SA-TIED YOUNG SCHOLARS PROGRAMME

Regional inequality and rural dependency in South Africa

How can opposing trends in regional inequality be explained?

Hannah Redders

UNITED NATIONS UNIVERSITY
UNU-WIDER

national treasury
Department:
National Treasury
REPUBLIC OF SOUTH AFRICA

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
IFPRI

planning, monitoring & evaluation
Department:
Planning, Monitoring and Evaluation
REPUBLIC OF SOUTH AFRICA

the dtic
Department:
Trade, Industry and Competition
REPUBLIC OF SOUTH AFRICA

SARS

TIPS

Young Scholars

This paper was produced as a part of the SA-TIED Young Scholars' programme. The programme is a part of SA-TIED's capacity building initiatives, designed to support the development of skills and capabilities in the research and policymaking aspects of economic development. Every year, the programme recruits Masters' level college students from Southern Africa to become SA-TIED Young Scholars.

SA-TIED Young Scholars work with top academics and officials in their research fields to complete original research projects as a part of the programme and as a part of the research component of their Masters' degree.

About the programme

Southern Africa –Towards Inclusive Economic Development (SA-TIED)

SA-TIED is a unique collaboration between local and international research institutes and the government of South Africa. Its primary goal is to improve the interface between research and policy by producing cutting-edge research for inclusive growth and economic transformation in the southern African region. It is hoped that the SA-TIED programme will lead to greater institutional and individual capacities, improve database management and data analysis, and provide research outputs that assist in the formulation of evidence-based economic policy.

The collaboration is between the United Nations University World Institute for Development Economics Research (UNU-WIDER), the National Treasury of South Africa, the International Food Policy Research Institute (IFPRI), the Department of Monitoring, Planning, and Evaluation, the Department of Trade and Industry, South African Revenue Services, Trade and Industrial Policy Strategies, and other universities and institutes. It is funded by the National Treasury of South Africa, the Department of Trade and Industry of South Africa, the Delegation of the European Union to South Africa, IFPRI, and UNU-WIDER through the Institute's contributions from Finland, Sweden, and the United Kingdom to its research programme.

Corresponding author: hannah.redders@googlemail.com

Regional inequality and rural dependency in South Africa.

How can opposing trends in regional inequality be explained?

Abstract.

South Africa is among the most unequal countries in the world in terms of income with a Gini of 0.65, and land with a Gini of 0.64. Inequality decreased in all provinces except for Limpopo and the Eastern Cape, that have experienced the slow pace of transition after apartheid. The study combines the South African *National Income Dynamics Study* (NIDS) panel (4 waves from 2008 – 2015) with data on regional inequality. The study shows that high inequality is found to persist under weak institutions. We find that land inequality accelerates urbanization through rural exodus. The Limpopo case is puzzling, since it exhibits paramount land inequality and relatively low but rising expenditure inequality. Given that the poor are equally poor, low levels of inequality fail to reveal anything about social mobility and equality of opportunity. A higher share of population active in subsistence agriculture significantly increases income dependency on grants and pensions. In addition, education is key in explaining economic dependency: a one percentage point increase in school attendance is associated with an average decrease of the dependency ratio by more than 6 percentage points *ceteris paribus*.

Table of contents

1   Introduction

Children have never been very good at listening to their elders, but they have never failed to imitate them.   James A. Baldwin

In Development Economics, there seems to be a great deal of excitement and academic discussion being generated concerning the question whether there could be systematic reasons for societies to evolve institutions more conducive to growth. The role of unequal distributions of endowments, wealth and income has been considered as such a systematic reason (Acemoglu & Robinson, 2008). Inequality in income and land come along with a lot of caveats for economic growth and social and political development. Literature calls attention to the potential of political destabilization, negative effects on growth, lower access two public goods and increasing crime rates - just some of the downsides of immense inequality. Income inequality is a persistent phenomenon in the case of South Africa and established through patterns of historic discrimination. South Africa is ranked one of the most unequal countries with an income Gini of 0.65 in 2015 (Maluleke, 2019) which coincides with a land Gini of 0.64 (Frankema, 2010).

Though it remains above 0.5, the income Gini coefficient for the Republic of South Africa has slowly but steadily decreased since 2008 – except for the two provinces Limpopo and the Eastern Cape. Inequality measured by the Gini coefficient increased in these two provinces between 2006 and 2015 from 0.56 to 0.61 and from 0.63 to 0.65 respectively. How can this opposing trend be explained?

The hypothesis of this paper is that high dependency, defined as the share of household income derived from remittances and social grants, is explained by the lack of social mobility, which in turn explains the increase in inequality in the respective Provinces.

South African inequality cannot be analyzed without considering persistent unemployment and some characteristics of rural poverty, such as subsistence agriculture. The novelty and motivation of this paper lies within exploring the status-quo of making a living as someone who is not integrated into the labour market and does not derive large shares of income through formal employment but through remittances and government grants - a phenomenon particularly pronounced among the rural poor. Given the gentrification strategies by the apartheid regime, land in these two provinces is extremely unequally distributed with

Limpopo standing out with a land Gini coefficient of 0.93 (Eastwood, 2006), which is one of the highest in the entire world. Land inequality translates into income inequality since land is a factor of production and identity for many black farmers (Lahiff and Manenzhe, 2012; Ladzani and Netswera, 2009; James, 2001).

Motivation lies in uncovering the following puzzle: All provinces except for Limpopo and Eastern Cape reported a decrease in their respective expenditure Gini coefficients. Surprisingly, while the Eastern Cape was the most unequal province, the expenditure Gini coefficient in Limpopo was among the lowest, despite having paramount land inequality. These observations suggest that inequality is driven by different factors in the country and that there are intra-provincial differentials in the determination of inequality.

2   Literature Review

Easterly (2007) claims South Africa as an evident example for structural inequality and that agricultural endowments predict inequality and inequality in turn predicts development. The Eastern Cape, Limpopo and Kwazulu-Natal reported the highest proportion of agricultural households within their provincial borders[1] and at the same time Limpopo (32,6%) and the Eastern Cape (32,2%) had the highest proportions of agricultural household heads with no income (Statistics South Africa, 2013). Being home of large rural population and incorporating the largest areas of former homelands or *bantustans[2]* during apartheid, these two provinces are also outstanding in terms of lacking behind in infrastructure and high levels of poverty (Statistics South Africa, 2013). These two provinces are both marked by weak capacity base, low educational and health care standards, low economic growth and a concomitant high dependence on state pensions and remittances for migrants (Ladzani and Netswera, 2009; Paumgarten et al., 2005). Even though inequality in terms of Gini might be lowest in certain regions in a national context, one has to bear in mind that it is still high and persistently high and that inequality in poor regions is lower because the poor are equally

---

[1] 35 percent in the Eastern Cape, 33 percent in Limpopo and 28 percent of households in KZN identify as agricultural households (ibid.).

[2] With the passage in 1951 of the Bantu Authorities Act, apartheid set in motion the development of ten bantustans, one of the most notorious racial ordering projects in South Africa. In an effort to legitimize the apartheid agenda and to strip black South Africans of their citizenship by establishing ten parallel "countries," the Bantustans were also known as "homelands" in official language (Laura Philipps, 2017).

poor. The persistence of poverty-related underdevelopment expresses the legacy of the apartheid system, since three of the ten former Bantustans (low-productive land assigned to black South Africans) were located in Limpopo[3] and the largest Transkei homeland in the Eastern Cape (Philipps, 2017).

Demombynes and Özler (2002) emphasize that inequality has to be analyzed on finer geographic levels than the national level, as its effects are particularly evident in local coexistence. Accordingly, the authors assess the inequality-crime-nexus among South African police station jurisdictions. Economic theory suggests a positive link between crime and inequality (Becker, 1968; Bourguignon, 2001), especially for property-related crimes (Ehrlich, 1973). In South Africa, burglary rates are indeed 20-30% higher due to local income inequalities (Demombynes and Özler, 2002). Being among the most unequal economies in the world, it is in turn no surprise that parts of South Africa are ranked among the most dangerous in the world (Newham, 2019), discouraging investment and hamper economic growth in the long-run. South Africa's overall land Gini mounts 0.64 (Frankema, 2010), with a bimodal farming sector (Wegerif and Guereña, 2020) and income Gini for South Africa was 0.65 in 2015 (Maluleke, 2019). Scholar identify a nexus between inequality in land and conflict (Couttenier and Soubeyran, 2014; Hidalgo and Nichter, 2008). In South Africa, mistrust and anger on the stagnating progress of land reform as well as severe economic shocks have translated into hate crimes, opening up a populist political discourse fired by proponents of expropriation (BBC, 2019; Newham, 2019). Farm attacks, including murder, surged by 340% between 2010 and 2019 (Claasen, 2019) and increasingly cause international media sensation (The New York Times, 2019). This does not only further undermine social cohesion but comes with an economic cost of increased fencing expenses, security and supervision costs, not to mention the psychological losses and reluctance to take risk (Claasen, 2017; Demombynes and Özler, 2002). Fear of crime is also found to foster brain drain (Dodson, 2002). Theories suggest a higher incidence of criminal behaviour in more unequal areas if income inequality is correlated with social mobility (Wilson and Daly, 1997). In addition, large inequality in land is found to rapidly disrupt efficient land governance, including rising occurrence of conflicts, and it hampers attempts to tackle environmental degradation and climate change (Wegerif & Guereña, 2020). Hence, if one also includes climate

---

[3] Namely Lebowa, Venda and Gazankulu.

Regional Inequality and rural dependency | Hannah Redders

change in the influencing factors of political destabilization, the discussion about land inequality becomes even more drastic.[4]

Adato et al. (2006) has recorded how historical social exclusion based on the racial discrimination limits new paths to social mobility by restricting people of the access to social networks beyond their own local community. Hence, even if some opportunities arise, they usually don't mobilize beyond a small geographic area. Louw et al. (2007) empirically link economic inequality and social mobility through educational data and explore the role parental education plays in the creation of children's human capital. The study finds that intergenerational social mobility within race groups improved over the 30-year-period. However, educational quality differentials among race remains substantial - thus, given the persistence of structural gentrification, also consolidates divergence among regions. Hausman and Szekely (1999) reveal that, in South Africa the educational achievement of parents is a more significant determinant of educational achievement of children than household income. The quality differential reflects in the rates of matriculation, since the likelihood to discontinue school attendance is highest with least educated parents (Louw et al., 2007; Van der Berg et al., 2016). Interestingly, especially in the rural areas it is not only parents' education that exerts a significant effect on children's wellbeing, but grandparents do as well. More than a quarter of black South African children under age five live with a pension recipient (Duflo, 2003). Analyzing the effect of this large cash transfer programme on grandchildren, disaggregated by gender, it is found that the pension had a significantly positive causal effect on anthropometric outcomes of girls (Duflo, 2003). Further, Duflo (2003) shows that observed effect was entirely driven by pensions received by women and unique for South Africa. Case and Deaton (1999) indicate that having a household head with completed secondary education may increase the educational achievement of a child by more than a quarter of a grade per year compared to children living in households led by individuals who have only completed primary education.

---

[4] The Potsdam Institute for Climate Impact Research predicts that Sub-Saharan Africa will experience major declines in crop yields at a temperature increase scenario of 1.5 until 2 degrees Celcius (Biewald et al., 2015). Couttenier & Soubeyran (2014) exploit a large data set of the Palmer Drought Severity Index, covering a period from 1946-2005. They show that drought has been a key factor in Sub-Saharan Africa's civil wars (after independence), and that the connection between drought and civil war in Sub-Saharan African countries is robust.

Expanding on labor migration models (Lewis 1954; Harris and Todaro, 1970), Oyvat (2016) causally links income and land inequality and empirically confirms that countries with egalitarian agricultural systems are predicted to experience a faster accumulation of human capital[5]. The study finds that high land inequality raises income Gini coefficients in both the urban and rural markets, not only by generating scarcity in the urban subsistence market.[6] Thus, but it is also these agrarian structures that can diminish the gap in service provision, such as education, posed by high levels of land and income Gini (OECD, 2015). In addition, the question of convergence is raised: land inequality in the 1960s has a significant positive impact on both, nowadays urban income inequality and overall income inequality. Overurbanization[7], which is triggered through unequal land access, again increases income inequality (Oyvat, 2016), creating a vicious cycle of inequality fostering inequality. Finally, countries with higher land inequalities are prone to larger shares of population engaged in low productivity urban subsistence activities (ibid.). This process materializes in an urban income Gini of 0.67 in 2008 in South Africa (Leibbrandt et al., 2010).

Gennaioli et al. (2014) empirically compare the speed of convergence of per capita income across and within 83 countries. The study finds evidence of significant barriers to factor mobility within countries. The speed of regional convergence increases with national GDP and integrated capital markets according to the authors. South Africa has been identified as a country, where migration lowers human capital in both low and high income regions (Gennaioli et al., 2014). On average, workers that choose to emigrate from one South African region to another have 1.2 more years of schooling than the natives from those regions. Consequently, the outflow of migrants tends to diminish the sending region's human capital. For the recipient region, effects are ambiguous depending on the unemployment rate and educational differentials among others (Louw et al., 2007; Oyvat, 2016). *Ceteris paribus*, the influx of migrants from other regions tends to increase the receiving region's human capital. However, this effect only holds if both regions have similar levels of human capital - an assumption likely to be breached if migration flows from poor to rich regions. Given the identified slow convergence, the effect of migration on human capital is low, but

---

[5] The nexus between human capital improvements and equalizing agriculture is also established by Easterly (2007).

[6] Data for South Africa depicts the situation in 2007-2008 and is taken from Leibbrandt et al. (2010) and World Bank Data.

[7] In the paper, 'overurbanization' is presented as a phenomenon of urban unemployment with negative spillovers and downward pressure on urban and rural wages (Oyvat, 2016).

relatively larger in the poorest regions of South Africa: The education differential in the bottom quartile of regional GDP per capita yields 0.09 more years of education (Gennaioli et al., 2014). In contrast, residents of rich regions in South Africa (in the top quartile of regional GDP per capita) have 0.01 less years of schooling than the natives of such regions (ibid,).

In a nutshell, inequality delimitates opportunities to access all three factors of production. Human **capital** is affected through the lack of social mobility and brain drain is a higher risk in the poorest provinces. The differentials in educational quality translate into a lack of opportunities in the **labour** market and, given low opportunities in the former homelands, foster urbanization. However, with unemployment being structurally high, a phenomenon of urban unemployment with negative spill overs emerges, coined as 'overurbanization' (Oyvat, 2016). The link between endowments and development of institutions is historically pronounced in South Africa and consolidates in **land** inequality with overlapping land rights in traditional authority areas (Cotula et al., 2009; Liversage, 2011). The overlapping land claims have been found to create conflict and constrained the land transition underlying the problems with crime and violence in South Africa today (Beinart et al., 2017). Consequently, it is imperative to understand the factors that create inequality, particularly where persistent historical institutions impede social mobility.

3   Data

3.1   Data Sources

The National Income Dynamics Study (NIDS) is the nationally representative panel study in South Africa (Department for Planning, 2020). For consistency with other data sources, this analysis covers the first four waves of NIDS, which have been conducted (1) 2008, (2) 2010/2011, (3) 2012, (4) 2014/2015. In this survey, a multi-stage sample method was used which is based on a stratified approach with probability proportional to the size range of primary sampling units in the first stage and the sampling of dwelling units with systematic sampling in the second (Brophy et al., 2018). The merged dataset compiles 42,562 observations on a household level and the individually derived NIDS covers 142,820 observations, which have been averaged to enable the analysis on a Provincial level.

The survey was further stratified by geography after redistribution of the survey to the provinces (primary stratification). Geographical information is based on the 2011 District

Council codes (Brophy et al. 2018). Data on Provincial inequality derives from Maluleke (2019), who uses primary data inputs from data which Stats SA officially collected and published, more precisely the Income and Expenditure Survey (IES), Living Conditions Survey (LCS), General Household Survey (GHS), and Quarterly Labour Force Survey (QLFS). Calculations on Provincial inequality are based on Statistics South Africa IES and LCS and matched in the best possible manner based on the time proximity to the NIDS waves. The same applies to crime data, which was taken from the South African Police Serve Crime Statistics (South African Police Service (SAPS), 2018). The report disentangles Police reports granularly by the type of crime and covers observations on a Provincial level.

3.2 Variable Measurement

This paper aims to connect Provincial inequality and its economic manifestation in terms of income dependency on remittances and social grants (pension, child support), resulting in two different dependent variables. The first variable to be explained is the income Gini coefficient based on expenditure per capita (Maluleke, 2019). The Gini is chosen because its intuitive interpretation and, given that vast amounts of observations are averaged to create Provincial observations, the Gini index is the most robust to data contamination or outliers (Klasen et al., 2016). The dependency ratio is calculated as the sum of remittances and government grants over the total household income, based on NIDS data:

$$\overline{dependency_{it}} = \frac{remittances_{it} + childgrant_{it} + pension_{it}}{total\ household\ income_{it}}$$

In the baseline model the following variables are used to explain the dependency ratio: *agri_pop* refers to the population active in agriculture in the past 12 months, without calling it a professional employment (derived from 'h_ag', NIDS). *Homeland* measures the share of population that lives in geographic units, former homalands (banthusthans), that are now defined as Traditional Authority Area: 'Communally- owned land under the jurisdiction of traditional leaders; settlement within these are villages' (Brophy et al., 2018: p.68). *School_attend* refers to the share of learners aged 6-18 years, based on the General Household Survey (GHS) (Maluleke, 2019).

The augmented model adds the share of *bank account* holders per province as well as the perceived *health* status, both based on NIDS data. The health status is ranked '1' for 'excellent', '2 – very good, '3 – good', '4 – Fair' and '5 – Poor'.

For descriptive statistics and robustness checks, further variables are used: crime ratio refers to the Provincial ratios of 'serious crime' (South African Police Service (SAPS), 2018). *Prop_crime* captures the all property related crime incidents (mainly break-ins and theft located at the own residency), adjusted for the population share of each province (Maluleke, 2019). Other variables, that capture the ownership of assets (e.g. cell phone ownership) are derived from NIDS.

3.3 Empirical Strategy

A credible regression model to explain the variation in Gini is complicated by two challenges. First, there are four time periods and nine provinces, resulting in small T and small N. In addition, the Gini coefficient shows little variation, which makes it difficult to estimate it using fixed effects estimation, because the within estimator is carried out through time demeaning, further decreasing the necessary variation in the variance matrix (Wooldridge 2010). However, it is imperative to at least be able to assess unobserved heterogeneity in a panel data framework, which random effects models assume to be orthogonal to the error term, a somewhat strong assumption. In other words, Fixed-effects will not work well with data for which within-cluster variation is minimal or for slow changing variables over time, such as the Gini coefficient.

For the given reasons, the paper proceeds with a descriptive approach to explaining inequality and using regression analysis to model the dependency ratio, which has been identified as a potential driver of increases in inequality and as a particularity of the two provinces, where inequality increased (see Figure 1).

The model specification is

$$dependency_{it} = \beta_0 + \beta_1 homeland_{it} + \beta_2 agricultural\_pop_{it} + \beta_3 school\_attend_{it} + \beta_i X_{it} + c_i + u_{it} ,$$

where X_{it} is a vector, which contains further control variables. For the random effects models, we assume orthogonality between the time-invariant unobserved heterogeneity c_i and the regressors, which is why it can be specified within a composite error term $v_{it} = c_i + u_{it}$, assuming strict exogeneity in conditional means $E(u_{it} \mid c_i, x_i) = 0$. Fixed effects models model the unobserved heterogeneity c_i, allowing for arbitrary correlation with the regressors. Derived from the same sources described above, additional control variables that were included in the model were unemployment, cell phone access, property crime rates, assault

ratio, share of households that experienced crop failure, median income of the population, household size and median education of the adult population, perceived health status. The variable 'bank account' refers to the share of population owning a bank account and captures the fact that a formal account is needed to receive government grants. From the NIDS interview language variable, an ethnolinguistic diversity score was created counting all the different languages spoken in the provinces in order to cater the finding that economic development can be explained by ethnic diversity (Easterly and Levine, 1997). However, except for the perceived health status and the bank account share, none of the control variable entered significantly nor improved the fit of the model. The baseline model and the augmented model with the best fit are further explained under 4.

4   Results and Discussion

4.1   Descriptive Results

Remittances have been identified as a potential driver of positive changes in the Gini coefficient, establishing a link to the puzzling inequality increase between 2006 and 2015 in Limpopo and the Eastern Cape amid an overall decline (see Figure 1).
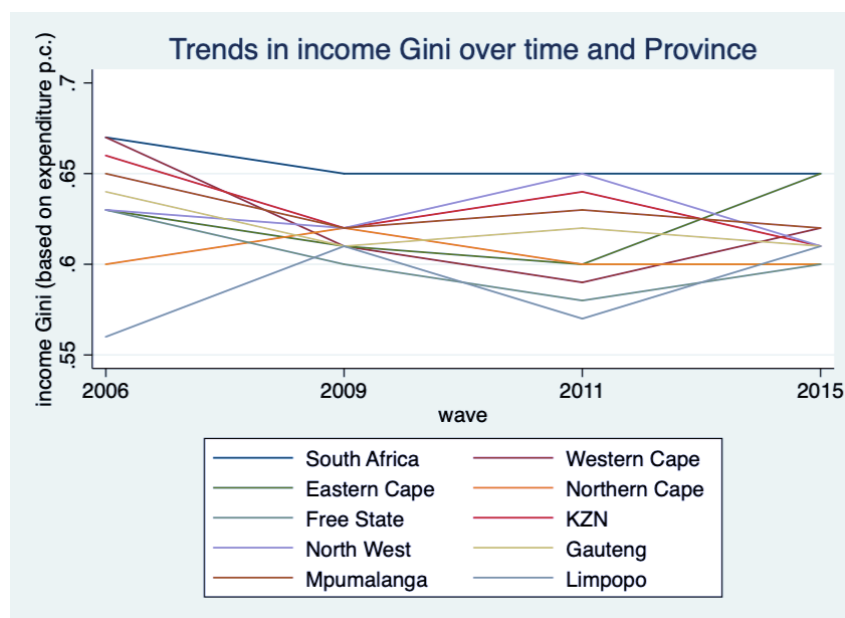

Figure 1. Trends in income inequality province comparison and national aggregate over time (2006-2015). Own elaboration based on Makulele 2019.

Economic realities are heterogeneous among South Africa (see Figure 2). The annual median income, which is already less susceptible for outliers vary by more than 15,000 ZAR

between rich and poor provinces and is lowest in the provinces where structural racial disadvantage was implemented in form of bantustans. To put the monetary, diverge into perspective: the annual median income in South Africa amounted 13,546 ZAR in 2015 (Maluleke, 2019).
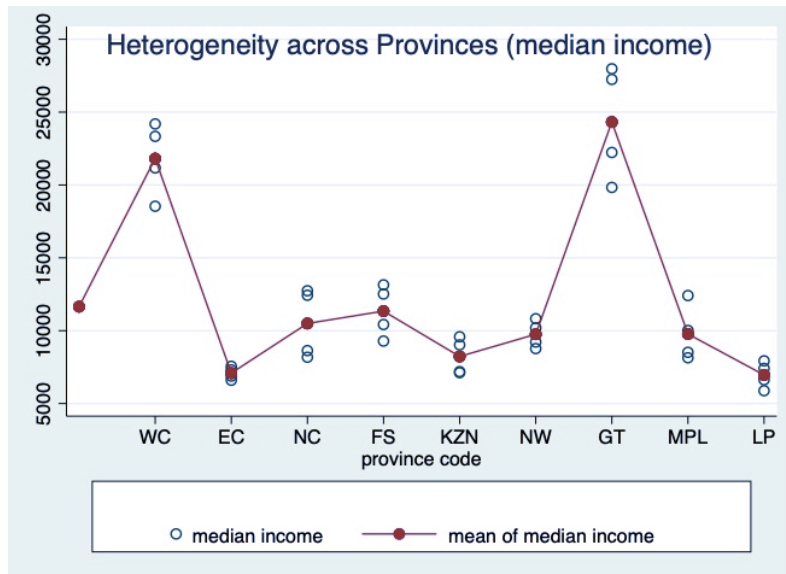

Figure 2. Heterogeneity in median income among provinces. Own elaboration based on Maluleke, 2019.

Figure 1 confirms a similar heterogeneity among the trends in income inequality among provinces. To shed light on drivers of Provincial Gini, correlation coefficients between are useful. Inequality, measured in terms of the income Gini, correlates to varying degrees and also in different directions with the variables at national versus provincial level. There are consistently positive connections between the Gini and the dependency ratio, the distance to water sources from dwellings und the unemployment rate. Where households are larger and large share of household income is derived from remittances, the income inequality tends to be consistently lower. None of the relationships described lay claim to causality. Compared to the national level, some variables exhibit inconsistent directions (+/-) compared to variable correlations measured at a Provincial level. For instance, Provincial inequality yields a negative correlation with the share of agricultural population, whereas the national correlation coefficient is positive and five times larger. Property related crime seems to be important to explain aggregate inequality and, surprisingly, correlates negatively with it.

For Limpopo province only, strong positive correlations[8] are scored between the Provincial Gini coefficient and the share of black population (0,9902), household size (0.9635), school attendance rate (0.9177), the share of elderly (0.9059), the average asset score (0.8660). Strong negative correlations exist between Limpopo's income inequality and the share of native Afrikaans speakers (-0.9889), households with access to a cell phone (-0.9856), the share of white population (-0.9665), the share of bank account holders (-0.9575), the ratio of assault crimes (-0.9382), and the ethnolinguistic diversity score (-0.8660).

In the Eastern Cape strong positive correlations are scored between the Provincial Gini coefficient and the share of English speakers[9] (0,9765), median income (0,9728), the unemployment rate (0,9397), the agriculturally active population (0,9192), the urban population share (0,8835), cell phone access in a household (0,8814), the share of former homeland residents (0,8648), and the elderly population ratio (0,8458). Strong negative correlations exist between the Eastern Cape's income inequality and property crime ratio (-0,9803), the share of population living on farms (-0,9398), social grant recipients (-0,9232), assault ratio (-0,8348), household size (-0,822) and the share of black population (-0,8129). At first glance, the directions of the correlation are in line with the assumption of low inequality among the poor.
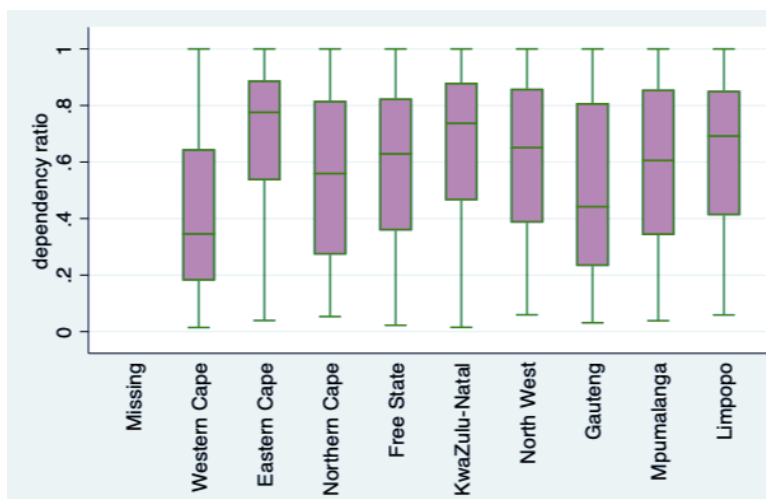

Figure 4. dependency ratio is defined as the share of total hh income that is derived jointly from remittances and government grants (pensions, child support). Own elaboration based on NIDS waves 1-5.
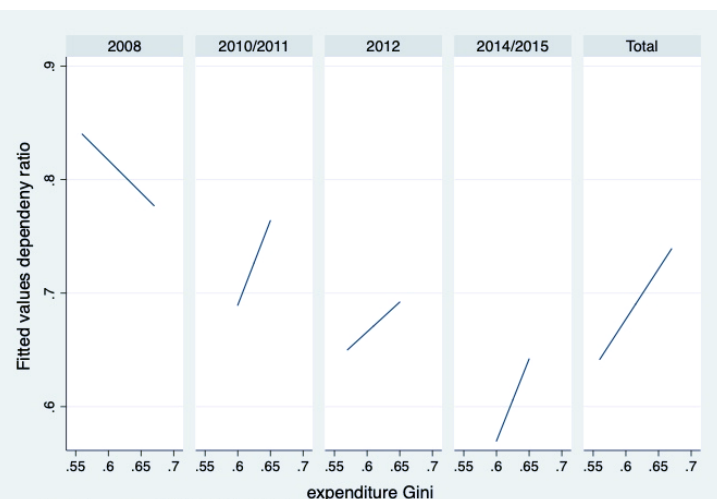
Figure 3. linear fit between dependency ratio (mean over provinces and years) based on NIDS and the regional expenditure Gini (Maluleke, 2019). Own elaboration.

---

[8] Above a 0.8-correlation-coefficient.

[9] Refers to English understood as preferred or first household language. See Data section and variable table in the appendix for details on language variables.

Figure 3 visualizes that the Eastern Cape and Limpopo exhibits an above average dependence on remittances and social grants and that even the low quintiles that generate some other income are relatively highly dependent.

Given that poor households are disproportionally affected by increases in food prices, land inequality is especially harmful for subsistence farmers (Statistics South Africa, 2013). The relationship between the dependency ratio and regional expenditure Ginis has turned around after the food price spikes in 2008[10] (Figure 4). Despite the increase in the dependency ratio for all South Africa over time, the linear fit between regional expenditure Gini coefficients and fitted values for the dependency ratio has become sharper. Ignoring any causal interpretations, Figure 4 confirms the findings from previous literature. There is an increasingly strong relationship between rural dependencies and regional inequality, which coincide with food price increases.

4.2 Econometric results

Starting off with a basic random effects model with robust standard errors, including only *remittances* and *property crime ratios* to explain the change in the Gini coefficient, we find that the latter two variables alone significantly explain 27% of the variation in Provincial and aggregate income inequality. On average, a 1% increase in remittances is associated with a decline of income inequality by 15,12% ceteris paribus. The nexus may not be causal, but according to the chi^2 statistics, the model is correctly specified at a 1% level and unbiased under RE.1- RE.2 (Wooldridge 2010).[11] Based on the link between inequality and external non-wage sources, linear random and fixed effects models are used to dig deeper into what constitutes dependency.

The regression equation

$$dependency_{it} = \beta_0 + \beta_1 homeland_{it} + \beta_2 agricultural\_pop_{it} + \beta_3 school\_attend_{it} + c_i + u_{it} ,$$

is first estimated by Pooled Ordinary Least Squares (POLS), which ignores the existence of unobserved heterogeneity ($c_i = 0$). All coefficients are statistically significant at a 5% level or higher, except for the share of agricultural population. The three explanatory variables

---

[10] In 2007–2008, the nominal prices of almost all food commodities increased by more than 50%. For an analysis of potential drivers see for example Tadasse et al. (2016).

[11] However, a Breusch-Pagan-Test reveals that it might be inconsistent due to either serial correlation or the lacking random effects structure in the variance matrix of the composite error term v_{it}.

alone already explain 43% of the variation in the dependency ratio, with only a 4.9% divergence between the R^2 and its adjusted version. A 1 percentage point increase in the share of population living in traditional authority areas (former 'homelands'), is associated with a 0.352 percentage points increase in the dependency ratio ceteris paribus. If the share of learners ages 6-18 increases in a province by 1 percentage points the dependency ratio on average decreases by 2.17 percentage points ceteris paribus.

If c_i is correlated with any element of the regressors, then pooled OLS is biased and inconsistent (Wooldridge, 2010). The Breusch-Pagan-test uses the Lagrange multiplier principle to test for heteroskedasticity (Breusch and Pagan, 1980) and thus enables to opt for either POLS or random effects (RE). Their test assesses the hypothesis $Var(c_i)=0$, meaning that the unobserved heterogeneity is the same across all individuals (Park, 2009). Here, we reject the H_0 at 5% level of significance, thus there is a significant unobserved heterogeneity across individuals, that we need to take into account. Since the POLS does not make any specifications about the structure of the error term, we would be better off with a RE model. Breusch-Pagan-test is not only understood as a test for heteroskedasticity, but also known as a 'test for random effects' (Park, 2009). Indeed, rejecting the null hints at the random effects structure of the variance matrix. However, this test statistic detects many forms of serial correlation in the composite error term (Hill et al., 2018; Wooldridge, 2010). Thus, only because we reject H_0, the variance matrix might not necessarily exhibit the random effects structure, which claims for using robust standard errors.

As a second step, the model is estimated using RE. In this more efficient estimation, all variables turn out to be statistically significant at a 1% level of significance and marginal effects increase. Ceteris paribus, a 1 percentage point increase in the population that has been active in subsistence agriculture lowers the dependency ratio on average by 0.3 percentage point ceteris paribus. Comprising a larger share of dwellers in former homelands within its borders, increases the economic dependency from external sources of people living in the respective province significantly. School attendance yields an opposite effect: All else equal, where school attendance increases by 1 percentage point, dependency decreases by approximately 5 percentage point on average.

To overcome the restrictions of the Breusch-Pagan test, Wooldridge (2010) suggests using tests that precisely test for random effects form of the variance-covariance matrix of the composite error term (see Baltagi and Li, 1995). The Baltagi-Li (1991) joint test for serial

correlation and random effects suggests that the possible misspecification is more likely due to the presence of serial correlation than random effects. Serial correlation can of course be created through the time-invariant part of the error term, thus the unobserved heterogeneity.

Thirdly, to allow for arbitrary correlation between the unobserved heterogeneity and the regressors, fixed effects estimation is applied to the model, yielding marginally larger and statistically significant coefficients among all three regressors. The directions of the effects are consistent among all models and according to chi-square- and F-test regressors are jointly significant and models significantly well specified. Especially the coefficients for 'homeland' and 'school attendance' gain importance in the fixed effects model and the within-estimation increases the fit to 75.1% of the variance in the dependency ratio being explained by the three regressors. Hausman (1978) proposed a test based on the difference between the two sets of coefficients estimated by RE and FE. Here, the null is rejected at a 5% level of significance, RE is not consistent and we use FE.[12] Thus, the Hausman-test confirms the results of the adjusted tests for random effects and the fixed effects model is unbiased and consistent under the necessary assumptions including strict exogeneity (Wooldridge, 2010).

Table 2 in the appendix contains the regression results of the augmented model

$dependency_{it} = \beta_0 + \beta_1 homeland_{it} + \beta_2 agricultural\_pop_{it} + \beta_3 school\_attend_{it} + bank\ account_{it} + health\ perception_{it} + c_i + u_{it}$.

Linear-Multiplier tests for heteroskedasticity and random effects are repeated, as well as the Hausman-test, leading again to a confirmation of substantial unobserved heterogeneity, making the fixed effects model the preferred option. Substantively, fixed-effects models are designed to study the causes of changes within a province.

The theta[13] of the random effects model is 0.51, thus in-between zero and one, which can be interpreted as a medium importance of time-invariant factors.

---

[12] Since FE is consistent when c_i and x_it are correlated, but RE is inconsistent, a statistically significant difference is interpreted as evidence against the RE assumption RE.1(b). Formally, the is b-B = 0 (diff between the two sets of coefficients). The first set of coefficients (b) is consistent under the Null and under the alternative Hypothesis (fixed effects in our case). The second set of coefficients (B) Is only consistent under the Null but more efficient than the other set.

[13] Adding the theta option in Stata thus tells us how much of the variation is due to the individual, time-invariant effect, since it is calculated based on rho. If theta is close to zero, the estimated coefficients of RE are

Comparing the three augmented models, the share of population, not professionally but for other purposes involved in agriculture gains significance and magnitude. In the pivotal fixed effects model, a 1 percentage points increase in this share is associated with an average decrease of 0.38 percentage points in dependency ceteris paribus. Subsistence agriculture is key in explaining economic dependency and seems to balance out the foregone income from external sources, e.g. family members usually sending money from cities facing a cash crunch.

Despite being statistically significant in both, the POLS und RE model, the own health perception is not significant any more in the fixed effects model. School attendance share is highly significant and a 1 percentage point increase in the latter, relates to an average decrease in dependency ratio of 6.14 percentage points, holding all else constant. Education turns out to be crucial in explaining economic dependency. Provinces with higher shares of learners aged 6-18 attending school appear to be able to create economic opportunities rather than deriving larger income from external, non-productive sources. This interpretation goes along with Case and Deaton (1999), as well as Louw et al. (2007): Dropping out of school is more likely for students from poor families – such as highly dependent households - facing constrained time out of economic necessities.

Of course, part of the dependency ratio is also generated through pensions and child support. The findings can also be interpreted in a way confirming the effect of pensions for child welfare as identified by Duflo (2003). Before they attend school, a quarter of black South African children live with a pension recipient. The empirical results suggest, that once they grow older, children (together with their parents) and adolescents migrate to other areas. If the migration route coincides largely with homeland-out-migration, we could explain the increase in significance for the school attendance, taking up the decrease in significance for the homeland population share. At the end of the day, the root cause of both effects is the same: Lack of opportunities in former homelands, resulting in skewed age structures in the rural communities with elderly, mostly female subsistence farmers and small children left behind (Lahiff & Manenzhe, 2012; Duflo, 2003).

---

similar to those of the pooled OLS, if it is unity, we actually estimate a fixed effects model (see Woldridge, 2010).

The results suggest evidence for long term patters of land inequality as drivers of changes in Provincial economic opportunities, which indirectly relate to income inequality. The share of people living in former homelands is statistically significant in all models, however the significance decreases to the 10% level in the fixed effects model with the highest explanatory power. Fixed effects models are designed to study the causes of changes within an entity. The share of homeland dwellers is quite time-invariant over time with an average change of less than 1% for Limpopo and the Eastern Cape. Opposed to that, especially between 2012 and 2015, the share of agriculturally active households has increased.

5 Conclusions

South Africa is afflicted with structural unemployment, high crime rates and persistent inequality. It is imperative to understand that these factors are deeply intertwined. For policymakers this opens up a variety of adjusting levers within the toolbox in order to tackle the lack of social mobility. Economic welfare is shaped by a large part at the local level. The study shows that inequality measures at national levels accordingly leave out a lot of the picture, drawn to understand drivers of inequality. A more granularly understanding of local contexts is key to realize first of all, what drives inequality in regions differently and secondly, to what share these drivers are structural patterns. Even achieving low levels of inequality has to be taken with a pinch of salt, because given that the poor are equally poor, low levels of inequality fail to reveal anything about social mobility and equality of opportunity or the prevalence of poverty. An increase in income inequality among the poorest provinces, however, is even more alarming, as it points to increasing inequality among the poor. It is precisely this fact that gives rise to concern in the face of rising trends for the Eastern Cape and Limpopo. South Africa has been identified as a country, where migration between provinces lowers human capital in both low- and high-income regions. Rural exodus of certain age groups spurs regional inequality through the channel of rural dependency and brain drain.

Little economic opportunities other than subsistence agriculture have been created in the former homelands. There is even ground for the concern that social grants maintain the structural underdevelopment active, since they are almost entirely responsible for sustaining the livelihoods in these regions. Without claiming causality, the strong relationship between rural dependencies and regional inequality has found to be increasing.

Education and infrastructural development are found to be key for the current disadvantaged generation to overcome historical disadvantage. Correspondingly, dependency on remittances and social support is found to be significantly lower, where the share of learners increases. The empirical analysis confirms that social mobility being particularly dependent on parental education. If living with a household head with only a primary school degree lowers children's achievements by more than a quarter of a grade per year compared to parents with secondary schooling. Historical trends of racism engender discrimination. An ensuing scenario in the former homelands, is that of children growing up with their grandparents which usually leaves these children deprived of their opportunities.

6 References

Acemoglu, D., & Robinson, J. A. (2008). Persistence of power, elites, and institutions. American Economic Review, 98(1), 267-93.

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

Afrobarometer Data. (2020). Rounds 1-7, available at http://www.afrobarometer.org.

Alvaredo F, Chancel L, Piketty T, et al. (2018). World Inequalit y Report 2018. World Inequality Report 2018. World Inequlity Lab.

Baltagi, B. H., & Q. Li. (1991). 'A Joint Test for Serial Correlation and Random Individual Effects, Statistics and Probability Letters 11, 277–280.

Baltagi, B. H., & Q. Li (1995). Testing AR(1) Against MA(1) Disturbances in an Error Component Model, Journal of Econometrics 68, 133–151.

BBC. (2019, April 30). *Julius Malema -South Africa's radical agenda setter*. https://www.bbc.com/news/world-asia-pacific-14718226.

Beinart, W., P. Delius, & M. Hay. (2017). Land Rights, A Guide to Restitution and Tenure Reform in South Africa. Johannesburg: Jacana.

Breusch, T. S., & A. R. Pagan. (1980). The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. Review of Economic Studies, 47(1):239-253.

Brophy, T Branson, N . Daniels, R.C. Leibbrandt, M. (2018). *National Income Dynamics Study Panel User Manual*. https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/712/download/9561.

Department for Planning, M. and E. R. of S. A. S. A. L. and D. R. U. of C. T. (2020). *The National Income Dynamics Study (NIDS)*. http://www.nids.uct.ac.za/.

Gennaioli, N., La Porta, R., Lopez De Silanes, F., & Shleifer, A. (2014). Growth in regions. *Journal of Economic Growth*, *19*(3), 259–309. https://doi.org/10.1007/s10887-014-9105-9.

James, D. (2001). Land for the Landless: Conflicting Images of Rural and Urban in South Africa's Land Reform Programme, Journal of Contemporary African Studies, 19:1, 93-109, DOI: 10.1080/02589000125070

Ladzani, W. & Netswera, G. (2009). Support for rural small businesses in Limpopo Province, South Africa. Development Southern Africa, 26(2), pp.225-239.

Lahiff, E., Davis, N., & Manenzhe, T. (2012). Joint ventures in agriculture: Lessons from land reform projects in South Africa.

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.

Louw, M., van der Berg, S., & Yu, D. (2007). Convergence of a kind: Educational attainment and intergenerational social mobility in South Africa. *South African Journal of Economics*, *75*(3), 548–571. https://doi.org/10.1111/j.1813-6982.2007.00137.x

Maluleke, R. (2019). Inequality trends in South Africa. In *Statistics South Africa* (Issue October). http://www.geocurrents.info/economic-geography/inequality-trends-in-south-africa

Newham, G. (2019). *Republic of South Africa Systematic Country Diagnostic Worldbank Background Note : Crime , Exclusion in South Africa Society*.

Organisation for Economic Cooperation and Development (OECD). (2015). South Africa Policy Brief: Regional, urban and rural development. www.oecd.org/policy-briefs.

Oyvat, C. (2016). Agrarian Structures, Urbanization, and Inequality. *World Development*, *83*, 207–230. https://doi.org/10.1016/j.worlddev.2016.01.019

Park, H. M. (2009). *Linear Regression Models for Panel Data Using SAS, Stata, LIMDEP, and SPSS. Mathematical Computing Department. Indiana University. 30*(812), 23–30.

Paumgarten, F., Shackleton, C., & Cocks, M. (2005). Growing of trees in home-gardens by rural households in the Eastern Cape and Limpopo Provinces, South Africa. *International Journal of Sustainable Development and World Ecology*, *12*(4), 365–383. https://doi.org/10.1080/13504500509469647.

Philipps, L. (2017). History of South Africa's Bantustans. In *Oxford Research Encyclopedias. African History.* Oxford Research Encyclopedias. https://oxfordre.com/africanhistory/view/10.1093/acrefore/9780190277734.001.0001/acrefore -9780190277734-e-80.

South African Police Service (SAPS). (2018). *Crime Statistics 2017/2018.*

Spaull, N., Van der Berg, S., Wills, G., Gustafsson, M., & Kotzé, J. (2016). Laying firm foundations: Getting reading right. Report to the PSPPD. University of Stellenbosch.

Statistics South Africa (StatsSA). (2013). *Census 2011 Agricultural households Key Highlights Report. Report No.*, 3–30. http://www.statssa.gov.za/publications/Report-03-11-01/Report-03-11-012011.pdf.

Statistics South Africa (StatsSA). (2020). *Stats SA releases Census of Commercial Agriculture 2017 Report*. http://www.statssa.gov.za/?p=13144.

Tadasse, G., Algieri, B., Kalkuhl, M., & Von Braun, J. (2016). Drivers and triggers of international food price spikes and volatility. In *Food price volatility and its implications for food security and policy* (pp. 59-82). Springer, Cham.

Van der Berg, S., Spaull, N., Wills, G., Gustafsson, M., & Kotzé, J. (2016). Identifying binding constraints in education. Available at SSRN 2906945.

Wegerif, M. C. A., & Guereña, A. (2020). Land inequality trends and drivers. *Land*, *9*(4), 1–23. https://doi.org/10.3390/land9040101.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Table 1: Dependency Regressions - POLS

	(1)	(2)	(3)
	RE and FE		
VARIABLES	dependency_ratio	dependency_ratio	dependency_ratio
agri_pop	-0.142	-0.299***	-0.393***
	(0.128)	(0.101)	(0.0394)
homeland	0.352***	0.466***	1.543***
	(0.066)	(0.121)	(0.370)
school_attend	-2.170**	-4.980***	-6.178***
	(0.957)	(0.748)	(0.915)
Constant	2.638***	5.270***	6.049***
	(0.896)	(0.713)	(0.906)
Observations	36	36	36
R-squared	0.480		0.751
Number of provincecode		9	9
Overall Rsq		0.420	0.464
Within Rsq		0.723	0.751
Between Rsq		0.334	0.610

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2: Extended Dependency Regressions – POLS RE and FE

	(1)	(2)	(3)
VARIABLES	dependency_ratio	dependency_ratio	dependency_ratio
agri_pop	-0.228*	-0.326***	-0.382***
	(0.122)	(0.0606)	(0.0222)
homeland	0.356***	0.426***	0.897*
	(0.0618)	(0.110)	(0.417)
school_attend	-1.675*	-3.788***	-6.135***
	(0.962)	(0.804)	(1.100)
health	0.192**	0.139**	0.109
	(0.0912)	(0.0571)	(0.0659)
bank_acc	0.585	0.226	-0.338
	(0.452)	(0.277)	(0.353)
Constant	1.427	3.724***	6.161***
	(0.964)	(0.791)	(1.135)
Observations	36	36	36
R-squared	0.585		0.776
Number of provincecode		9	9
Overall Rsq		0.513	0.490
Within Rsq		0.719	0.776
Between Rsq		0.445	0.520

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1