in the raw data, as shown in column 1.[2] Standard Mincerian, demographic and labour market composition controls (as outlined in

section 3) explain more than half of wage inequality across individuals. Staying with a global model, estimates of inter-regional variation in inequality ($Var(\hat{\mu})$) are about 95% of the size of variation of individual earnings ($Var[\log(earnings)]$). Spatial variation in earnings is therefore a strong reflection of individual wage inequality. Estimating this same quantity with a local geographically weighted model including only a constant, raises this figure to 97.95%. The improvement in fit results from smoothing within the bandwidth, which accounts for spatial autocorrelation.[4]

Table 3 Spatial inequality decomposition

		% of variation explained relative to:		
		(1)	(2)	(3)
Model type	Controlling for	Individuals	Regions	Regions
Global	Micro controls	53,55%		
	All within-region variation (=inter-regional spatial inequality)	95,38%		
Local	Spatial autocorrelation	97,95%	55,59%	
	+ contemporary labour market	99,30%	84,86%	65,91%
	+ history and geography	98,51%	67,70%	27,27%
	+ all regional controls	99,46%	88,24%	73,51%

NOTES: Column (1) refers to individuals in the raw data, column (2) refers to regions before controlling for spatial autocorrelation, column (3) refers to regions after controlling for spatial autocorrelation

Given the high proportion of individual variance already explained by between-region inequality, I now focus on explaining only the spatial, cross-regional element. Column 2 explains inequality relative to $Var(\hat{\mu})$ obtained from a global model that does not adjust for spatial autocorrelation. Moving from the global to a local model (without adding controls, as in the top left panel of Figure 2) accounts for spatial autocorrelation in the error terms. This explains approximately 56% of spatial inequality. Local spill-over effects and interconnectedness of adjacent regions re-inforce inequality, as wealthier regions tend to interact with wealthier regions, and poorer with poorer.

---

[4] I use a cross-validated bandwidth. Other methods of bandwidth selection are more suitable to solve spatial autocorrelation problems (Cho, Lambert & Chen, 2010). However, the change in fit noted here can only be attributed to accounting for spatial autocorrelation, as no covariates are included in this part of the analysis.

In column 3 the departure point is inequality that already accounts for spatial autocorrelation. Off this basis, an additional 66% of inequality can be explained by contemporary local labour market conditions. These effects represent agglomeration. Historical and geographic controls account for less than half of spatial inequality compared to modern labour market variables. While history plays a role in explaining spatial inequality, its direct effect is smaller than indirect channels through which it also operates.

5   Conclusions

South Africa has battled with persistently high inequality for centuries, despite becoming the most developed economy on the African continent in modern times (Fourie & von Fintel, 2010, 2011; Leibbrandt, Woolard & Woolard, 2009). This pattern not only defies the Kuznets (1955) hypothesis, but also the predictions of Acemoglu and Robinson (2002) who posit that democratisation is an impetus for greater equality. Contrary to this proposition, inequality has grown in the democratic era. Overall inequality is mirrored closely by the spatial distinctions introduced by separate development legislation in South Africa. My estimates show that this overlap is strong. Former apartheid homelands, which are located in the regions reserved for black African land ownership by the 1913 Land Act, remain low wage regions with sluggish labour demand; despite the removal of influx controls and rapid increases in migration flows, spatial imbalances have not averted after democratisation.

This paper shows that about half of earnings differences across local labour markets can be explained by positive spatial autocorrelation. It emphasises the strong localisation of labour market earnings: better-paying regions are clustered in areas adjacent to other better-paying districts; poverty in one region reinforces the poverty experienced by neighbours. This strong spatial concentration and polarization is the product of separate development. A large part of inequality that is not explained by this pattern can be attributed to historical settlement patterns, geography and climate shocks. Initial conditions continued, as economic activity agglomerated around existing market potential, except in the former apartheid homelands. Initial settlement before the formalisation of separate development (measured in 1911) is strongly correlated with positive rainfall shocks and terrain ruggedness. It supports a hypothesis that black Africans flocked to regions that favoured cultivation and were shielded from conflicts induced by the *mfecane*. The

rise of peasant farming in this period suggests that this was a successful settlement strategy. However, this pattern became detrimental in the long run, as the 1913 Land Act and subsequent separate development legislation locked populations into these areas that were separate from the economic core. Under-investments in infrastructure by colonial and apartheid-era governments represent an institutional failure that prevented agglomeration in these regions. As a consequence of geographic isolation from urban labour markets and a changing economy which moved away from agriculture, labour demand remained too low to yield living wages to most of the population in former homelands. Hence, current local labour market conditions explain a large proportion of earnings inequality, even after the dissolution of formal separate development. While people may relocate to more favourable local labour markets, firms have not moved to areas that remain geographically isolated. The result is that despite formal changes in institutions and large migration of people, former economic inequalities have remained intact.

New economic geography theories explain why spatial inequalities arise and grow. This paper has shown that agglomeration effects could potentially have manifested in former apartheid homelands had institutional barriers not prevented the normal course. Historically, therefore, firms did not locate to the former homelands. While market potential still remains high due to dense settlement, external economies of scale remain absent. New firms are also unlikely to establish operations close to the former homelands, so that spatial inequalities may continue into the future. Failed industrial decentralisation during the apartheid period bears testament to the difficulties in remedying existing spatial inequalities. South Africa therefore faces a unique situation in which densely populated areas have not grown into industrialised local economies. Solutions to spatial inequality may therefore not rely on standard economic forces.

In the democratic era, the welfare of residents in former homelands has improved due to the expansion of government cash transfers (Pienaar & von Fintel, 2014). The amounts of these government grants are, however, too low to bridge spatial imbalances across regions. Together with continued low labour demand in rural homeland areas, and relatively high unemployment in urban areas, it is furthermore unlikely that urbanisation can alleviate spatial inequality. Hence, I argue that despite the removal of formal separate development policies, their economic legacy perpetuates the spatial inequalities that they were originally designed to create.

References

Abel, M. 2016. Long-run effects of forced resettlement: Evidence from apartheid South Africa. Harvard University. https://scholar.harvard.edu/abel/publications/long-run-effects-forced-removal-under-apartheid-social-capital.

Acemoglu, D. & Robinson, J.A. 2002. 'The political economy of the Kuznets Curve', *Review of Development Economics,* **6**(2), 183–203. https://doi.org/10.1111/1467-9361.00149.

Acemoglu, D. & Robinson, J.A. 2012. *Why nations fail: The origins of power, prosperity and poverty.* Profile Books.

Bigsten, A. 2016. 'Determinants of the evolution of inequality in Africa', *Journal of African Economies*, December. https://doi.org/10.1093/jae/ejw028.

Bundy, C. 1979. *The rise and fall of the South African peasantry.* University of California Press.

Burger, R., Van der Berg, S., Van der Walt, S. & Yu, D. 2017. 'The long walk: Considering the enduring spatial and racial dimensions of deprivation two decades after the fall of apartheid', *Social Indicators Research,* **130**(3), 1101–23. https://doi.org/10.1007/s11205-016-1237-1.

Burke, M., Hsiang, S.M. & Miguel, E. 2015. 'Climate and conflict', *Annual Review of Economics,* **7**(1), 577–617. https://doi.org/10.1146/annurev-economics-080614-115430.

Cho, S-H., Lambert, D.M. & Chen, Z. 2010. 'Geographically weighted regression bandwidth selection and spatial autocorrelation: An empirical example using Chinese agriculture data', *Applied Economics Letters,* **17**(8), 767–72. https://doi.org/10.1080/13504850802314452.

Dell, M. 2010. 'The persistent effects of Peru's mining mita', *Econometrica,* **78**(6), 1863–1903. https://doi.org/10.3982/ECTA8121.

Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton. 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Wiley.

Fourie, Johan, and Dieter von Fintel. 2010. 'The Dynamics of Inequality in a Newly Settled, Pre-Industrial Society: The Case of the Cape Colony'. Cliometrica **4**(3), 229–67. https://doi.org/10.1007/s11698-009-0044-1.

———. 2011. 'A History with Evidence: Income Inequality in the Dutch Cape Colony'. Economic History of Developing Regions **26**(1), 16–48. https://doi.org/10.1080/20780389.2011.582990.

———. 2012. 'The Fruit of the Vine? An Augmented Endowments-Inequality Hypothesis and the Rise of an Elite in the Cape Colony'. In The Role of Elites in Economic Development, edited by Alice Hoffenberg Amsden, Alisa DiCaprio, and James A. Robinson. OUP Oxford.

Hannaford, Matthew, Grant Bigg, Julie Jones, Ian Phimister, and Martial Staub. 2014. 'Climate Variability and Societal Dynamics in Pre-Colonial Southern African History (AD 900–1840): A Synthesis and Critique'. Environment and History **20**, 411–45.

Harris, John R., and Michael P. Todaro. 1970. 'Migration, Unemployment and Development: A Two-Sector Analysis'. The American Economic Review **60**(1), 126–42.

Herranz-Loncán, Alfonso, and Johan Fourie. 2018. '"For the Public Benefit"? Railways in the British Cape Colony'. European Review of Economic History **22**(1), 73–100. https://doi.org/10.1093/ereh/hex010.

Kanbur, Ravi. 2017. 'Structural Transformation and Income Distribution: Kuznets and Beyond'. 266. African Development Bank Working Paper. African Development Bank Group.

Kingdon, Geeta Gandhi, and John Knight. 2006. 'How Flexible Are Wages in Response to Local Unemployment in South Africa?' Industrial and Labor Relations Review **59**(3), 471–95.

Krugell, Waldo. 2014. 'The Spatial Persistence of Population and Wealth During Apartheid: Comparing the 1911 and 2011 Censuses'. Economic History of Developing Regions **29**(2), 336–52. https://doi.org/10.1080/20780389.2014.957907.

Krugman, Paul. 1991. 'Increasing Returns and Economic Geography'. Journal of Political Economy **99**(3), 483–99.

Kuznets, Simon. 1955. 'Economic Growth and Income Inequality'. The American Economic Review **45**(1), 1–28.

Lakner, Christoph, and Branko Milanovic. 2016. 'Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession'. The World Bank Economic Review **30**(2), 203–32. https://doi.org/10.1093/wber/lhv039.

Leibbrandt, Murray, Ingrid Woolard, and Chris Woolard. 2009. 'Poverty and Inequality Dynamics in South Africa: Post-Apartheid Developments in the Light of the Long-Run Legacy'. In

South African Economic Policy under Democracy, edited by Janine Aron, Brian Kahn, and Geeta Kingdon. OUP Oxford.

Lindert, Peter H. 2000. 'Chapter 3 Three Centuries of Inequality in Britain and America'. In Handbook of Income Distribution, 1:167–216. Elsevier. https://doi.org/10.1016/S1574-0056(00)80006-8.

Lowenberg, Anton D. 1997. 'Why South Africa's Apartheid Economy Failed'. Contemporary Economic Policy **15**(3), 62–72. https://doi.org/10.1111/j.1465-7287.1997.tb00478.x.

Magruder, Jeremy R. 2012. 'High Unemployment Yet Few Small Firms: The Role of Centralized Bargaining in South Africa'. American Economic Journal: Applied Economics **4**(3), 138–66. https://doi.org/10.1257/app.4.3.138.

Mudiriza, G. 2018. Regional wage disparities in post-apartheid South Africa: Spatial patterns, convergence dynamics and causes. PhD Dissertation. University of Cape Town.

Nunn, N & Diego, P. 2012. 'Ruggedness: The blessing of bad geography in Africa', *The Review of Economics and Statistics,* **94**(1), 20–36.

Pienaar, L. & Von Fintel, D. 2014. 'Hunger in the former apartheid homelands: Determinants of convergence one century after the 1913 land act', *Agrekon,* **53**(4), 38–67. https://doi.org/10.1080/03031853.2014.929014.

Redding, S. 2000. 'A blood-stained tax: Poll tax and the bambatha rebellion in South Africa', *African Studies Review,* **43**(2), 29–54. https://doi.org/10.2307/524983.

Sokoloff, K.L. & Engerman, S.L. 2000. 'Institutions, factor endowments, and paths of development in the new world', *Journal of Economic Perspectives,* **14**(3), 217–32. https://doi.org/10.1257/jep.14.3.217.

Union of South Africa. 1912. Census of the Union of South Africa 1911. Pretoria: Government Printer.

Von Fintel, D. 2017. 'Institutional wage-setting, labour demand and labour supply: Causal estimates from a South African pseudo-panel', *Development Southern Africa,* **34**(1), 1–16. https://doi.org/10.1080/0376835X.2016.1269637.

Von Fintel, D. & Moses, E. 2017. 'Migration and gender in South Africa: Following bright lights and the fortunes of others?', *Regional Science Policy & Practice,* **9**(4), 251–68. https://doi.org/10.1111/rsp3.12106.

Von Fintel, D. & Fourie, J. 2018. *Complicating history: The persistence or reversal of subnational fortune in South Africa.* Mimeo. Stellenbosch University.

Willmott, C.J. & Matsuura, K. 2015. 'Terrestrial air temperature and precipitation: monthly and annual time series (1950 - 2014)'. 2015. http://climate.geog.udel.edu/~climate/html_pages/download.html#P2014.

Wilson, F. 2011. 'Historical roots of inequality in South Africa', *Economic History of Developing Regions,* **26**(1), 1–15. https://doi.org/10.1080/20780389.2011.583026.