Distribution dynamics of regional income per worker in South Africa, 2001-2011.

Gibson Mudiriza[1]

Supervisors

Lawrence Edwards and Asha Sundaram

Abstract.

This paper examines the distribution dynamics of regional income per worker in South Africa, a country with one of the highest inequality in income distribution. While existing evidence has made use of the classical convergence approach, we employ the distribution dynamics approach using the 2001 and 2011 census data. The paper also contributes to the literature by identifying the role of space, which has generally been ignored in previous studies. Using univariate kernel density estimates to explore the shape dynamics of the entire cross-section distribution, we find that regional dispersion in income per worker decreases with time. The decrease was driven by increases in relative income per worker of poorer regions, as well as slight decrease of rich regions. Finally, we use a spatially conditioned scheme to assess the role of space in the distribution dynamics. The results suggest that space is important in explaining the convergence process and its effects persist overtime.

Key words: Regional income dispersion, distribution dynamics, geographical location.

JEL classification: D31, C14, R12.

1. Introduction

The issue of regional income disparities has been the subject of a lot of empirical research for developed and developing nations alike. Consequently, a large body of empirical literature has accumulated, initially applying the classical convergence approach (Barro & Sala-i-Martin, 1992; Meliciani & Peracchi, 2006; Pfitzner & Lang, ; Young, Higgins, & Levy, 2008) and later using the distribution dynamics approach (Ezcurra, Pascual, & Rapún, 2007; Johnson, 2000; Juessen, 2009; Magrini, 1999) While offering valuable insights, the geographical dimension of the data underlying empirical analysis of regional income disparities is largely ignored.

[1]PhD Candidate, School of Economics, University of Cape Town.
mdrgib001@myuct.ac.za

Regions are treated as independent entities, as if their geographical location and potential interregional linkages did not matter. However, theoretical insights notably from the new economic geography literature (Fujita et al., 2001; Helpman, 1998; Krugman, 1991), suggests that spatial dependence between regions can be critical in explaining the distribution dynamics of regional incomes. As such modelling and incorporating spatial dependence is of paramount importance, as it relaxes the unrealistic independence assumption and highlights the importance of geographical location in the distribution dynamics for connected economies. In this paper, we examine the distribution dynamics of income per worker, paying particular attention to spatial dependence (geographical location of one region relative to other regions).

The concept of spatial dependence refers to the absence of independence between geographic observations, and is defined as the correlation of a variable with itself in space (Anselin & Bera, 1998). The correlation can arise in a number ways. Economic interactions through mechanisms such as movement of factors of production, interregional trade, technology and knowledge spillovers as well as institutions can tie economic fortunes of connected regions (Buettner, 1999; Rey & Janikas, 2005). Furthermore, spatial dependence can be an outcome of measurement error and omitted variables. This implies that, economic outcomes for a given region, such as income, may be correlated with those of its neighbouring regions, meaning that incomes of connected regions tend to influence each other directly or through common shocks (Anselin & Rey, 1991).

Indeed, empirically, spatial dependence is acknowledged to play a critical role in regional income disparities and convergence process (Maza & Villaverde, 2009; Maza, Villaverde, & Hierro, 2009; Rey & Montouri, 1999; Rey, 2001). And, ignoring its presence can have serious implications on both the classical convergence approach and the distribution dynamics approach as it can invalidate their inference basis (Rey & Janikas, 2005). Apart from being an empirical issue, as it corrects for a possible statistical bias, integration of spatial dependence is also seen as a real empirical issue with a basis in theory (Figleton, 2003b). It provide a basis for evaluating the neoclassical theories, as well as the new economic geography theory.

The recognition of the importance of spatial dependence in the analysis of regional income disparities has seen several studies employing the classical convergence approach incorporating it in recent years. To date quite an extensive literature now exist (Bosker, 2007; Dall'Erba & Le Gallo, 2008; Ertur, Le Gallo, & Baumont, 2006; Rey & Montouri, 1999). (Abreu et al., 2004; Magrini, 2004; Rey & Janikas, 2005) have provided a survey of this literature. These studies have generally confirmed the need to integrate spatial dependence into the classical convergence approach. Although quite illustrative, the classical convergence approach has been criticised on several grounds and its appropriateness has generally been questioned. The approach not only present several econometric problems (Bosker & Krugell, 2008), but it also fails to capture some potentially interesting features of the distribution dynamics of the entire cross-sectional distribution (D. T. Quah, 1996b) (Fischer & Stumpner, 2008; Magrini, 2004; D. T. Quah, 1996a)[2].

As an alternative approach Quah (1993a; Quah, 1993b) proposed using the distribution dynamics approach, which uses non-parametric techniques. The approach is seen as empirically attractive as it examines the entire cross-sectional distribution dynamics without imposing prior restrictive assumptions as done by classical approach (Bosker & Krugell, 2008; Magrini, 1999). However, unlike the case of the classical approach to convergence, where the issue of spatial dependence has attracted considerable attention, its importance and implications has generally been ignored in the distribution dynamics approach. Only a few studies among them (Rey, 2001; Rey, 2004) for U.S, (Bosker, 2009; Le Gallo, 2004; Maza & Villaverde, 2005; Quah, 1996b) for E.U have accounted for space within the distribution dynamics approach. All these studies gave overwhelming evidence of the importance of space.

Despite the recognition of the importance of space, quite paradoxically, there is very little research accounting for space dependence within the distribution dynamics approach among SSA countries. Only a handful of studies exist and even fewer use data at sub-national level among them,(Bosker & Krugell, 2008). Thus, despite the significant differences in incomes

[2] It suffers from the Galton's fallacy problem(D. Quah, 1993b), and also fails to reveal such features as spatial aspects/core-periphery (Fujita, Krugman, & Venables, 2001; Krugman, 1991), polarisation and club convergence (D. T. Quah, 1996a) Moreso, it does not show the shape dynamics of the distribution of income across regions and how the shapes evolves overtime and the mobility of regions across the distribution (Rey & Janikas, 2005).

within countries and evidence of the importance of geographical location, studies examining distribution dynamics of income at the sub-national level, later alone incorporating space are still scarce in SSA. There is therefore a huge gap in the literature which must be addressed in order to provide a broader picture of regional income dispersion in SSA, especially at sub-national level.

Against this background, this paper is an attempt to fill this gap, by examining the distribution dynamics of regional income per worker at sub-national level for a unique SSA country, South Africa. Apart from being one of the most unequal economies in the world, South Africa is an interesting case study for a number of reasons. First, its regional imbalances are a product of a complex set of factors and what is more certain is that history has had an important role in the generation and reproduction on these imbalances in space and time. While geographic factors determined the initial unequal development of the economy, the legacy of apartheid has had and still has overwhelming implications on these disparities. Concentration of income and wealth and the low levels of education prevailing today are, to a large extent, a legacy of apartheid racial oppressive policies.

Second, while personal and household income inequality[3] has been and remains the major focus for empirical research in South Africa, its regional income imbalances have received little attention. Only a handful of researchers have decided to get a deeper insight into South Africa's sub-national income disparities using both the classical convergence approach (W. F. Krugell, 2005; Naudé & Krugell, 2003; Naudé & Krugell, 2006) and distribution dynamics approach (Bosker & Krugell, 2008; W. F. Krugell et al., 2005).While we learn a lot from these studies, with the exception of (Bosker & Krugell, 2008) these studies have largely neglected the role of space in the explaining regional income disparities and convergence. Considering that (Bosker & Krugell, 2008) did find evidence of spatial dependence, results from other studies can be questionable given that analysis in all these studies was done at the same geographical level (magisterial districts).

[3] These include studies by (Bhorat & Kanbur, 2005; Hoogeveen & Özler, 2006; Leibbrandt, Levinsohn, & McCrary, 2005; Leibbrandt, Levinsohn, & McCrary, 2010; Leibbrandt, Wegner, & Finn, 2011; Leibbrandt, Finn, & Woolard, 2012; Van der Berg, Burger, Burger, Louw, & Yu, 2006; Van der Berg, Louw, & Yu, 2008; Van Der Berg, 2011; van der Berg, 2014; Wittenberg, 2014).

Following the spirit of (Bosker & Krugell, 2008) study, the aim of this paper is to contribute to improving our understanding of the spatial dimensions of dispersions in income per worker in South Africa. To do so, the paper begins by developing a regional database from the 2001 and 2011 national census datasets that is geographically consistent overtime. This database is then used to answer the following specific question: How dispersed are regional income per worker in South Africa and how has it changed overtime? Accordingly, non-parametric kernel density estimates will be used to examine how the overall shape dynamics of the distribution (spread and skewness features) have changed over time. Such changes will reveal straight evidence on whether dispersion has widened, narrowed or persisted. To evaluate the role of space, the paper compares the results from the "aspatial" distribution of regional income per worker against the "spatial" distribution obtained from the spatially conditioned scheme. In so doing, the paper not only reveals the role of spatial dependence but also underlines its effects in dispersion and convergence of regional income per worker in South Africa.

The remainder of this paper is structured as follows. Section 2 presents the economic convergence literature, while section 3 provides the empirical strategy. Section 4 provides a description of the data used in our empirical analysis. In section 5 the empirical findings are presented and interpreted. And, finally, some concluding remarks are made in Section 6.

2. Literature.

A number of economic theories have been put forward in the quest to uncover income disparities across, as well as within countries. Firstly, the neoclassical growth theory (Solow, 1956; Swan, 1956) is the starting point for most explanations for dispersions in growth rates and income levels across, as well as within countries. Due to decreasing returns to capital and similar technologies, the theory predicts decreasing disparities in income levels across, as well as within countries (Magrini, 2004; Mankiw, Romer, & Weil, 1992). Income disparities are seen only as short-run phenomena, deriving from exogenous population and technology shocks (Monastiriotis, 2006). Under these mechanisms high returns in capital allow poorer regions to grow faster than wealthy regions where returns to capital are low. Predictions of decrease in regional income disparities are also supported by classical trade theories that predict that openness to trade leads to factor and market price equalisation with complete specialisation of production.

In contrast, new growth (endogenous growth) and trade (new trade and new economic geography) {{39 Krugman, Paul R 1991; 43 Fujita, Masahisa 2001}} theories cast doubt on the neoclassical predictions of decreasing disparities and eventual convergence in regional incomes. Through increasing returns to scale due to technology differences across regions, the new growth theories points to persisting and even increasing income disparities across space. This outcome is also supported by new trade theory and new economic geography {{39 Krugman, Paul R 1991; 43 Fujita, Masahisa 2001}}. In these theories the process is driven by the interplay of increasing returns to scale and transport costs which can lead to a cumulative causation process for factor of production {{128 Monastiriotis, Vassilis 2007}}[4].

While it is quite clear in theory what disparities (increasing or decrease) and convergence entail, measuring these disparities empirically is not so straight forward as different concepts and strategies exists. For many years, regional income studies have mostly centred on beta-and sigma-convergence hypotheses, with the beta-convergence concept dominating the "economics convergence literature". Using regression models, beta-convergence exists when regions with low initial incomes experience faster growth rates than initially well-off regions. In the regression this is confirmed by a negative beta coefficient and is referred to as unconditional convergence, while if additional factors put as controls negative beta is interpreted as conditional convergence. On the other hand, sigma-convergence arise when the dispersion of income across a group of economies falls overtime. It can be seen as revealing the evolution of the distribution of income per capita across regions. For sigma-convergence several measures have been employed, among them unweighted standard deviation and the coefficient of variation of the log of income per capita (in our case income per worker).

[4] As predicted by {{39 Krugman, Paul R 1991}}, the cumulative causation process can lead to well-defined patterns of clustering and core-periphery income structures, defining multiple equilibria states as firms and workers agglomerate in regions with high market access.

These two concepts come from the neoclassical growth models and have grown to be known as the classical convergence approach[5]. While a large body of empirical literature using this approach at sub-national level now exists in developed and developing countries alike[6], empirical evidence regarding South Africa is very limited. This is not only limited to South Africa, but SSA as a whole and this could be attributed to lack of adequate data at finer geographical levels. In South Africa, empirical studies have mainly focused on income inequality at the national level[7].

This paper aim to build on existing studies in South Africa (Naudé and Krugell, 2003b; 2006; Krugell, 2005; Naudé and Krugell, 2004; Krugell et al., 2005 and Bosker and Krugell, 2008) and contribute to the literature which addresses regional income disparities at smaller geographical level.

Using panel data regression model to beta-convergence at the magisterial district level over the period 1998-2002, Naudé and Krugell (2003b; 2006) find no evidence of absolute convergence, but slow conditional convergence, with poorer regions very slowly catching up with the richer ones. Their result are also supported by Krugell (2005) who also used panel data regression model to beta-convergence at the magisterial district level over the period 1998-2002. Krugell (2005) also tests for sigma-convergence and find little evidence of convergence and his results are in line with results found by Naudé and Krugell (2004) in other study.

[5] And, Sala-i-Martin (1996) pointed out, these two hypotheses are related: Increase in incomes of poor regions enables them to catch-up with rich country, which mean evidence β-convergence. In other words, this implies that the dispersion in incomes between the rich and poor regions has decreased, which confirms evidence β-convergence. It's important to note that beta-convergence can exist without sigma-convergence, while for sigma-convergence, beta-convergence is a necessary condition.

[6] Building on the works of Baumol (1986) and Barro and Sala-i-Martin (1991; Barro, 1991) some of the studies that have applied the classical convergence approach include (Badinger et al., 2004; Paas and Schlitte, 2006) for sub-national for a group of countries, as well as (Levy, and Young, 2006; 2008, Ferreira, 2000; Nagaraj et al., 2000; Cikurel (2002) for sub-national regions for different countries. While most of the studies above confirmed evidence of conditional convergence, their results depend largely on the method used (cross-section, time series or panel data), as well as country under study.

[7] These include studies by (Bhorat & Kanbur, 2005; Hoogeveen & Özler, 2006; Leibbrandt, Levinsohn, & McCrary, 2005; Leibbrandt, Levinsohn, & McCrary, 2010; Leibbrandt, Wegner, & Finn, 2011; Leibbrandt, Finn, & Woolard, 2012; Van der Berg, Burger, Burger, Louw, & Yu, 2006; Van der Berg, Louw, & Yu, 2008; Van Der Berg, 2011; van der Berg, 2014; Wittenberg, 2014).

While we learn much from these studies, the classical convergence approach has been criticised for a number of failures {{92 Quah, Danny T 1996; 26 Quah, Danny T 1996; 38 Quah, Danny T 1997; 93 Fischer, Manfred M 2008}} and its appropriateness has generally been questioned[8]. One way to overcome the problems associated with the classical convergence approach is to apply the distribution dynamics approach pioneered by Quah (1993a, b). This approach focuses on the evolution dynamics of the entire cross-sectional distribution, addressing both the shape and internal mobility dynamics. It is an attractive empirical method as it reveals the distribution dynamics without imposing prior restrictive assumptions as done by classical convergence approach {{25 Magrini, Stefano 1999}}[9]. Accordingly, using the distribution dynamics approach (kernel density estimates), Krugell et al. (2005) examines the shape of the relative income distribution and how it evolved over the period 1996 to 2004. Their results confirm evidence of increasing regional income disparities.

While both the classical convergence approach and the distribution dynamics approach have offered valuable insights on regional income disparities, their major limitation is treating regions as 'isolated entities'. This implies that possible interactions that take place across regions are ignored and the observed geographic pattern of economic activity and resulting income levels are considered to be merely random. However, from a theoretical perspective, the importance of regional interdependence[10] in the determination of regional incomes is an issue that has been highlighted in the new economic geography literature (Krugman, 1991; Puga 1999)[11]. As such, integrating space into the distribution is of paramount importance, as it

---

[8] The classical convergence approach has been criticised for several econometric problems among them "Galton's fallacy of regressions toward the mean" {{12 Bosker, Maarten 2008}}. Furthermore, by focusing on average behaviour of a representative region it fails to capture some potentially interesting features of the distribution dynamics of the entire cross-sectional distribution (see Quah, 1993a, 1996a; Fischer and Stumpner, 2008; Magrini, 2009), such as spatial aspects/core-periphery (Krugman, 1991; Fujita et al., 1999), polarisation and club convergence (Quah, 1993a; 1996a) of the distribution. By concentrating on average behaviour rather entire distribution, it also fails to show the shape and mobility dynamics of regions across the distribution (Quah, 1996).

[9] To date a number of studies have applied the distribution dynamics approach using mainly non-parametric stochastic kernel density functions to examine regional income distribution dynamics (among them Bianchi, 1997; Magrini, 1999; Carrington, 2003; Johnson, 2000; Le Gallo 2004; Pittau and Zelli, 2006; Ezcurra et al., 2007; Bosker, 2009). While different conclusions are reached in these studies, generally these studies have mostly found little to no evidence for regional income convergence.

[10] Regions are usually linked and interact through different mechanisms such as factor mobility, interregional trade, technological and knowledge spillovers and institutions (Buettner, 1999; Maza and Villaverde, 2010).

[11]The NEG literature predicts that the interplay of increasing returns to scale and transport costs generate agglomeration and dispersion forces such as factor mobility (Krugman, 1991a; Puga, 1999), transport infrastructure (Krugman and Venables, 1995), housing stocks (Helpman, 1998; Hanson, 2004). These factors have a geographical element which can offer importance insights of the observed regional income distribution dynamics and may lead to a core-periphery income structure (Krugman, 1991).

relaxes the unrealistic assumption that economies are independent entities and highlights the importance of economic linkages and spillovers for the economic development of interconnected regions. Empirically, Anselin (1988), Anselin and Getis 1992, Getis et al. 2004, as well as Rey and Anselin (2006) have provided extensive evidence of the need to take the spatial aspects of georeferenced data into account. And, Rey and Janikas (2005) concluded that failure to account for possible spatial aspects in the data may invalidate results for both classical convergence and the distribution dynamics approaches.

The growing awareness of the importance of space has led many researchers to integrate spatial analysis into both the classical convergence and distribution dynamics approaches. While spatial econometrics (Anselin, 1988; 1998; 2001) have mainly been used within the classical convergence approach, the spatially conditioning approach (D. T. Quah, 1996b; D. T. Quah, 1997a; D. T. Quah, 1997b) and the spatial Markov chain approach (Rey, 2001) have mostly been applied in distribution dynamics approaches.[12] Empirical works incorporating geographical location in the analysis of regional income disparities are almost none-existing in South Africa. Naude et al (2010) and Bosker and Krugell (2008) are among the few exceptional studies to date. While Naude at el (2010) accounted for space within the classical convergence approach[13], the overall criticisms of the classical convergence approach continue to hold. Innovatively, within the distribution dynamics approach, Bosker and Krugell (2008) integrated space using both the spatially conditioning approach and the spatial Markov chain technique[14].

[12] These approaches not only incorporate space, but also offer a possible explanation for the observed regional income disaprities (Magrini, 2004). In explaining cross-sectional income distribution dynamics, Quah (1997a) emphasised the importance of trade patterns and geographical spillovers and in a different study (Quah, 1997b), he concluded that while both macro factors and geographical spillovers must be considered, geographical spillovers play a significant role. However, relatively few studies have accounted for spatial dependence in the distribution dynamics approach, yet spatial dependence can bias the standard deviation which is the central parameter in the kernel density estimators {{1 Rey, Sergio J 2005}}. And, {{1 Rey, Sergio J 2005}} concluded that "*While the use of kernel density estimators to analyse regional income distribution has attracted a great deal of attention, the properties of these estimators in the presence of spatial dependence are unknown".*

[13] Studies to also account for geographical location within the classical convergence approach in other countries includes (among them Fingleton 1999; Bosker 2007; Rey and Montouri 1999; Badinger et al. 2004; Baumont et al., 2002; Ertur et al. 2006; Le Gallo and Dall'Erba 2006) for developed nations, (among Ying, 2003; Turkey (Gezici and Hewings, 2002; Magalhaes et al., 2000; Ahmed, 2011; Resende, 2012) for developing countries. Another group of studies have instead focused on merely providing evidence of spatial dependence (see López-Bazo et al. 1999; Le Gallo and Ertur 2003).

[14] Generally, compared to classical convergence approach, studies incorporating space in the distribution dynamics approach are still relatively few among the (among them, Quah, 1996b, Le Gallo, 2004, Bosker 2006; Rey, 2001; Mossi et al., 2003; Liao and Wei, 2012). While, Quah, 1996b, Le Gallo (2004, Bosker 2006 applied the spatially conditioning scheme for regions within a group of E.U countries, Maurseth (2013) applied it across-countries for E.U. Rey (2001) applied the spatial Markov chain for U.S regions. Interestingly, for a developing country, Mossi et al., (2003), as well as Bosker and Krugell (2008) applied both the conditioned scheme and spatial Markov chains for the case of Brazil and South Africa respectively, while Liao and Wei (2012) employed

Their results confirmed the importance of geographical location in the distribution dynamics of regional incomes.

In summary, while we learn a lot from existing South African studies, these studies face a number a shortcomings: First, no study go beyond 2004 as existing studies covers the period 1990-2004. Second, with the exception of Naude et al (2010) and Bosker and Krugell (2008), these studies neglect the role of space and treat regions are independent entities. However, as pointed out in the literature this can lead to misleading conclusions if the data used has a geographical dimension as most of the datasets used. Third, GDP instead of personal income data is used in all studies[15]. This paper's methodological approach allows us to move beyond these limitations and adds to the growing number of studies in the literature, that have dealt with regional income disparities at sub-national level in other developing countries such as Brazil, China, India, amongst others.

Building on Bosker and Krugell (2008) who integrated space in examining regional income disparities, the main contributions of this paper are threefold. First, apart from extending the evidence found in Bosker and Krugell (2008) to 2011, this paper also provide an up-to-date analysis of regional income disparities in South Africa[16]. Second, the paper moves away from the usual applied datasets that are aggregated and interpolated and use the 2001 and 2011 census datasets. Apart from being collected at finer geographical levels required for this study, the census data also allows us to use income per worker which offers a more accurate picture of South Africa's regional income disparities compared to GDP per capita[17]. Third, inconsistent

the spatial Markov chains for China. Fischer and Stumpner (2014) also applied a different technique to deal with spatial dependence, the spatial filtering approach for a group for E.U countries. Lopez-Bazo et al (2004), Rey and Janikas (2005), as well as Migrini (2004) presents an overview of this literature. While few in numbers these studies have already provided evidence pointing to potentially important consequences for neglecting the spatial aspects underlying most regional datasets. Regardless of the region or country under studies, all these studies gave convincing evidence of the important role that space plays in the distribution dynamics of income across regions. An important conclusion that can be drawn from these studies is that incomes/growth is spatially contagious and normally leads to spatial clustering.

[15] Apart from being aggregated and interpolated from different sources, GDP per capita/GDP per worker is not equivalent to total individual/household income in a given region, it's simply an approximation. Moreover, GDP per capita is affected by a commuting bias arising from the effects of commuters who live in one region and work in another. Given the problems associated with GDP data, using income per worker offers a more accurate picture of South Africa's regional income disparities.

[16] The 2001-2011 period is not only interesting as it extend the study period beyond what has currently been covered in the literature, but it also coincide with the time South Africa implemented some of its SDIs and this can be any opportunity to evaluate the impact of such initiatives.

[17] While widely used in the literature GDP per capita/GDP per worker is not equivalent to total individual/household income in a given region, it's simply an approximation. Furthermore, GDP data is hardly collected at smaller geographical units in many countries including South Africa. And, it's available is through

and continuously changing geographical units have hampered analysis overtime, especially using the census data. To overcome this problem a novelty of this paper is the application of the areal-weighted interpolation technique to come up with compatible geographical units' overtime.

3. Empirical strategy

The objective of this paper is to examine the distribution dynamics of regional income per worker in South Africa, paying particular attention to the role of space. Our approach involves two steps. We explore how the entire cross-section distribution of income per worker evolves overtime. We focus on the shape dynamics of the distribution, concentrating on its spread and skewedness features, and how they have changed overtime. These two features will reveal how dispersed regional income per worker is in South Africa, how it has changed overtime (whether it increased, decreased or persisted), as well as regions driving the process (poor or rich).

To capture dispersion in regional income per worker we have the option to use the absolute regional income per worker ("unconditioned distribution"), or alternatively regional income per worker relative to national average income per worker ("nationally conditioned distribution"[18]). Although there are merits for using the unconditioned distribution, it is more informative to apply the nationally conditioned distribution when considering income changes over time. Nationally conditioned distribution allows us to control for business cycles and other random events that might affect regional income levels over time.

A natural approach to evaluate the shape dynamics is to estimate the entire cross-section income distribution using non-parametric kernel density estimates {{22 Silverman, Bernard W 1986(D. Quah, 1993a; D. Quah, 1993b)}}[19]. This approach is attractive as it avoids the strong restrictive assumptions which are central in parametric estimation. Thus, given a sample

---

aggregation and interpolation of data obtained from different data sources collected mostly at higher geographical levels. While important, such data is likely to generate GDP figures that are not very accurate given the assumptions and errors made in the aggregation and interpolation process.

[18] Regional income per worker is normalised as follows: $y_{it} = Y_{it}/Y_t$, where Y_{it} is real regional income per worker in region i (i=1,…, 354), in period t (t=1,…., 2), while Y_t is the national average income per worker. Essentially y_{it} enable us to compare each region's income relative to national average, and we conclude whether overtime more or less regions are below or above national income.

[19] In our empirical analysis we complement all the kernel density estimates with boxplot estimates. These plots clarities some distributional aspects which might not be very clear on the kernel densities.

realization, $y = \{y_1, y_2, \ldots, y_n\}$, from an unknown income distribution with density f, a nonparametric estimate of this density is given by:

$$f(y) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{y - Y_i}{h}\right) \quad (1)$$

Where n is the number of regions, i, $K(\bullet)$ is the kernel function, and h denotes the bandwidth parameter which determines the degree of smoothing. While the choice of the kernel function is considered to be a minor issue in the literature {{22 Silverman, Bernard W 1986}} the selection of the bandwidth is of great importance as the results from distribution dynamics strongly depend on the bandwidth parameters chosen. Small bandwidths yield small bias but large variance, while large bandwidths lead to large bias and small variance. A balance is needed between having too large or small a bandwidth. The consistently good performance over a wide range of kernel functions of the plug-in method based on (Silverman, 1986) rule-of-thumb, suggests it as the appropriate method to select the optimal bandwidth.

While informative, the non-parametric kernel estimation rests on the implicit assumption that regions are independent entities and each region provides unique information that can be used to generate the cross-sectional income distribution. As discussed in section 2, regions are not independent as they tend to interact. These interactions lead to the violation of the independence assumption central to the non-parametric kernel density approach. This may lead to misguided inferences and interpretations (Rey & Janikas, 2005). Accordingly, there is need to explicitly integrate (or remove) space into the cross-sectional income distribution. However, before doing so there is need to confirm that, indeed the distribution is linked to a spatial dimension. This can be achieved by testing for presence of spatial dependence in the distribution using exploratory spatial data analysis (ESDA) techniques.

A central issue in addressing spatial dependence is how to capture the spatial arrangement of the underlying cross-section data. In the literature these interactions are normally captured by a spatial weight matrix, W (in our case a 354x354 matrix) whose elements, w_{ij}, show the degree of spatial interactions between region i and region j, with closer pairs are given a higher weight (Anselin & Bao, 1997) [20]. While there are various ways to specify a spatial weight matrix in the literature, the simplest way is through a binary contiguity[21] relationship where regions interact if they share a common boundary ($w_{ij} = 1$, meaning $i\ and\ j$ are neighbours and 0 otherwise) (Moran, 1948). Although, the binary contiguity based matrices have widely been used, these matrices have the disadvantage of limiting spatial interactions to those regions that are physically adjacent (share a border only) to each other. However, through such factors as interregional trade, factor mobility, knowledge and technology spillovers spatial interactions go beyond adjacent regions. An alternative way to capture spatial interactions between regions is to use distance-based spatial weight matrices. Apart from incorporating spatial interactions beyond contiguous regions, distance-based weight matrices also have a solid theoretical foundation rooted in the gravity trade models (see(Getis, 2009) Given a distance-based weight matrix intensity of the interactions between regions decrease with increasing distance between them. This relationship is capture by an inverse function of the distance between regions.

While the inverse distance spatial weight matrix assumes that all regions interaction, as distance increase these interactions become negligible. As such a critical distance cut-off point above which spatial interactions are zero is usually implemented in the literature (Dall'Erba, 2005). This gives a binary inverse distance weight matrix[22] given as follows:

$$W = \begin{cases} w_{ij} = 0\ if\ i = j \\ w_{ij} = 1/d_{ij}{}^{\alpha}\ \ if\ d_{ij} \leq D_i \\ w_{ij} = 0\ if\ d_{ij} > D_i \end{cases} \tag{2}$$

where, d_{ij} is the great circle distance[23] between region $i\ and\ j$. α is a distance decaying parameter taking positive values (we use α=1) which ensures that spatial interactions fall with

[20] This idea of greater weight to closer regions is summed up by Tobler's First Law of Geography which states that "Everything is related to everything else, but closer things are more related" (Tobler, 1970, p.234) (Fischer & Wang, 2011). The weight matrix is normally standardised such that rows add up to one, meaning W ranges between 0 and 1

[21] The binary contiguity matrix takes many forms among them rook, double rook, bishop and queen weight matrix (see (LeSage, 1999) for more details). Rook contiguity refers to regions that share a common side with the region of interest, bishop refers to regions that share a vertex with the region of interest, while queen means regions share a common border on either side.

[22] Apart from the binary distance matrix, the inverse-distance weight matrix can also be manipulated into squared inverse distance, binary squared inverse distance and k-nearest neighbours.

[23] Distance between geospatial units is commonly computed from the latitudes and longitudes of the units under analysis and is given by $d_{ij} = \sqrt{\sum_{t=1}^{q}(x_t[i] - x_t[j])^2}$ which gives the straight-line or great-circle distance (Drukker, Peng, Prucha, & Raciborski, 2013) . However {{117 Drukker, David M 2013}} noted that when the regions under analysis are located on the surface of the earth and coordinate variables represent the geographical

increasing distance between regions. D_i is the critical distance cut-off above which spatial relations are assumed negligible. To ensure that relative, rather than absolute distance is considered the spatial weight matrix is transformed to raw-standardised ($w_{ij}^* = w_{ij} / \sum_{j=1}^{n} w_{ij}$) to ensure that each row sums up to one. Such transformation ensure easy use, computation and interpretation of spatial autocorrelation results. Given the sensitivity of spatial dependence results to the spatial weight matrix used, we carry out robustness checks taking different distance cut-off points, different decay functions, as well as different weight matrix specifications, mainly the binary continuity matrix.

Given the spatial weight matrix, the overall presence of spatial dependence in the data can be tested using the well-known standardised Moran's I statistic (Moran, 1948) , which is defined formally as a weighted correlation coefficient (Anselin & Bao, 1997) given by:

$$I = \frac{n \sum_i \sum_j w_{ij}^* (y_i - \overline{y})(y_j - \overline{y})}{\sum_i \sum_j w_{ij}^* \sum_i (y_i - \overline{y})^2} \qquad (3)$$

where n is the total number of geographical units[24] indexed by i and j; y_i and y_j is the income per worker for region i and j, while $\overline{y}$ is the national average income per worker. Moran's I statistic ranges from +1 to -1, with +1 indicating perfect positive spatial association, which implies clustering of regions with similar income levels, either high or low (also referred to as hot or cold spots). Positive spatial association implies that regional income disparities have a geographical dimension and income of a given region is influenced by incomes of its neighbouring regions which leads to formation of spatial clusters regimes[25]. On the other hand, a Moran's I statistic of -1 reveals perfect negative spatial association (spatial outliers), which shows that dissimilar parameter values cluster together (either high - low values or low - high values). However, lack of evidence of spatial clustering suggests a random distribution of values and this is revealed by a Moran's I statistic of 0.

coordinates of the spatial units it's best to use great-circle distance as it captures the shape of the surface areas (as such, takes into account effects of mountains). As such in this paper we use the great-circle distance.
[24] Just like the spatial weight matrix, the unit of analysis is also critical for the revealed level of spatial association and an optimal geographical unit of analysis is needed and our choice will be explained in the data section.
[25] However, from a technical point of view evidence of positive spatial clustering might be highlighting a measurement error problem (Rey & Montouri, 1999) or omitted variable problem (Abreu, De Groot, & Florax, 2004).

Evidence of spatial dependence implys violation of the independence assumption central to the non-parametric kernel density estimations. This may lead to misleading conclusions, as such to account for spatial dependence, we opt to integrate space into the cross-sectional distribution using the spatially conditioning scheme proposed by (D. T. Quah, 1996b)[26]. In his approach Quah suggests integrating space into the cross-sectional distribution by normalising regional income by a weighted sum of neighbouring regions' income to get spatially conditioned relative income distribution (neighbour relative income distribution). As with the Moran's I statistic a central issue in constructing the regionally conditioned distribution is who to integrate space into the distribution. As revealed earlier there are several ways to do this and in this paper we continue to use the binary distance spatial weight matrix in equation 2 to have:

$$y_i^* = \sum_{i \neq j} w_{ij}^* \, y_j \qquad (4)$$

Where w_{ij}^* is an element of the spatial weight matrix which is defined as before. y_j is the regional income per worker for region j, while y_i^* is the newly derived spatially conditioned variable. This variable not only allows us to understand the dynamics the entire cross-sectional distribution, but also to quantify geographical spillovers and spatial interaction across regions.

To determine whether space matters, (D. T. Quah, 1996b) suggests using the nationally conditioned distribution as the benchmark distribution and compare it with the spatially conditioned one. Given the conditioning scheme, both distributions have three possible outcomes: a region's income can either be smaller, equal to or greater than (<1, 1 and >1) the overall national income or weighted sum of neighbouring regions income. Implying that 1 is the desirable outcome which suggests income are uniformly distributed across space or regions have similar income levels with their neighbouring regions[27]. Comparing the two distribution at a given time, local spatial factors are said to matter if the spatially conditioned distribution departs from the identified map of the nationally conditioned distributed (Magrini, 2004; D. T.

---

[26] To date the approach has now been used widely in the literature and some of the studies to apply it includes (Bosker, 2007; Bosker & Krugell, 2008; Bosker, 2009; Le Gallo, 2004; Mossi, Aroca, Fernández, & Azzoni, 2003). Another approaches are the spatially conditioned Markov distributions proposed by (Rey, 2001) and spatial filtering approach by Getis (1990; 1995) (Fischer & Stumpner, 2008; Maza & Villaverde, 2009). The motive to control for space using regionally conditioned distribution by (D. T. Quah, 1996b) was driven by is simplicity to integrate space yet giving reliable results.

[27] While a value of 0.5 indicates a region's income half the national average income (its neighbour's income) and 2 indicates a region's income twice the national average (its neighbour's income), and so on. Overall we can evaluate whether relatively poor or rich regions remains relatively poor or rich overtime compared to national or neighbouring regions.

Quah, 1996b)[28]. (D. T. Quah, 1996b) described the spatially conditioned distribution as that parts of regional income unexplained by physical-location factors. He argued that, space does not matter in the event that physical location factors account for all the observed dispersion in regional income. In that case the spatially and nationally conditioned distributions will be similar.

4. Data and descriptive statistics

The analysis of regional income disparity depends essentially on the availability of consistent and reliable data at finer geographical levels. Such data is a challenge to get in most developing countries, especially in SSA. However, South Africa is one of the few exceptional cases where a number of spatially disaggregated data sources are available. It's interesting to note that Quantec and Global Insight Southern Africa through their regional integrated databases known as regional economic explorer (REX), as well as Statistics South Africa (Stats SA) through the national census datasets collects data at finer geographical levels desirable for our analysis. While, Quantec and Global Insight Southern Africa databases have the advantage of having a series of data covering many years (providing a panel), the major drawback of these databases is that they are based on projections from different data sources (W. F. Krugell, 2011)[29]. Implying that the databases are not based on actual survey data and there is room for error in the process of aggregating the data to the preferred geographical unit. On the other hand, the national census datasets collected by Stats SA are based on actual survey data collected at finer geographical levels.

This paper utilises the full national census data for 2001 and 2011. Its major strength is its large sample size and availability of a wide range of information collected at different geographical levels. The geographical levels range from highly aggregated to highly disaggregated units (provincial, district, municipality, magisterial districts, main place and even ward level)[30].

[28] While, (D. T. Quah, 1996b) additional carried out the comparison using transition probability matrix of the two distributions, in this study we simply compare the shape dynamics of the two distributions. We concentrate on the location, spread and skewedness features of the distribution.

[29] The projections are based on data from national census, household and labour survey surveys, government departments, as well as development agencies and in the process of reconciling to have one unique dataset many assumptions are made which makes the data far from perfect. A discussion and comparison of these databases with the census data is contained in (De Klerk, 2012).

[30] Accompanying these census datasets are different metadata files (shape files which show the demarcation of South Africa different geographical units) that enable us to explore the spatial aspects of the data. The metadata

However, just like any other dataset, the census datasets are far from ideal, especially in two major ways[31]. First, the major challenge for working with geographical referenced data is the need to have a compatible geographical unit of analysis overtime (Martin et al, 2002). However, due to on-going demarcations of South Africa's administrative boundaries, its geographical units have been changing overtime. This entails that comparison overtime is impossible. As such, to account for these changes and come up with a compatible geographical unit overtime, ArcGIS overlay tool was used to overlay 2011 sub-place boundaries to 2001 magisterial district boundaries using the areal-weighted interpolation technique. From the overlay results, a simple areal-weighted ratio was created to aggregate 2011 data at sub-place level to 2001 magisterial district boundaries (354 units). This makes magisterial districts our unit of analysis.

The use of magisterial districts as our unit of analysis is motivated by the factor that, they define the location of cities and towns in South Africa. It's our belief that it's in these cities and towns where most economic activities take place, hence define the local labour market well[32]. Through statistical analysis, (Von Fintel, 2014) found that magisterial districts define South Africa's local labour market well. Furthermore, on a technique level, a balance is needed between having too many or too few units as either extremes can lead to serious underestimations (Von Fintel, 2014). For example, apart from the weight matrix used, the spatial autocorrelation results depend heavily on the spatial scale used. With 354 units, magisterial districts provide a reasonable and optimal spatial unit of analysis and its appropriateness has seen it being used in most sub-national studies in South Africa.

Second, while a continuous income measure is needed Stats SA collects income data in brackets as a way to increase responses rate. As such, to obtain a continuous income measure the midpoint approach was used[33]. Since the top end bracket was open-end, to get the midpoint

---

enables calculation of distance between regions, which is critical for incorporating space in the analysis of regional income disparities.

31 A detailed explanation of the census data problems and solutions is provided in Appendix 1. The appendix also provides the definition and summary statistics of all the variables to be used in this chapter.

32 While 2011 has many geographical units we could have used, we used sub-places to ensure accuracy of our results by aggregating from a smaller unit to a higher one.

33 For robustness checks we used the 10% census sample data to derive a continuous income measure using the multiple imputation technique. This was done to see whether the statistics from the midpoint and multiple imputations approach were significantly different. Our results showed that the results from the two approaches

we applied StatsSA decision rule which assigns a mid-point which is two times the lower bound of the top bracket. Apart from being bracketed, the income data is also characterised by high values of missing and zero income (this could be attributed partly to children included in the sample). To minimise this problem we limit our sample to employed individuals only (age 15-64 years) who normally have a positive and better reported than income[34]. While this looks arbitrary, most economic theories (growth and trade models) are based on production functions, whose implications relate more closely to income per worker than income per capita (Durlauf, Johnson, & Temple, 2005). To allow comparison over time we adjust regional income for cost of living. We deflated 2001 nominal income was deflated with the consumer price indices data supplied by Stats SA[35]. It's important to note that while we have all these issues with regard to income data, its availability at lower geographical unit is a big advantage as it provides a better proxy for regional income than GDP which is used in most regional studies not only in South Africa, but in many other countries.

From the constructed database, the key variable for our analysis is regional income per worker. Regional income per worker data is normalised by South Africa's average income per worker. This allows us to abstract from the growth and other random shocks of the South African economy during the period under study. Moreover, the normalisation also accounts for common changes in inflation (Juessen, 2009). The complete sample consists of 354 regions comprising of extremely rich and poor magisterial districts. However, in the event of extreme outliers in the sample we do not drop them from our analysis because these regions reflects the true structure of South Africa's economy. From an economic point of view it is not appealing to simply delete these observations (D. T. Quah, 1997a).

are consistence and led to the same conclusions. For the multiple imputation technique we draw from Rubin's (1987) works on coarse data and follow Royston (2007) and Daniels (2012) works to derive a continuous individual income measure. A univariate multiple imputation algorithm for bounded course data which uses each individual characteristics in our data that are known to best predict individual income or wage to derive a continuous income measure is used. The set of exogenous factors, used to derive the continuous income measure are age, age2 (capturing work experience), marital status, education level, employment status, area of residence, population group, gender and province. These factors are drawn from existing literature (see Daniels, 2012) and are in line with the usual factors for the extended mincerian earnings equation and the continuous income measure is derived by replacing missing and bracketed income by random draws from the posterior predictive distribution.

[34] Focusing on income for employed individuals only is also important as it's the variable most theories that try to give possible explanation for income differences use.

[35] We converted 2001 nominal income to real income using December 2012 as the base year. With a CPI of 52.6 for 2001 and 94.2 for 2011, real income for 2001 is obtain by: Real_2001= Nominal_2001*(94.2/52.6).

Table 1 displays summary statistics for regional income per worker by census year and geographical scale (national, top and bottom 20 regions, homeland and non-homeland regions, as well as richest and poorest regions). Overall, monthly income per worker increased from 7475 rands in 2001 to 8396 in 2011 representing a 12 percent increase. Across different geographical groupings, the poorest region experienced the highest increase (73 %), while the income for the richest region only increased marginal (3%). It's clear that the aggregate national income, masks significant differences across different geographical scales. While the 20 richest regions had average income 1.59 and 1.56 times that of national income in 2001 and 2011 respectively, bottom 20 regions had average income 0.2 and 0.3 times that of national income[36]. Furthermore, the top 20 percent regions accounted for 35 and 33 percent of total income in 2001 and 2011 respectively, while the lowest 20 percent regions contributed 11.5 and 12.6 percent to national income over the same period.

Table 1: Summary statistics for monthly income per worker across regions.

		Regional income		**Income Δ**	**Relative income**	
Region	Number of regions	2001	2011	% Δ	2001	2011
National	354	7475	8396	12	1	1
Top 20	20	11878	13104	10	1.59	1.56
Bottom 20	20	2265	3155	39	0.30	0.38
Richest	1	20059	20708	3	2.68	2.47
Poorest	1	1460	2529	73	0.20	0.30

Notes: Regional income is monthly income per worker in rands. National – South Africa's average monthly income per worker, Top and Bottom 20 - richest and poorest 20 regions, Richest and poorest - richest and poorest region.

The statistics in Table 1 tend to increase with geographical scale. While, the richest region (Randburg both years) had a monthly income per worker 2.7 and 2.5 times that national average income in 2001 and 2011, the poorest regions, Clocolan in 2001 and Joubertina 2011 had income per worker 0.2 and 0.3 times that of national income. These statistics seem to suggest wide dispersion in regional income per worker in South Africa, which tend to decrease over the study period.

[36] The list of top and bottom 20 regions in shown in Table 1A in appendix 1A.

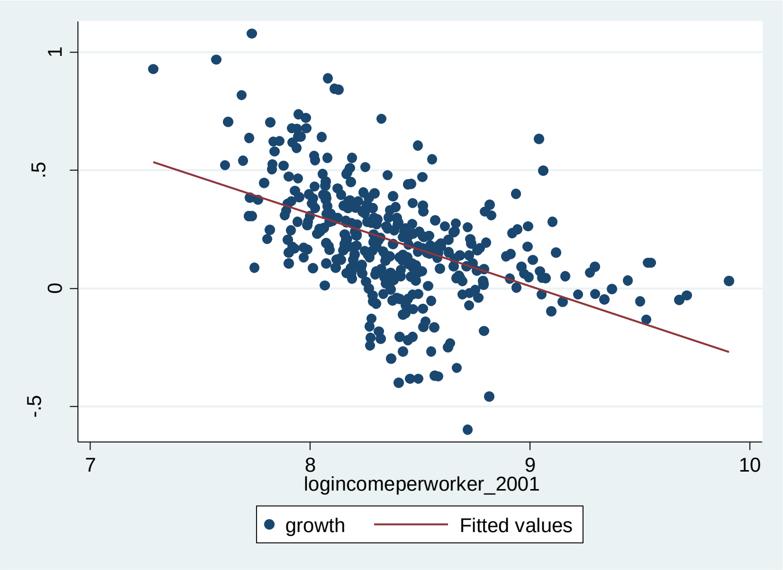
5. Empirical results

From the preliminary evidence from these descriptive statistics above, it is clear that levels and changes in regional income per worker widely dispersed across regions in South Africa. Here, using non-parametric kernel density estimates we present and discuss evidence of the evolution of the entire cross-sectional distribution of regional income per worker distribution among South Africa's 354 magisterial districts. To reveal levels of dispersion in regional income per worker, we focus of the shape dynamics, particularly the spread and skewedness features of the distribution and how they have changed overtime. In the first step we estimate univariate kernel density functions of unconditioned and nationally conditioned regional income per worker for 2001 and 2011. To check whether these distributions are not affected by space dependence for its presence using the Moran's I statistic. Evidence of spatial dependence leads us to estimate a spatially conditioned relative income distribution which allows us to explicitly integrate space into the distribution.

5.1 The distribution dynamics of regional income per worker

While the beta and sigma-convergence hypotheses have been heavily criticised in the literature, to provide a background to our distribution dynamics analysis, we provide evidence of unconditional beta and sigma-convergence first. To show evidence of unconditional beta-convergence, Figure 1, presents a scatterplot of the growth in income per worker (2001 to 2011) against initial income per worker (logged 2001). The negative slope in the scatterplot reveals strong evidence of a process of unconditional beta-convergence across magisterial districts in South Africa. This implies that on average over the 10 year period poor regions grew faster than relatively richer regions, which provide evidence of unconditional beta-convergence. This is evident from the plot as convergence is an outcome of relatively strong growth in low income regions of Free State and KwaZulu-Natal, as well as relatively weak growth city region in Gauteng.

Figure 1: Unconditional convergence in regional income per worker, 2011-2011[37]



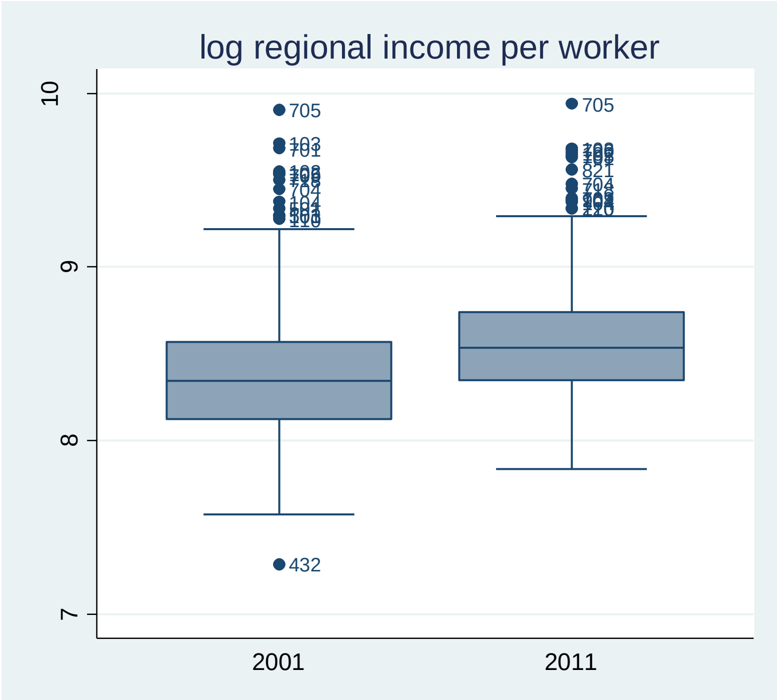
Turning to sigma-convergence, which implies fall in dispersion of regional income per worker overtime and this is confirmed by a fall in regional income per worker relative to mean regional income per worker. Indeed, a look at the standard deviation of log of real income per worker for all regions reveals evidence of sigma-convergence. It declined from 0.394 in 2001 to 0.346 in 2011, giving a 12.2 percentage fall in dispersion of income per worker over the 10 year period. These results are in line with findings by Naudé & Krugell (2003), who found a 13 percent decline in the dispersion in GDP per worker over a 10 year period from 1990 -2004. We use Tukey boxplots[38] for log regional income per worker, Figure 2, to visualise this. This allows us to see whether increase in relative incomes of poor or decrease in income of relative richer regions is driving sigma-convergence[39].

[37] The significance of the negative relationship is confirmed in Table 2A in Appendix 1A showing the simple OLS regression results for growth rate and initial income per worker in 2001.

[38] The boxplot consists of three main elements: the box, the whiskers and the extreme values. The box captures the 25th and 75th percentiles of the data, hence gives the interquartile range (IQ) and the line within the box represents the median income value. The whiskers captures the lower and upper values in the data, while values outside the lower and upper values are considered to be outliers or extreme values. Overall, the boxplot gives an indication of the variability of the data. The size of the box, position of the median value and length of the whiskers, as well as existence of extreme values tell us whether the distribution is symmetric or skewed, either to the right or left. Above all highlights how dispersed incomes are and how the dispersion has changed overtime.

[39] A look at the max/min ratio also provide further evidence of sigma-convergence. While the richest region in 2001 had income per worker 13.7 times more than that of the poorest region (Randburg/Clocolan), this ratio decreased to 8.2 in 2011 (Randburg/Joubertina),

Figure 2: Boxplots of dispersion in regional income per worker, 2011-2011



Notes: Boxplots for South Africa's 354 magisterial districts.

From the plot clear evidence of sigma-convergence is confirmed by a decrease in the size of the box and interquartile range. This decrease is driven by increase in incomes of relatively poor regions, implying that poor regions are growing faster than relatively richer regions an outcome which also supports beta-convergence. This seem to confirm arguments in the literature that beta-convergence (growth of poor regions) tends to generate sigma-convergence (reduction in dispersion) (Naudé, Krugell, & Matthee, 2010; Sala-i-Martin, 1996)[40].

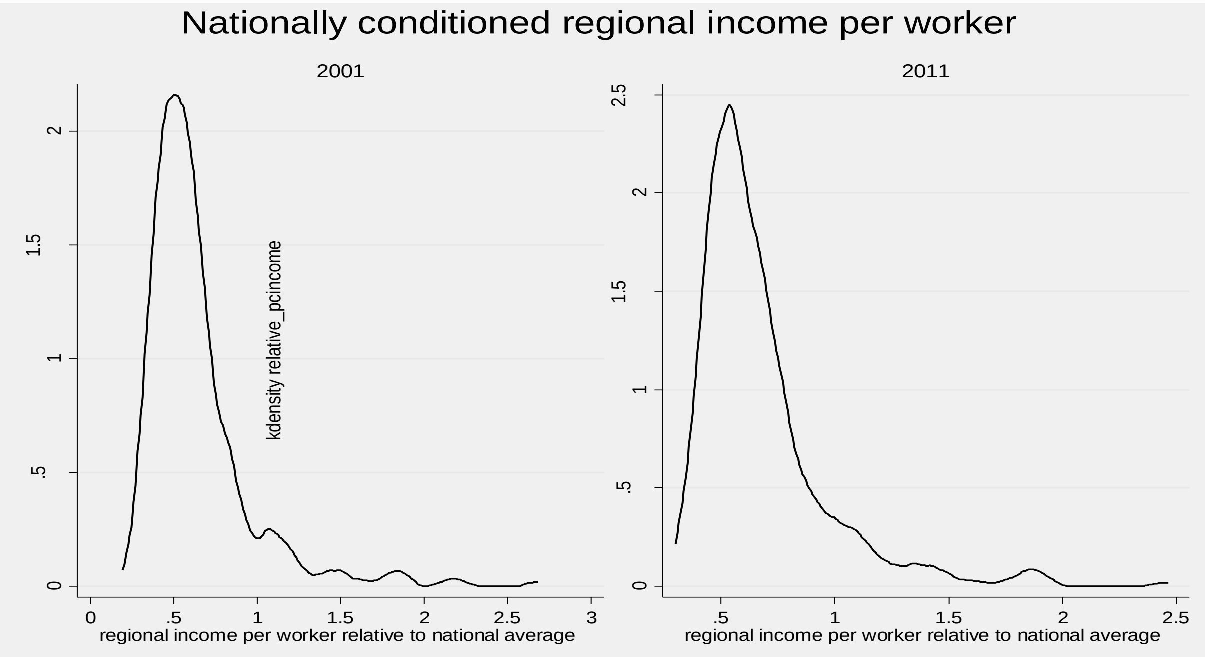
While giving important insights about the convergence process of regional income per worker across regions, beta and sigma-convergence hypotheses hide important features of the evolution dynamics of the entire cross-sectional distribution. To explore the distributional dynamics of the distribution we turn to the distribution dynamics approach. We concentrate on the shape dynamics of the distribution, examining its spread and skewedness. The distribution dynamics approach allows us to examine the cross-section, as well as the time dynamics of the distribution of regional income per worker. Starting with the cross-section feature, Figure 3

[40] As (Sala-i-Martin, 1996) pointed out, beta and sigma-convergence hypotheses are related. And, if there is β-convergence, increase in incomes of poor regions enables them to catch-up with rich country and this implies a decrease of the dispersion between the two countries. While beta-convergence can exist without sigma-convergence, for sigma-convergence, beta-convergence is a necessary condition.

presents the nationally conditioned relative distribution for South Africa's 354 regions for 2001 and 2011. The densities reveal wide dispersion in income per worker across regions in both years. While the richest region had income per worker 2.68 of national average in 2001, for the poorest it was only 0.2. On the other hand for 2011 for the richest it fall slightly to 2.47 times, while for poorest increased to 0.3.

Perhaps the most pronounced feature in these densities is the high degree of positive skewedness in both years. The positive skewedness suggests that the bulk of the regions were gravitated towards the bottom part of the distribution. This highlights that most regions in South Africa had incomes below national average (less than 1 and almost half the national average) and this is also confirmed by other studies in South Africa using GDP per capita (Bosker & Krugell, 2008). Only a few regions had incomes greater than national average in both years. While about 10 percent of the regions had incomes (35 out of 354) greater than national average in 2001, this figure increased slightly to 11 percent in 2011 (39 districts).

Figure 2: Cross-sectional distribution of regional income per worker.



Notes: Gaussian kernel densities with bandwidth chosen using (Silverman, 1986) rule-of-thumb. By nationally conditioned we mean regional income per worker relative to national average income per worker.

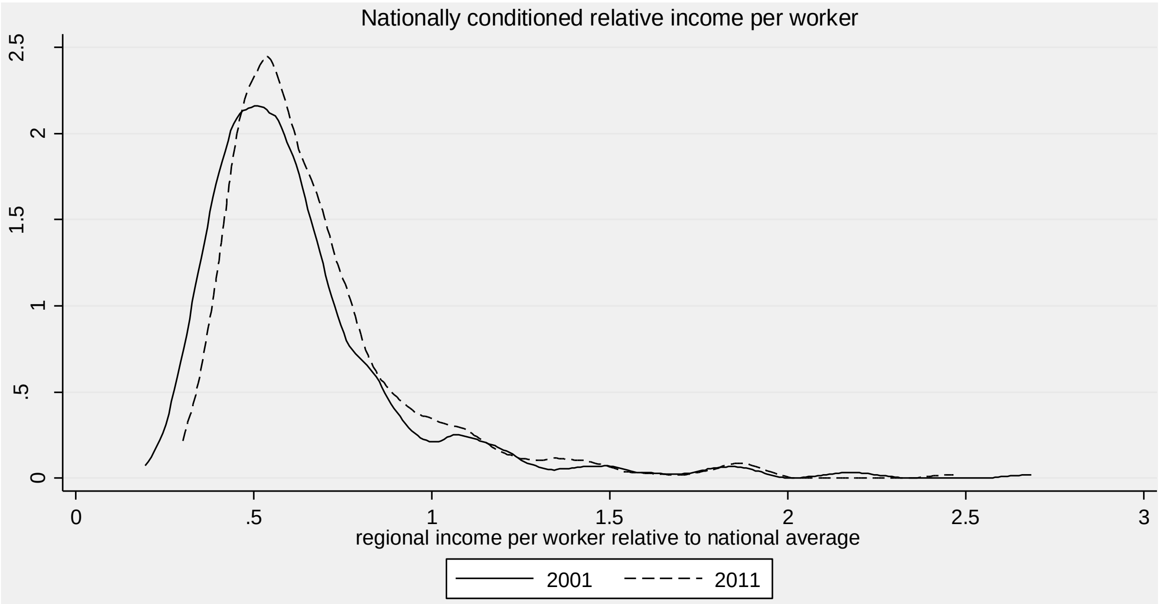
It's interesting to note that while both densities had a huge mass around 0.5 of national average, some small dumps more pronounced in 2001 can be seen at the upper end of the distributions. This might be indicating the presence of two groups in the population of regions, which can be defined as all other regions and the extreme outlier regions (mainly wealthy regions of Gauteng and Western Cape). However, there is a tendency of the small dumps to disappear overtime and this might be the first sign for convergence as the distributions move progressively towards unimodality.

To get the distribution dynamics of regional income per worker, Figure 3 presents densities for 2001 and 2011 nationally conditioned distributions for income per worker[41]. Contrary to the unconditioned densities, the nationally conditioned densities seem to suggest a decline in dispersion in income per worker across regions. This is highlighted by an inward shift in both the right and the left tails of the 2011 density, which confirms reduction in variability. This is further reflected in the increase in the peak for 2011 density, which implies higher concentration of regions. This phenomena is driven by two distinct sources: increase in relative incomes of poorer regions, as well as slight fall in relative incomes of richer regions. Implying that poorer regions experienced faster growth than richer regions over the study period. Overaly, by revealing faster growth of poor regions and decrease in variability, these densities confirm our earlier results for beta and sigma-convergence among South Africa's 354 regions.

Furthermore, it's interesting to note that over the 10 year period the overall shape of the distributions seem to have changed. The most visible change is the reduction in visibility of the small second mode (at 1) in 2011 observed in 2001. Moreso, the slight changes in the share of total income accounted for by the top and bottom 20% regions (fall from 35% to 33% for top 20 regions and increased from 11.5% to 13% for poorest 20 regions) also confirm these slight dynamics.

[41] In Appendix 1B, Figure 1B and 2B we also present cross-sectional and distribution dynamics density functions for unconditionally regional income per worker. These densities have the same distribution shape and features as those derived from the nationally conditioned densities and seem to suggest an increase in dispersion of regional income per worker overtime. However, these densities can be very misleading as the results are only valid when the years under study experienced the same economic conditions, which is highly unlikely. Overtime economics experience business cycles, inflation and other economic shocks, as such these densities are not very informative for distribution dynamics. This motivates us to use conditioned distributions as they remove these effects.

Figure 3: Distribution dynamics of regional income per worker.



Note: By nationally conditioned we mean regional income per worker relative to national average income per worker and on the x-axis 1 present national income which is the desirable outcome.

While the analysis this far seem to confirm evidence of convergence in regional income per worker, the methods adopted essentially ignore the spatial aspects that characterise our dataset. This implies that regions are treat as independent entities and the observed income per worker is assumed to be merely random. However, as discussed in section 2 there is a broad agreement in the literature that the process of income dynamics and convergence is characterised by a spatial dimension. And, interactions or spillovers across regions are likely to lead to spatial dependence across space, violating the independence assumption (Abreu et al., 2004; Anselin & Bera, 1998). Existence of spatial dependence imply that the income of a given region is influenced by the income levels of its neighbouring regions. This relationship can have important implications in the distributional dynamics of regional income per worker as it might alter the shape dynamics (spread and skewedness) of the entire cross section distribution. As such, without a proper assessment and account of space, confidence over our earlier results on reduction in dispersion in income per worker diminishes.[42]

[42] However, Frenken and Hoekman (2006), pointed out that usual convergence conclusions continue to hold when space is integrated into the analysis. Also, (Le Gallo & Dall'Erba, 2008) state that, while undertaking a sigma-convergence, the conventional approach and the spatial one do not produce contradictory conclusions.

At this point the question that arise is: is the cross-sectional distribution of income per worker across regions somehow linked to a geographical dimension? In other words, is there spatial dependence? Preliminary evidence revealed by the maps in Figure 1B in Append 2B seem to confirm presence of spatial dependence (spatial clustering) in the distribution.[43] We test this claim formally using the Moran's I statistic. And, the results presented in Table 2, do confirm evidence of positive spatial dependence for nationally conditioned relative income and its percentage change[44]. However, overtime the intensity of spatial dependence decreased from 0.312 in 2001 to 0.255 in 2011. This might be suggesting that overtime geographical location is playing a lesser role in influencing the distribution of income across regions. As suggested by the NEG theory, this might be highlighting the effects of congestion costs, arising from agglomeration of firms and workers.

Table 2: Global spatial autocorrelation results for regional income.

	Regional income		% Δ in rincome
Variables	2001	2011	2011-2001
I	0.312	0.255	0.205
E(I)	-0.003	-0.003	-0.003
sd(I)	0.015	0.015	0.016
Z	20.517	16.708	13.342
p-value*	0.000	0.000	0.000

Notes: The global spatial autocorrelation results were derived using an inverse distance-weighted spatial matrix with 205 kms as the cut-off point. For robustness, different cut-off points, as well as the contiguity and nearest neighbour spatial weight matrices were also used and the conclusion of positive and significant spatial association continue to hold.

Evidence of spatial dependence is a clear conformation that overally income per worker is positively spatially clustered across space in South Africa. Implying that regions with higher (lower) than national average income tend to cluster in space and this finding is in line with

[43] From the maps, one can immediately see striking differences in income across regions, which tend to cluster in space (regions of high (or low) income cluster together) in both 2001 and 2011. While, the wealthiest regions tend to cluster around Gauteng and Western Cape, the poorest regions were concentrated in inland regions of the country around Free States in 2001, as well in rural parts of KwaZulu-Natal and Eastern Cape in 2011. These clusters seem to suggest a rural-urban division in line with the core-periphery income structure predicted by the NEG (Krugman, 1991) .

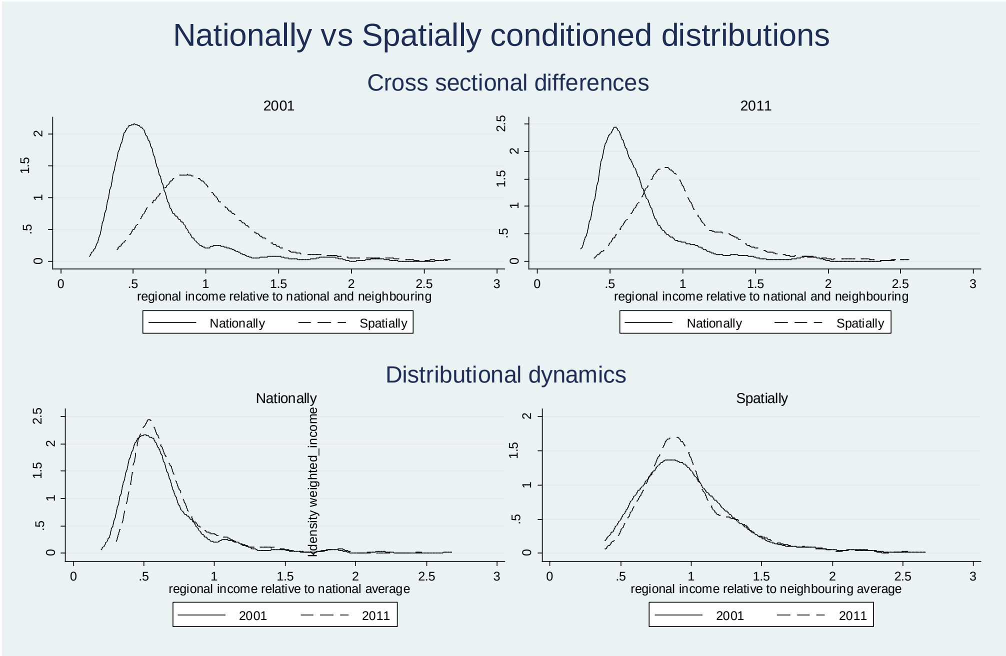
[44] To reach a conclusion the significance of the test compare the calculated Moran's I value with the Moran's I expected valued. For example for 2001 we compare I=0.312 and E(I)=-0.003 and given that the I>E(I) and the p-value is statistical significant this confirms evident of positive spatial dependence.

most findings in the literature (Celebioglu & Dall'erba, 2010; Le Gallo, 2004; Rey & Montouri, 1999) and more importantly confirm earlier findings South Africa by (Bosker & Krugell, 2008)

Evidence of positive spatial dependence imply that South Africa's regions are not independent entities. Implying that the independence assumption central in the "aspatial" non-parametric kernel density estimations does not hold. This may lead to misguided inferences and interpretations (Rey & Janikas, 2005) . Implying that our earlier results from the nationally conditioned distribution might be misleading. Accordingly, we re-estimate the non-parametric kernel densities using the spatially conditioned relative income distribution (or a neighbour relative income distribution). The distribution explicitly integrate space (spatial dependence) into the cross-sectional income distribution. As such we expect the spatially conditioned variable to be randomly distribution in space, implies that the independence assumption is not violated. Indeed, Figure 4B in Appendix 2B which maps the spatial distribution of the variable seem to suggest that it's randomly distribution. This is confirmed with the Moran's I test which shows the variable is statistical insignificant.

Given the spatially conditioned variable, to evaluate the role of space (spatial interactions) among spatially contiguous regions, we compare the kernel density estimates from the spatially conditioned distribution with those from the nationally conditioned distributions (use it as benchmark for comparison). If local spatial factors matter in the cross-sectional distribution of regional income per worker across regions, then the spatially conditioned density will depart from the identified nationally conditioned density. Indeed, the first panel of Figure 4 conveys precisely this message.

Figure 4: Kernel density estimates for nationally and spatially conditioned schemes.



By plotting the nationally and spatially conditioned densities for each year, we see marked differences between the two functions. While both distributions are positively skewed it's evident that the spatially conditioned densities are less skewed, much more symmetric and centred around 1, with a larger spread. By centring around 1 (compared to 0.5 for national) the spatially conditioned densities confirms that once space is integrated into the analysis, neighbouring regions tend to have very similar levels of income per worker. In other words, while regions have different levels of income per worker relative to national average, on average they had very similar levels with their neighbouring regions. This implies that a substantial part of the variation in income per worker is attributed to local spatial factors, rather than national (Magrini, 2004) . This means the economic fortunes of a given regions depend largely on the economic performance of its neighbouring regions, a factor which highlights the importance of geographical location for economic outcomes. However, there are some regions with much high incomes than their neighbouring regions (between 1 and 2.5 times). This might be suggesting that South Africa's high income regions are scatter around representing regional growth poles. This is largely expected given South Africa's history which led to the development of a few regions at the expense of other regions. For example, the development of South Africa 6 largest regions that emerged out of its history (Johannesburg, the East Rand, Durban, Cape Town, Pretoria and Port Elizabeth).

Evidence of a larger spread for the spatially conditioned distribution seem to suggest that spatial dependence leads to a less homogenous distribution, with geographical location favouring economic prosperity of some regions and at the same time hindering for others. This further confirms the importance of space in explaining disparities in regional income per worker in South Africa[45], a result which affirms earlier findings for South Africa by (Bosker & Krugell, 2008). Evidence of the importance of geographical location, supports the theoretical predictions of the new economic geography theory (Fujita et al., 2001; Krugman, 1991).

Panel 2 of Figure 4 reveals information on the distribution dynamics of regional income per worker. As with the nationally conditioned densities a fall in dispersion in regional income per worker is also evident from the spatially conditioned densities. However, a test of the significance of these distribution reveals that while the changes in the nationally conditioned distribution are significant, those for the spatially conditioned distribution were not (tests of the test are in Table 3A in Appendix 1A[46]). This implies that South Africa's regions have essentially maintained their initial incomes per worker levels relative to neighbouring regions weighted average ones. This indicates persistence in dispersion of levels of income per worker across regions. As such, once spatial dependence is integrated in the analysis the picture that emerges appears to lend little support to the classical convergence approach and "aspatial" density results. This indeed confirms (Rey & Janikas, 2005) arguments that failure to account for space may lead to misleading conclusions.

Overally, the picture emerging from this analysis provides interesting facts about South Africa's regional income disparities. First, the economy is characterised by huge dispersion in levels of income per worker, with most regions' incomes far below national income, while only

[45] In the event that space did not matter the distributions for nationally and regionally conditioned distributions will be the same, which is not the case in panel 2.

[46] We apply the two-sample Kolmogorov–Smirnov test (KS) (Smirnov, 1933), to determine if there are any differences in the distribution functions of income per worker in 2001 and 2011 for both nationally and spatially conditioned distributions. While the first line tests the hypothesis that 2001 distribution has smaller values than 2011, the second line tests the hypothesis that 2011 distribution has smaller values than 2001 distribution and then third line shows results for a combined test. The results in Table 3A in Appendix 1A reveals that the while then nationally conditioned distribution functions for 2001 and 2011 were indeed statistical significantly different, the distribution functions for spatially conditioned distributions we not significant. This implies that failing to control for spatial dependence can lead one to falsely conclude falling disparities, yet in actual fact dispersion in income per worker tend to persist overtime.

a few regions have income above national average[47]. However, it's interesting to note that, during the 2001 – 2011 period regional income disparities decreased, mainly due to increase in relative incomes of poor regions. This implies convergence between rich and poor regions[48]. Comparing the nationally and regionally conditioned distributions show significant differences in terms of their spread and skewedness features. These differences highlights the importance of space in the distribution dynamics of regional income per worker in South Africa. However, once space is accounted for, rather than decreasing disparities, evidence of persistence dispersions in observed. This result highlights the need to properly controlling for spatial dependence to avoid arriving at misleading conclusions.

6. Conclusions

In this paper, we examined the evolution of regional income per work in South Africa over the period 2001-2011, and evaluated the role of space in the process. To avoid the drawbacks of the classical beta and sigma-converegence approach, the distribution dynamics approach was employed. To reveal how dispersed regional income per worker, the paper employed non-parametric kernel density estimates. These are used to reveal the shape dynamics (spread and skewedness features) of regional income per worker. First, using "aspatial" non-parametric kernel density estimates we examined the evolution of regional income per worker relative to national average income. Second, we utilised the exploratory spatial data analysis (ESDA) to check for presence of positive in the distribution of income per worker. Third, we integrated space into the regional income worker variable using the spatially conditioning scheme pioneered by (D. T. Quah, 1996a; D. T. Quah, 1996b). Using the new variable (income per worker relative to neighbouring regions weighted average income per worker) "spatial" non-parametric kernel densities we estimated to evaluate the role of geographical location in the distribution dynamics of regional income per worker.

---

[47] This led (Bosker & Krugell, 2008) to conclude that the distribution of GDP per capita in South Africa was far more unequal in than in other countries at the same level of development with it like Brazil, China and India.

[48] This result contradict earlier findings by (W. F. Krugell, Koekemoer, & Allison, 2005) , as well as (Bosker & Krugell, 2008) who found evidence of increased regional income disparities (divergence). This could be attributed to different data sources, and more importantly the difference could be also any outcome of different time periods (these studies focused on 1996 – 2004 period). A lot could have happened between 2004 and 2011 leading to convergence. For example, this period coincides with the implementation of South Africa's spatial development programmes (Industrial Development Zones, 2001; National Spatial Development Perspective, 2003; 2006).

Three main conclusions emerged from this paper. First, the spread and skewedness of the "aspatial" income per worker distribution revealed evidence of wide dispersion in regional income per worker. A large number of regions had income per worker far below (around 0.5 times) national average, with only a few above national average. While the positive skewedness persisted overtime, the spread of the distribution slightlt fall, confirming evidence of decreasing dispersion across regions over the 10 year period.The decrease was largely driven by increases in relative incomes for poor regions and slight fall in relative incomes of wealthy regions. This confirmed evidence of converegence in income per worker across regions overtime. Second, the distribution of regional income per worker was characterised by positive and significant spatial dependence which tend to decrease overtime. This implies that economic activity and resulting levels of income per worker are not randomly distributed across regions and in particular, regions with income per worker higher (lower) than national average tend to cluster together.

Finally, recognising that neglecting the presence of spatial dependence may give rise to misleading (Rey & Janikas, 2005), we integrated space into the income per worker distribution and estimated "spatial" non-parametric kernel densities. By comparing the "aspatial" and "spatial" density estaimates, the paper revealed that space plays an important role in the evolution dynamics of regional income per worker. In particular, not only was the spread wider and skewedness lower when spatial dependence was considered, but the densities were also more centred around 1. This indicates that while income per worker is widely dispersed across regions in South Africa, on average neigbhouring regions tend to have similar income levels.

From these densities it can be concluded that a substantial part of the shape dynamics can actually be attributed to spatial dependence embedded in the income per worker variable. The picture emerging from this analysis seem to lend little support of the converegence conclusions revealed by the classical converegence approach and the "aspatial" distribution densities. The results obtained fail to find evidence of decreasing dispersion, but rather persistence dispersion pointing to the need to properly account for spatial dependence to avoid misleading conclusions (converegence when in fact its not the case). In essence, the main point made in this paper is that geographical locational is key in explaining the distribution dynamics of regional income per worker in South Africa. The presence of spatial depedence leads to a less homogenous distribution with persisting disparities in regional income per worker.

From a policy perspective, evidence of the importance of geographical location in the distribution dynamics of regional income per worker, implies that regional development needs to pay close attention to space, as a regional intervention will not only benefit the region in question, but its neigbhouring regions as well. Moreover, a shock in one regions will most likely spillover to its surrounding regions. For future research, apart from conditioning regional income by national average and neighbouring regions weighted average incomes, the conditioning can also be done using different factors known to affect regional income per capita such as education, health, gender, among other factors. In trying to understanding fully the factors behind the observed distribution dynamics, results from this paper can be used to improve model specification in the classical convergence approach. For example, evidence of the importance of space points to the need to incorporate the spatial aspects underlying the data when modelling determinants of regional income per worker.

References

Abreu, M., De Groot, H. L., & Florax, R. J. (2004). Space and growth: A survey of empirical evidence and methods.

Anselin, L., & Bao, S. (1997). Exploratory spatial data analysis linking SpaceStat and ArcView. *Recent developments in spatial analysis* (pp. 35-59) Springer.

Anselin, L., & Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. *Statistics Textbooks and Monographs, 155*, 237-290.

Anselin, L., & Rey, S. (1991). Properties of tests for spatial dependence in linear regression models. *Geographical Analysis, 23*(2), 112-131.

Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy,* , 223-251.

Bhorat, H., & Kanbur, R. (2005). Poverty and well-being in post-apartheid south africa: An overview of data, outcomes and policy.

Bosker, M. (2007). Growth, agglomeration and convergence: A space-time analysis for european regions. *Spatial Economic Analysis, 2*(1), 91-100.

Bosker, M. (2009). The spatial evolution of regional GDP disparities in the 'old'and the 'new'Europe*. *Papers in Regional Science, 88*(1), 3-27.

Bosker, M., & Krugell, W. (2008). REGIONAL INCOME EVOLUTION IN SOUTH AFRICA AFTER APARTHEID*. *Journal of Regional Science, 48*(3), 493-523.

Buettner, T. (1999). The effect of unemployment, aggregate wages, and spatial contiguity on local wages: An investigation with german district level data. *Papers in Regional Science, 78*(1), 47-67.

Celebioglu, F., & Dall'erba, S. (2010). Spatial disparities across the regions of turkey: An exploratory spatial data analysis. *The Annals of Regional Science, 45*(2), 379-400.

Dall'Erba, S. (2005). Distribution of regional income and regional funds in europe 1989–1999: An exploratory spatial data analysis. *The Annals of Regional Science, 39*(1), 121-148.
Dall'Erba, S., & Le Gallo, J. (2008). Regional convergence and the impact of european structural funds over 1989–1999: A spatial econometric analysis*. *Papers in Regional Science, 87*(2), 219-244.
De Klerk, M. (2012). *An Analysis of sub–national Economic Databases in South Africa/de Klerk M.,*
Drukker, D. M., Peng, H., Prucha, I. R., & Raciborski, R. (2013). Creating and managing spatial-weighting matrices with the spmat command. *Stata Journal, 13*(2), 242-286.
Durlauf, S. N., Johnson, P. A., & Temple, J. R. (2005). Growth econometrics. *Handbook of Economic Growth, 1*, 555-677.
Ertur, C., Le Gallo, J., & Baumont, C. (2006). The european regional convergence process, 1980-1995: Do spatial regimes and spatial dependence matter? *International Regional Science Review, 29*(1), 3-34.
Ezcurra, R., Pascual, P., & Rapún, M. (2007). The spatial distribution of income inequality in the european union. *Environment and Planning A, 39*(4), 869.
Fischer, M. M., & Stumpner, P. (2008). Income distribution dynamics and cross-region convergence in europe. *Journal of Geographical Systems, 10*(2), 109-139.
Fischer, M. M., & Wang, J. (2011). *Spatial data analysis: Models, methods and techniques* Springer Science & Business Media.
Fujita, M., Krugman, P. R., & Venables, A. (2001). *The spatial economy: Cities, regions, and international trade* MIT press.
Getis, A. (2009). Spatial weights matrices. *Geographical Analysis, 41*(4), 404-410.
Helpman, E. (1998). The size of regions. *Topics in Public Economics,* , 33-54.
Hoogeveen, J. G., & Özler, B. (2006). Poverty and inequality in post-apartheid south africa: 1995-2000. *Poverty and Policy in Post-Apartheid South Africa,* , 59-94.
Johnson, P. A. (2000). A nonparametric analysis of income convergence across the US states. *Economics Letters, 69*(2), 219-223.
Juessen, F. (2009). A distribution dynamics approach to regional GDP convergence in unified germany. *Empirical Economics, 37*(3), 627-652.
Krugell, W. F. (2011). *Putting the economy in its place: Geographical economics in south africa* Potchefstroom: Noordwes-Universiteit, Potchefstroomkampus (Suid-Afrika).
Krugell, W. F., Koekemoer, G., & Allison, J. (2005). Convergence or divergence of south african cities and towns? evidence from kernel density estimates. *Biennial Conference of the Economic Society of South Africa: Development Perspectives: Is Africa Different,*
Krugell, W. F. (2005). *The Geographical Economy of South Africa/WF Krugell,*
Krugman, P. R. (1991). *Geography and trade* MIT press.
Le Gallo, J. (2004). Space-time analysis of GDP disparities among european regions: A markov chains approach. *International Regional Science Review, 27*(2), 138-163.
Le Gallo, J., & Dall'Erba, S. (2008). Spatial and sectoral productivity convergence between european regions, 1975–2000*. *Papers in Regional Science, 87*(4), 505-525.
Leibbrandt, M., Finn, A., & Woolard, I. (2012). Describing and decomposing post-apartheid income inequality in south africa. *Development Southern Africa, 29*(1), 19-34.
Leibbrandt, M., Levinsohn, J. A., & McCrary, J. (2010). Incomes in south africa after the fall of apartheid. *Journal of Globalization and Development, 1*(1)
Leibbrandt, M., Levinsohn, J., & McCrary, J. (2005). *Incomes in South Africa since the Fall of Apartheid,*
Leibbrandt, M., Wegner, E., & Finn, A. (2011). The policies for reducing income inequality and poverty in south africa.

LeSage, J. P. (1999). The theory and practice of spatial econometrics. *University of Toledo.Toledo, Ohio, 28,* 33.
Magrini, S. (1999). The evolution of income disparities among the regions of the european union. *Regional Science and Urban Economics, 29*(2), 257-281.
Magrini, S. (2004). Regional (di) convergence. *Handbook of Regional and Urban Economics, 4,* 2741-2796.
Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics, 107*(2), 407-437. doi:10.2307/2118477
Maza, A., & Villaverde, J. (2009). Spatial effects on provincial convergence and income distribution in spain: 1985–2003. *Tijdschrift Voor Economische En Sociale Geografie, 100*(3), 316-331.
Maza, A., & Villaverde, J. (2005). A spatial analysis of wage differentials for the spanish provinces. *Available at SSRN 835264,*
Maza, A., Villaverde, J., & Hierro, M. (2009). Regional productivity distribution in the european union: Which are the influencing factors?
Meliciani, V., & Peracchi, F. (2006). Convergence in per-capita GDP across european regions: A reappraisal. *Empirical Economics, 31*(3), 549-568.
Moran, P. A. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society.Series B (Methodological), 10*(2), 243-251.
Mossi, M. B., Aroca, P., Fernández, I. J., & Azzoni, C. R. (2003). Growth dynamics and space in brazil. *International Regional Science Review, 26*(3), 393-418.
Naudé, W., Krugell, W., & Matthee, M. (2010). Globalization and local economic growth in south africa, 1994 to 2004. *Local Economies and Global Competitiveness,* , 45-73.
Naudé, W. A., & Krugell, W. F. (2003). An inquiry into cities and their role in subnational economic growth in south africa. *Journal of African Economies, 12*(4), 476-499.
Naudé, W. A., & Krugell, W. (2006). Sub-national growth rate differentials in south africa: An econometric analysis*. *Papers in Regional Science, 85*(3), 443-457.
Paas, T.Modelling regional income convergence in EU-25.
Pfitzner, C. B., & Lang, S. D.ECONOMIC CONVERGENCE: EVIDENCE FROM COUNTIES IN THE CAROLINAS.
Quah, D. (1993a). Empirical cross-section dynamics in economic growth. *European Economic Review, 37*(2), 426-434.
Quah, D. (1993b). Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics,* , 427-443.
Quah, D. T. (1996a). Convergence empirics across economies with (some) capital mobility. *Journal of Economic Growth, 1*(1), 95-124.
Quah, D. T. (1996b). Regional convergence clusters across europe. *European Economic Review, 40*(3), 951-958.
Quah, D. T. (1997a). Empirics for growth and distribution: Stratification, polarization, and convergence clubs. *Journal of Economic Growth, 2*(1), 27-59.
Quah, D. T. (1997b). *Regional cohesion from local isolated actions: II. conditioning* LSE.
Rey, S. J. (2001). Spatial empirics for economic growth and convergence. *Geographical Analysis, 33*(3), 195-214.
Rey, S. J. (2004). Spatial dependence in the evolution of regional income distributions. *Spatial Econometrics and Spatial Statistics,* , 194-213.
Rey, S. J., & Janikas, M. V. (2005). Regional convergence, inequality, and space. *Journal of Economic Geography, 5*(2), 155-176.
Rey, S. J., & Montouri, B. D. (1999). US regional income convergence: A spatial econometric perspective. *Regional Studies, 33*(2), 143-156.

Sala-i-Martin, X. X. (1996). The classical approach to convergence analysis. *The Economic Journal,* , 1019-1036.
Silverman, B. W. (1986). *Density estimation for statistics and data analysis* CRC press.
Van Der Berg, S. (2011). Current poverty and income distribution in the context of south african history. *Economic History of Developing Regions, 26*(1), 120-140.
van der Berg, S. (2014). Inequality, poverty and prospects for redistribution. *Development Southern Africa, 31*(2), 197-218.
Van der Berg, S., Burger, R., Burger, R., Louw, M., & Yu, D. (2006). Trends in poverty and inequality since the political transition.
Van der Berg, S., Louw, M., & Yu, D. (2008). POST-TRANSITION POVERTY TRENDS BASED ON AN ALTERNATIVE DATA SOURCE1. *South African Journal of Economics, 76*(1), 58-76.
Von Fintel, D. (2014). *Spatial Heterogeneity, Generational Change and Childhood Socioeconomic Status: Microeconometric Solutions to South African Labour Market Questions,*
Wittenberg, M. (2014). *Wages and wage inequality in south africa 1994-2011: The evidence from household survey data*
Young, A. T., Higgins, M. J., & Levy, D. (2008). Sigma convergence versus beta convergence: Evidence from US County-Level data. *Journal of Money, Credit and Banking, 40*(5), 1083-1093.

Appendix 1A

Table 1A: 20 highest and lowest income regions in 2001 and 2011.

2001		2011	
20 richest regions			
MD_NAME	rincome	MD_NAME	Rincome
Randburg	20059	Randburg	20708
Cape	16533	Cape	16031
Pretoria	16004	Bronkhorstspruit	15898
Goodwood	14065	Goodwood	15667
Wynberg	13854	Wynberg	15420
Germiston	13776	Pretoria	15239
Roodepoort	13405	Nelspruit	14203
Johannesburg	12679	Johannesburg	13102
Simonstown	11779	Roodepoort	12707
Durban	11382	Germiston	12061
Pinetown	10916	Bellville	11948
Bellville	10897	Pietersburg	11937
Somerset West	10652	Simonstown	11749
Stellenbosch	10093	King William's Town	11378
Boksburg	9541	Somerset West	11367
Kempton Park	9404	Durban	10857
Wonderboom	9134	Pinetown	10653
Pietersburg	8999	Wonderboom	10636
Benoni	8940	Highveld Ridge	10461
Kuilsrivier	8702	Boksburg	10054
20 poorest income regions			
MD_NAME	rincome	MD_NAME	rincome

Clocolan	1460	Joubertina	2529
Hoopstad	1945	Calitzdorp	2989
Ngotshe	2027	Ndwedwe	3008
Bultfontein	2051	Kenhardt	3023
Zastron	2187	Wakkerstroom	3063
Marquard	2202	Jansenville	3077
Wakkerstroom	2257	Koppies	3112
Kranskop	2265	Bolobedu	3151
Botshabelo	2268	Kirkwood	3179
Rouxville	2285	Paulpietersburg	3190
Koppies	2289	Ntabethemba	3204
Joubertina	2316	Mapumulo	3207
Ventersburg	2353	Hewu	3225
New Hanover	2418	Alfred	3284
Kenhardt	2455	Ventersdorp	3285
Kirkwood	2480	Mthonjaneni	3294
Smithfield	2489	Nkomazi	3297
Fouriesburg	2511	Mkobola	3316
Vrede	2515	Keiskammahoek	3326
Winburg	2523	Lady Frere	3332

Notes: rincome is monthly regional income per worker in rands.

Table 2A: Unconditional convergence in regional income per worker, 2011-2011

	OLS Model
VARIABLES	growth_rate
logincomeperworker_2001	-0.307***
	(-0.0301)
Constant	2.769***
	(-0.253)
Observations	354
R-squared	0.245

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3A: Two-sample Kolmogorov-Smirnov test for equality in distributions

Distribution	Nationally conditioned			Spatially conditioned		
Smaller group	D	P-value	Corrected	D	P-value	Corrected
Relative income per worker 2001	0.1384	0.001		0.0706	0.171	
Relative income per worker 2011	-0.0085	0.975		-0.0452	0.485	
Combined K-S	0.1384	0.002	0.002	0.0706	0.340	0.307

Notes: Looking first at the nationally conditioned relative income per worker distribution function: A significant p-value (0.001) in the first line confirms that income per worker values for 2001 were smaller than those for 2011 and as expected the p-value (0.975) in the second line is insignificant showing that 2011 values were not smaller than 2001. Finally, the p-value (0.002) for the combined test is significant, which confirms that the distributions are different. On the other hand looking at the line, second and third lines for the spatially conditioned distribution function reveals that the distributions for 2001 and 2011 are not significantly different. This confirms evidence of persistent regional disparities.

Appendix 2B: Additional Figures

Figure 1B: Cross-sectional distribution of regional income per worker.

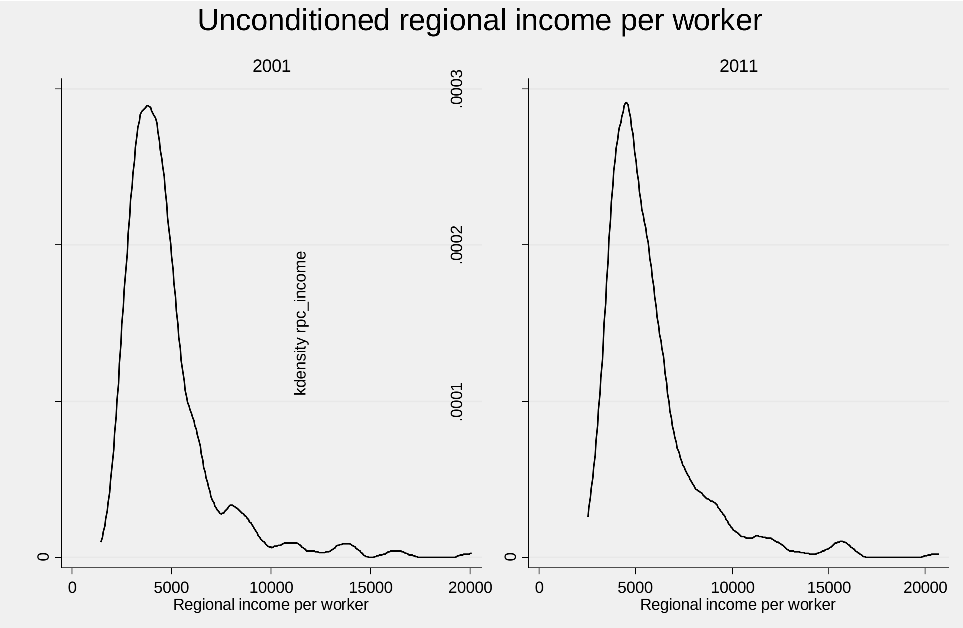


Figure 2B: distribution dynamics of regional income per worker.

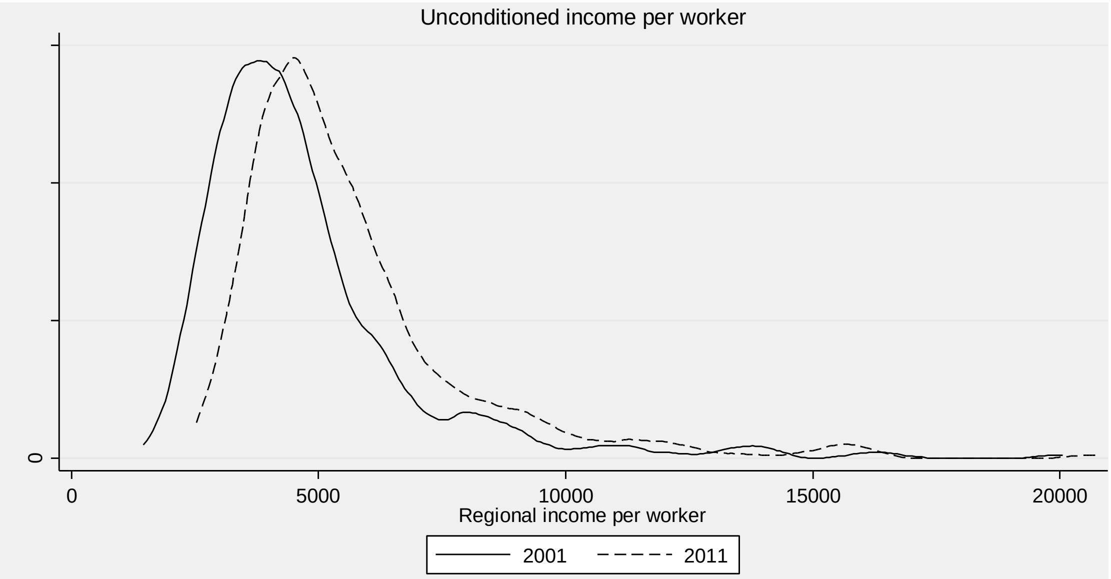
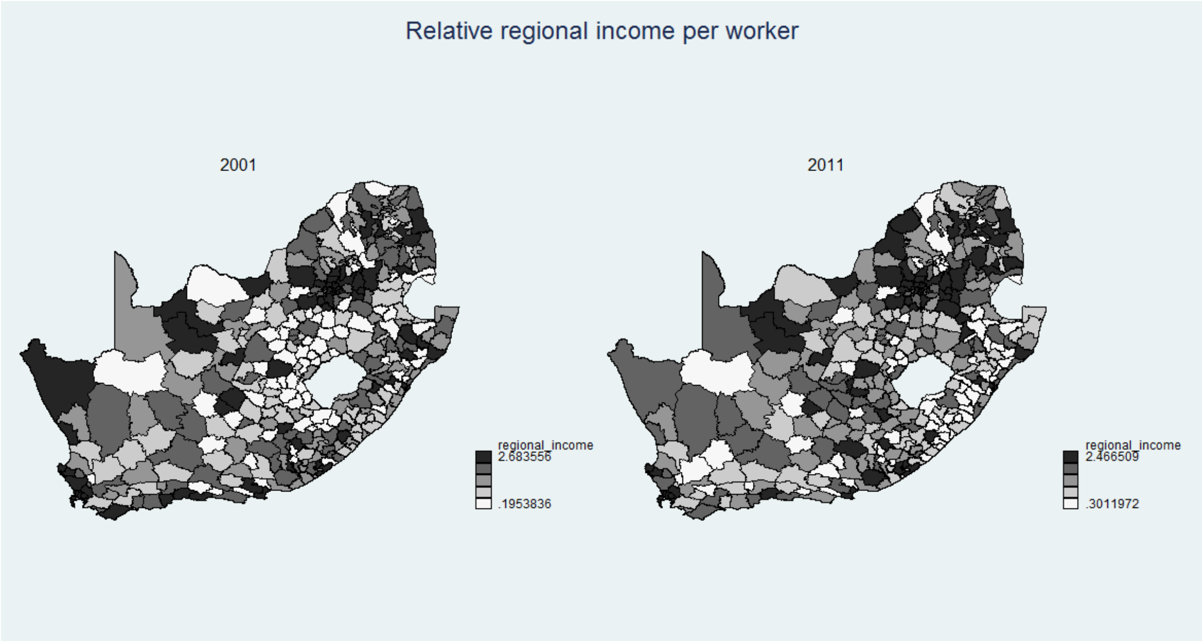


Figure 3B: The spatial distribution of regional income per worker, 2001-2011.



Notes: On the map colour intensity increases with the level of income, implying that high income regions are shown by dark colour, while low income regions are shown by light shade. From the legendary key, the top 20 percent regions have incomes 2.68 and 2.47 times of national average in 2001 and 2011, while the bottom 20 percent regions had incomes 0.20 and 0.3 times of national average.

Figure 4B: The spatial distribution of spatially conditioned income.