

Hakizimana, Jean-Mari; Geyer, Hermanus

**Conference Paper**

## Socio-economic inequality in South Africa according to different disparity indices

54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Hakizimana, Jean-Mari; Geyer, Hermanus (2014) : Socio-economic inequality in South Africa according to different disparity indices, 54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124233>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# **Socio-economic inequality in South Africa according to different disparity indices**

J-M. V. Hakizimana  
Statistics South Africa  
Pretoria  
South Africa

H. S. Geyer (Corresponding author)  
Centre for Regional and Urban Innovation and Statistical Exploration (CRUISE)  
Department of Geography  
Stellenbosch University  
Stellenbosch  
South Africa  
hsgeyer@sun.ac.za

# **Socio-economic inequality in South Africa according to different disparity indices**

## **Abstract**

The purpose of the paper is to identify areas in South Africa where social-economic disparity exists using 2011 census data. Different indices are used to measure spatial disparity with the aim of finding the most appropriate approach for measuring disparities under different circumstances. The following measures were used in the study: the multidimensional composite index of deprivation; range ratio; relative mean deviation; standard deviation of logarithms; Gini coefficient; Kuznets ratio; Theil inequality index, mean logarithmic deviation, and the Atkinson index. In the study settlements are either regarded as individual settlements or contiguous settlements as delimited at the main place level in the census. Due to the fragmenting impact of apartheid on the South African society, different measures tend to be the most appropriate in different parts of the country. This implies that different policy interventions are needed to address area-specific challenges.

*Keywords:* Spatial disparity, Poverty, Deprivation, Inequality, Indices, South Africa  
*JEL:* D63, I32

## **Introduction**

World-wide there is a growing concern that spatial inequality is on the increase, globally and locally (UNDP, 2013; KANBAR and VENEABLES, 2005; GOSCHIN et al., 2008; HERRMANN-PILLATH et al., 2002; PENG et al., 2006; MOHIUDDIN and HASHIA, 2012; REY and JANIKAS, 2005; SAHN and STIFEL, 2003; VENKATANARAYANA, 2004; to list only a few). This is partially due to the gradual shift in global dynamics, driven by the fast-growing economies of the developing world (WORLD BANK, 2009), but it is also caused by abrupt political economic changes that could occur at any specific point in time. Increasing inequalities due to the expansion of the EU (EUROPEAN UNION, 2007) and the democratisation of South Africa are examples of the latter cases in point. Paradoxically, growing convergence is an expected outcome of development. As economies grow, economic inequality tends to follow suit. When production increases it tends to become spatially more concentrated since particular locations (such as countries that are consistently maintaining higher levels of productivity than others, connected cities and resource rich areas, to name but a few) are generally preferred by producers (WORLD BANK, 2009). It is the challenge of governments of such countries or regions to institute policies that will make living standards of people more uniform across space. Exactly that is the expressed aim of the South African government's National Development Plan (NPC, 2011), an initiative that took its cue from earlier similar

initiatives by the governments of India and China and one that is aimed at incorporating ideals set out in the country's earlier Industrial Policy Action Plan of the Department of Trade and Industry and the Economic Development Ministry's New Growth Path (ZARENDA, 2013). The aim of this paper is to assist in this process by finding ways of accurately computing spatial inequality in South Africa.

Spatial statistics can be used as a tool to measure the levels of spatial disparity between different communities as well as to take stock of the efficacy of public policy in addressing socio-economic disparities (WISHLADE and YUILL, 1997). Because there are often practical difficulties associated with the measuring of disparities due to local specific circumstances, it is important to find an appropriate approach in measuring spatial inequality by looking at significant differences between settlements in different regions.

This paper looks at different approaches of measuring inequality in order to find the most appropriate measure for particular circumstances in South Africa. It provides a comparative analysis of different measurements of spatial disparities with the aim of selecting the most appropriate approach based on available population and housing census data. These measures make it possible to identify areas of extreme disparity based on the spatial clustering of local communities displaying negative results.

The paper starts by explaining the current understanding of disparities in order to properly define the concept and remove possible ambiguities with reference to other aspects of measurement of spatial disparities in the literature. In the second section, it presents different approaches used in measuring social-economic disparities based on the available data for the specific place. Local variations in the spatial disparities are presented in the third section of the paper based on the clustering of the local Moran statistics. The fourth section explains the approach used in determining the extreme disparities appropriate for a particular place and the final section gives the results applied to South Africa using the 2011 population census (STATISTICS SOUTH AFRICA, 2013).

### **Approaches to spatial disparity**

The concept of *spatial disparity* depends on one's understanding and perception of its meaning based on social-economic factors as well as the type of settlements requiring appropriate policy intervention. The International Labour Organisation (ILO) defines spatial disparities as the differences that are observed in the levels of economic performance and welfare between countries or regions. The OECD defines it as differences in the intensity of economic phenomena in a region or country (OECD, 2002; 2003).

In terms of the Lisbon Treaty, the European Union's management policy is aimed at monitoring progress made in the level of cohesion in the development of different social groups and regions over time. The aim is for cities and regions to move closer to the average point of convergence which would require less political intervention as opposed to those places that require more intervention to address divergences (MOLLE, 2007). Areas falling within particular levels of the deprivation index are designated as *poor* (KUTSCHERAUER at al., 2010).

Theoretically, the type of spatial disparities affecting a particular area may be detected on the basis of hypothetical judgements with reference to particular characteristics. On the one hand areas have to be identified, and changes in their structure assessed, in terms of characteristics that are regarded as negative. Such an assessment allows one to obtain a better understanding of the specific reasons for comparative advantages and disadvantages. In this manner one may define negative spatial disparities based on weaknesses (i.e. conditions of vulnerability or lack of abilities) whilst the positive spatial disparities would be based on the strengths (competitive advantages with regard to specific competences).

In a chain of cause and effect the forms of spatial disparities determine the characteristics that are chosen, and that determines the approach that is used in determining spatial disparities. Some approaches focus on the selection of indicators either by using specific indicators of the sub-area or using selected indicators aggregated for the entire area. The most important factor to consider is to use the approach that is providing the most relevant and usable results to the planners and politicians (KUTSCHERAUER at al., 2010).

The manner in which spatial disparity is presented, could result in it being viewed as a multidimensional problem which may influence choices made in spatial policy. Therefore, the type of spatial disparity that is appropriate for a particular area depends on the spatial disparity classification and its associated relevant attributes (KUTSCHERAUER at al., 2010). Disparities may be determined from place-related issues and could be classified as either physical disparities (associated with geographical or natural conditions), economic disparities (concerned with differences in the quality or quantity of output of a region), or social disparities (concerned with the standard of living of the population).

Inequalities between regions are often the cause of social problems (STILLWELL et al, 2010). Social disparities affect different people in a country differently because they impact their physical, social and economic mobility, their ability to improve their well-being, and could negatively affect the quality of an individual's social, cultural and religious experiences.

The measurement of spatial disparities could be used as a tool to guide local policy intervention. The aim of local policy intervention is to reduce the effect of spatial disparities by increasing the economic competitiveness of regions and to stimulate diffusion of development (VITURKA, 2006; 2010). The other approach is to set standards and determine the spatial disparities as deviations from the standard.

Looking at some examples, Papua New Guinea has used an index of development of districts to identify less developed areas (D'SA, 1986). In Tunisia, a study on spatial disparities shifted from the economic to the non-monetary dimensions of inequality by using a multidimensional perspective in computing the Spatial Composite Index of Welfare (KRIAA et al., 2011). In China, using economic census data from 1978 to 2005, spatial disparity was based on a trend analysis, differentiating between urban and rural and decomposing it into within-group and between-group components (LI and XU, 2008). In the EU a policy has been adopted targeting harmonious, balanced and sustainable spatial development. These goals are based on the European Union's policy of economic, social and spatial cohesiveness as stated in the Treaty of Lisbon (EC, 2007). The European Spatial Development Perspective highlights the spatial disparity challenges in terms of output, productivity and employment across countries where the local level is considered as an essential analytical unit of spatial development (LENGYEL 2009).

The South African example is published by a team of researchers from the Oxford University calculating the Index of Multiple Deprivation (SAIMD) on small areas using the 2001 Census (NOBLE *et al.*, 2010). Subsequently, similar indices have been computed by the same team such as the South African Index of Multiple Deprivation for Children 2001 at municipal level based on a 10 per cent sample of the 2001 Census (BARNES *et al.*, 2007).

There are other approaches beyond the local average at a specific place where spatial disparities are measured, taking distance-decay functions into account. Here it is argued that proximity (or distance) between higher and lower income areas could have a significant impact on socio-economic disparities. This is done by conducting a spatial autocorrelation analysis and determining the level of similarity (positive spatial autocorrelation) or dissimilarity (negative spatial autocorrelation) as clustered in a particular area. Therefore, the local Moran indicator of spatial association is used as an approximation of spatial disparities. This means that there is a neighbourhood disparity or spatial disparities at place names level when there is some difference in value for a given variable between neighbouring places or suburbs (CHAKRAVORTY, 1996).

## Measuring disparity

In this paper, several approaches of measuring disparity are followed using the 2011 census results. The kinds of disparity measures presented are spatially and socio-economically inclined making it possible to focus on a variety of key determining factors. The different approaches considered for measuring spatial disparities in South Africa are:

- The Multidimensional Composite Index of Deprivation (ALIKE, et al. 2011);
- The range ratio;
- The relative mean deviation;
- The standard deviation of logarithms;
- The GINI coefficient;
- The Kuznets ratio R;
- The Theil inequality index
- The mean logarithmic deviation
- The Atkinson index

The above measures of disparity provide a large spectrum of the attributes as determinants of spatial disparities. However, they have some inherent limitations that have to be taken into consideration when determining the appropriate measure of disparity in a particular area. When computing the above measures of disparity, one would be interested to know the main causes underlying the observed spatial disparity patterns. This will help identifying the socio-economic characteristics that were causing such conditions and will better inform local planners of the kinds of spatial disparities that are prevalent in the area and the kinds of steps that need to be taken to improve the situation.

## Indices of deprivation

Measuring deprivation at the *main place level* (of the 2011 South African census) allows the detection of geographical patterns of social advantage or disadvantage following social change, economic variation and change in living spaces. The spatial measure of multi-dimensional deprivation also allows targeting the most deprived areas through specific policies (SMITH, 1999; KLEINMAN, 1999; SMITH *et al.*, 2001) and assists in the detection of variations in the mix of problems in an area (NOBEL et al., 2009).

Based on the definition by TOWNSEND (1987), the measure of deprivation is a status representing communities lacking in areas such as clothing, housing, household facilities and educational, working and social conditions. Therefore, *deprivation* is defined as the status of people not meeting their needs as oppose to *poverty* which refers to a lack of resources required for their needs. By extension, the collection

of each type of deprivation into one aggregated model is conceptually considered as multiple deprivations (TOWNSEND, 1987).

The concept of multiple deprivation permits the use of different measurements from the diverse elements of deprivation such as education deprivation and health deprivation. In the case of income, the deprivation from income may not be considered as components of multiple deprivation measure. This is because the deprivations of material or services may be regarded as proxy measures of low income (NOBEL et al., 2009).

In the study, a data set was first prepared by identifying the different variables necessary for computing from the census data. The proposed variables that will reflect the multidimensional character of the process are shown in Table 1. After more variables have been considered, the multidimensional composite index of deprivation is computed using weighted scores derived from a factor analysis.



Table 1: Deprivation indicators and specification of the deprivation condition

Sector	Indicator	Deprived if
Education	Primary year of schooling	Number of household members that have completed 7 years of schooling
	Secondary year of schooling	Number of household members that have completed 12 years of schooling
	Child school attendance	Any school-age child not attending school until grade 7
Health	Child mortality	Child less than 5 year who died a year before the Census reference
	Adult mortality	Person above 20 year who died a year before the Census reference
	Parenthood surviving	Any family not having both parents alive
Living standard	Electricity	Household without any electricity
	Sanitation	No access to flush toilet (use of bucket, others)
	cleaned water	No Access to piped water
	Poor construction material	House in wood, plastic, cardboard, mud, wattle, thatch, other
	Non sustainable energy for cooking	Household cooking using wood, animal dung, charcoal, others
	Lack of asset ownership	Household not owning household assets : telephone, refrigerator, car, computer, cellphone, electric stove
Housing	No ownership of the house	House not owned
	Value of property	Amount in Rand
	Age of property	Number
Economic	Lack of employment	Proportion of unemployed household members
	Income	Rand
	Potential youth dependency ratio	Total of persons that are less than 15 and those between 15 and 65
	Real youth dependency ratio	Total of those that are less than 15 and those between 15 and 65 gainfully employed
	Potential elderly dependency ratio	Total of those above 65 and those between 15 and 65
	Real elderly dependency ratio	Total of those above 65 and those between 15 and 65 gainfully employed
	Minimal total dependency ratio	Total of those that are less than 15 and those above 65 divided by those between 15 and 65
	Real total dependency ratio	Total population divided by those employed
	Potential dependency ratio	Total population divided by those between 15 and 65
Demographic	Household size	Number of persons
	Age of the head of household	Age
	Gender differential	Female = 1 and Male = 0
	Majority as Black	Value 1 or 0
	Majority as Coloured	Value 1 or 0
	Majority as Indian	Value 1 or 0
	Majority as White	Value 1 or 0
	Citizenship of head	Non South Africa = 1, South African 0

Using a similar approach in the calculation of the South African Index of Multiple Deprivation (SAIMD), it was necessary to make sure there is no issue of reliability due to small numbers. There is a slight limitation of the South African censuses as they have been adjusted for undercounts accurate within the post-enumeration survey domain of estimation (i.e. province, urban and non-urban). For instance, the use

of the *enumeration area* or *sub-place* as the smallest geographical unit may require aggregating the numbers toward a larger area (*main places*, for instance) in order to avoid small cells numbers. The aggregation does not mean that the indicator becomes smaller (i.e. less deprived). The extent of the shift depends on the reliability of the indicator and the heterogeneity of the larger area (NOBLE *et al.*, 2009).

The multidimensional composite index of deprivation is computed using the socio-economic dimension – health, education, housing, demographic, economic and living standard – as a single summary measure expressed in value (e.g. proportions of people or of households experiencing a type of deprivation). The single index of deprivation for each area is calculated using

$$I_{ij} = \frac{\sum_{k=1}^{n_{ij}} x_{ijk}}{\sum_{k=1}^{n_{ij}} n_{ijk}}$$

where  $x_{ijk}$  is an indicator variable and  $x_{ijk} = 1$  means the lack of a particular need for a household or person  $k$  in terms of variable  $j$ , at place  $i$ . The deprivation within each dimension at a given place will be the proportion of the people within the households within the selected dimension- or domain-variable at a given place.  $n_{ijk}$  represents the total number of persons for household  $k$  within  $j$  at  $i$ . The above-mentioned deprivation index of the selected dimension is 1 if all people (in households) are deprived while it will be zero if the characteristics considered as a measure of the dimension are not available (i.e. nobody is lacking in the characteristics of the expressed needs). Therefore,  $0 \geq I_{ij} \leq 1$ .

The multi-dimensional index of deprivation ( $C_i$ ) is computed as

$$c_i = \sum_{j=1}^d w_{ij} I_{ij}$$

where  $w_{ij}$  represents the weight or relative importance of dimension-variable  $j$  with the total number of dimension-variable  $d$  and  $\sum_{j=1}^d w_{ij} = 1$ .

Even though each socio-economic variable may be considered as an independent experience of deprivation, each deprivation measure contributes to a greater understanding of the concept of multiple deprivations. Therefore, their contribution depends on their relative importance. The use of proportion permits a transformation of those socio-economic indices into a common distribution with the same range and with the maximum and minimum ranging between 0 and 1 at unitary scale.

The implicit weights are derived from the underlying distributions of the data obtained from the factor analysis of the variables representing the different elements of deprivation needs. The reason of using the

factor analysis is to accurately identify the most significant socio-economic variables that could impact the deprivation as well as reduce the number of variables initially under consideration. Therefore, the loading of the variables on the significant factors (corresponding to weights associated to each deprivation dimension  $w_{ij}$ ) provides the relative importance of each socio-economic variable in each factor.

### Indices of dispersion of income

Using the using information from the censuses, the statistical units counted are the number of households and the value of income (Rand). The lowest spatial unit is the main place having  $n$  households,  $i = 1, \dots, n$ , where  $y_i$  is a variable such as income of the household  $i$ . The appropriate measure of disparity using income is to standardise using the average household income. The mean or average household income is calculated for place  $j$  as

$$\mu_j = \frac{1}{n_j} \sum_{i=1}^{n_j} y_{ij}$$

The range ratio ( $R_j$ ) for each main place  $j$  is a crude but simple measure of disparity computed by dividing the difference between the highest and lowest income by the mean income for each place  $j$

$$R_j = \frac{(\text{Max}_i y_{ij} - \text{Min}_i y_{ij})}{\mu_j}$$

Those households having the same income give an  $R_j = 0$  whilst the other extreme gives  $R_j = n_j$  (where  $n_j$  is the size of the population in each place  $j$ ). The disadvantage of using the range is that it does not take into account the distribution between the extremes (PAUW, 2003).

### Relative mean deviation

The relative mean deviation ( $M_j$ ) for each main place  $j$  is based on a comparison of the income of each observation and the mean income by using the sum of the absolute differences between the income of each household and the mean income divided by the total income (mean income times number of observations) for each place  $j$

$$M_j = \frac{1}{n_j \mu_j} \sum_{i=1}^n |y_{ij} - \mu_j|$$

If one household is contributing to the overall income of the place,  $M_j = 2(n_j - 1)n$ . However, when the income is equally distributed across all households around the mean income,  $M_j = 0$ . The disadvantage of the relative mean deviation is the violation of the principle where the contribution of income from poor households to rich households implies the increase of the inequality (PAUW, 2003).

## Coefficient of variation

The coefficient of variation ( $CV_j$ ) of each main place  $j$  is calculated by taking the standard deviation divided by the mean. The coefficient of variation is part of the generalised entropy with the property of using equal weights to transfers at different levels of income.

$$CV_j = \frac{\sqrt{V_j}}{\mu_j}$$

Therefore, it uses the standard deviation which is the square root of the variance computed as a deviation from the mean with squared differences accentuating the magnitude of the differences further away from the mean (SEN, 1997)

$$V_j = \frac{1}{n_j} \sum_{i=1}^{n_j} (y_{ij} - \mu_j)^2$$

The use of the variance is not respecting the income scale independence axiom as for instance the increase of the income results in a further increase of the variance (i.e. the doubling of all the income results in quadrupling of the variance).

## Standard deviation of logarithms

The standard deviation of logarithms ( $H_j$ ) of each main place  $j$  gives greater importance to lower income levels as staggered and it is calculated as the logarithm of the arithmetic mean. The standard deviation of logarithms attaches more weight to transfers at the lower end of the income distribution (LITCHFIELD, 1999).

$$H_j = \sqrt{\frac{1}{n_j} \sum_{i=1}^{n_j} (\log(y_{ij}) - \log(\mu_j))^2}$$

## Measuring inequality

### Gini coefficient

The Gini coefficient ( $G_j$ ) uses the concept of income shares of groups or households with reference to the Lorenz curve for each place of analysis where the cumulative percentage of households is plotted against the cumulative share of income. It widely used as it has some advantages such as giving one value reflecting the overall income differential as well as providing the platform for international comparison and for decomposition analysis at the local level (Xiaoli, 2010). There are various formulas that can be

used to calculate the Gini coefficient. SEN (1997: 31) indicated that the Gini coefficient is equal to one half of the relative mean difference, which is defined as the arithmetic mean of the absolute values of differences between all pairs of incomes (PAUW, 2003).

$$G_j = 1 + \left(\frac{1}{n_j}\right) + \left(\frac{2}{n_j^2 \mu_j}\right) [y_{1j} + 2y_{2j} + \dots + n_j y_{n_j j}] \text{ for } y_{1j} \geq y_{2j} \geq \dots \geq y_{n_j j}$$

### **Kuznets ratio**

The Kuznets ratio ( $K_j$ ) for each main place  $j$  is computed as the sum of absolute value of the difference between the proportion of income ( $p_{ij}$ ) and population share ( $f_{ij}$ ) giving a ratio between 0 and 1.

$$K_j = \sum_{i=1}^{n_j} |f_{ij} - p_{ij}|$$

The index permits the detection of extreme disparity using the income share of 20 per cent and above. For each population group or place, the income share is based on the different proportions of the number of people in each group or place compared with the corresponding shares of the income. The benefit of using aggregated data is that the missing value or incomplete data has little effect on the measurement of distribution of the disparity of different areas (XIAOLI, 2010).

### **Theil index**

The Theil index of inequality ( $T_j$ ) measures economic inequality by taking the sum of the ratio of each household's income relative to the mean of all household incomes in a given place, and then multiplies this by the log of the same ratio

$$T_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \left(\frac{y_{ij}}{\mu_j}\right) \log\left(\frac{y_{ij}}{\mu_j}\right)$$

### **Mean logarithmic deviation**

The mean logarithmic deviation ( $GE_j$ ) also uses the income share in the population:

$$GE_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \log\left(\frac{\mu_j}{y_{ij}}\right) = \log \mu_j - \frac{1}{n_j} \sum_{i=1}^{n_j} \log y_{ij}$$

### **Atkinson index**

The Atkinson index ( $A_j$ ) is also a measure of income inequality with the ability to highlight movements in particular segments of the distribution

$$A_{j\epsilon} = 1 - \left[ \frac{1}{n_j} \sum_{i=1}^{n_j} \left( \frac{y_{ij}}{\mu_j} \right)^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}} \text{ Where } 0 < A_{j\epsilon} < 1$$

The index uses an adjustment factor of a particular segment of the income distribution which represents the level of *inequality aversion* denoted by  $\epsilon$ . The range of the inequality aversion parameter,  $\epsilon$ , is  $0 < \epsilon < \infty$ , making the Atkinson index more sensitive to changes at the lower end of the income distribution. The reduction of the level of inequality aversion toward zero influences the Atkinson index with more sensitivity due to changes in the upper end of the income distribution (MONFORT, 2008). If the value of  $\epsilon$  is higher, the community is more concerned with the issue of inequality (LITCHFIELD 1999). When the value  $A_{j\epsilon} = 0$ , this implies there is perfect equality in the community having achieved the full fair distribution of income. The calculated Atkinson index can be decomposed into population groups having a similar total income disparity as the aggregated income disparity within group income and among groups (XIAOLI, 2010).

### **Spatial clustering of local variation**

The different disparity indicators computed for South Africa represent a generalisation value associated with each target area (main place) using either the household average income or proportion of deprivation derived from census data. Spatial autocorrelation was used to confirm the robustness of the clustering of areas plagued by disparity (KRIAA et al, 2011).

To effectively portray the spatial pattern of disparity in the country it was important to take the differentials in spatial inequalities and contiguity between polygons (areas) into account. Each polygon representing a place or area has one value of disparity assigned to it as part of the map generalisation. Therefore, polygons with a higher disparity index value than its neighbours were regarded as disadvantaged or peripheral in terms of needs as opposed to advantaged areas. These spatial relationships are determined using spatial autocorrelation and spatial heterogeneity techniques.

Measuring spatial relationships is often done by comparing the relative positions of spatial observations at specific locations with those of neighbouring locations using the global and local spatial autocorrelation function Moran's I statistic is well-known for measuring global spatial autocorrelation. However, the

global Moran's *I* statistic does not permit the assessment of the existence of spatial autocorrelation for small areas. Instead, local spatial variation and extreme observations are measured using the *local Moran* commonly known as the *local indicators of spatial association* (LISA) (ANSELIN, 1995). It tests the hypothesis of random distribution by comparing the values of each specific location with the values of nearby locations (KRIAA et al, 2011). The computed local indicator of spatial association is used mainly as the indicator measuring the significance of the local spatial clustering of the different locations. Also, the local indicator of spatial association assists in testing the diagnostics for local variation and detection of the significant places considered as outliers.

### **Places displaying extreme disparities in South Africa**

Main places in the 2011 Census, 14039 in total, were classified into clusters of high or low values for each indicator of disparity. A two-pronged approach was used. First, spatial disparities for each place were calculated for all indicators of disparity. Each place associated with category HH was marked among all indicators after they have been put through the local association of grouping into clustered categories. Next, places were either given *disparity* or *extreme disparity* status using voting rules. When a particular place attained a HH outcome based on one disparity index only, it was given a *disparity* designation. If based on more than one, it was designated *extreme disparity*. The disparity index contributing mostly to that particular designation is determined by the index resulting in the smallest z-score and the highest p-value. This disparity index also serves as a broad indication of the kind of policy intervention that would be required to improve conditions there.

In the computation of a multidimensional composite index, a factor analysis had to be conducted to determine the variance explained by each factor. The multidimensional composite index is a result of the weighted factors where each weight corresponds to the proportion of the variance explained by each factor. However, there are critical validation processes that assist in the computation of the scores loading that might influence the ranking of the main places based on their socioeconomic inequalities and outcomes (KRISHNAN, 2010).

One of the issues to check when calculating the spatial disparities index is to find the appropriate aggregation strategy combining the multidimensional variables into one composite index for any given main place. The other problem is to be able to standardise to avoid giving variables with different measurement units, disproportionate ranges and removal of skewness (GILTHORPE, 1995).

To check multicollinearity problems the Kaiser-Meyer-Olkin (KMO) measure is usually used to ensure that the data are appropriate for factor analysis using sampling adequacy prediction based on correlations and partial correlations. The KMO measure compares the magnitudes of the observed correlation

coefficients to the magnitudes of the partial correlation coefficients. The KMO gives a value of 0.815 which allow rejecting the null hypothesis of uncorrelated variables based on the Bartlett's Test of Sphericity. The factor analysis results, based on the variables shown in Table 1, are shown in table 2. To save space only the variables loading the highest on each factor are shown. The variance explained by each factor is shown in the last row of the matrix (Table 2).

Table 2: Factor Matrix

Z-score variables	Factors				
	1	2	3	4	5
Potential youth dependency ratio	.908	-.094	-.183	.021	-.224
% of household with those under 14	.892	-.256	-.202	.017	-.147
Minimal total dependency ratio	.876	.225	-.255	-.004	-.252
% of household with those between 15 and 64 years old	-.867	-.184	.319	.030	.189
% of those achieving at least Matric and above	-.764	.024	.107	-.073	-.444
% of Household's female Headed	.729	-.045	-.181	-.166	-.054
% of household with No flush toilet	.699	.051	-.191	.000	.294
% of household with Non good energy for cooking	.684	.049	.323	.158	-.185
% of person with No school attendance	-.625	.342	.012	.264	.231
% of household with poor construction material	.620	.190	.337	.129	.152
% of household with No access to piped water	.564	.099	.321	.152	.056
% of household with those 65 +	-.033	.953	-.178	-.049	-.080
Potential elderly dependency ratio	.130	.924	-.220	-.035	-.139
Mean age of persons	-.559	.697	.070	.008	.069
% of household with No electricity	.519	.118	.564	.367	-.064
% of household with Lack of asset ownership	.450	.197	.554	.329	-.021
% unemployed	-.271	-.174	-.430	.675	-.012
% not economically active	.454	.100	.322	-.654	.338
% of those achieving at least primary and some secondary	.447	.082	-.342	.224	.667
% variance explained	39.9	14.1	9.2	7.1	6.1

The percentage of the total variation explained by each factor represents the weight to be applied on the factor score coefficients when computing the composite index for each main place.



The results of different types of measurement of disparity indices are shown in Table 3.

Table 3: Comparison of 2011 disparities indices per province

SPATIAL DISPARITIES INDEX	SA	WC	EC	NC	FS	KNT L	NW	GP	MP	LP
Multidimensional Composite Index (Cj)	0.00	-0.83	0.18	-0.44	-0.61	0.07	-0.43	-1.13	-0.47	0.04
Standardised Range Ratio (Rj)	-0.01	0.42	-0.32	0.02	0.66	0.06	0.41	1.11	0.72	0.21
Relative mean deviation (Mj)	0.86	0.80	0.86	0.83	0.85	0.84	0.81	0.82	0.97	0.89
Coefficient of variation (CVj)	2.43	2.58	2.08	2.39	2.86	2.68	2.76	2.85	3.06	2.72
Standard deviation of logs (Hj)	0.17	0.11	0.20	0.15	0.12	0.15	0.13	0.08	0.12	0.15
Gini Coefficient (Gj)	0.60	0.62	0.58	0.61	0.63	0.62	0.64	0.65	0.65	0.63
Kuznet ratio(Kj) (*0.001)	0.05	0.26	0.01	0.03	0.11	0.03	0.06	1.13	0.08	0.03
Atkinson index (Aj) (*0.1)	9.98	9.87	9.98	9.97	9.96	9.99	10.0	10.0	9.96	9.98
Mean log deviation (GEj) (*0.1)	0.18	0.11	0.25	0.16	0.11	0.14	0.11	0.06	0.11	0.13
Theil index (Tj)	1.94	2.31	1.70	1.97	2.39	1.96	2.12	2.73	2.64	2.18
Proportion households (PHj)	100%	11%	12%	2%	6%	17%	7%	28%	7%	10%
Proportion income (PIj)	100%	15%	8%	2%	4%	14%	5%	41%	6%	6%

Note: SA=South Africa; WC=Western Cape Province; EC=Eastern Cape Province; NC=Northern Cape Province; FS=Free State Province; KNTL=KwaZuluNatal; NW=North West Province; GP=Gauteng Province; MP=Mpumalanga Province; LP=Limpopo Province

Prominent composite results from Table 3 are the following:

- The multidimensional composite index of deprivation indicates that the settlements in Eastern Cape, KwaZulu-Natal and Limpopo are the most deprived associated to very poor areas.
- The range ratio presents varying disparities in Limpopo, Western, KwaZulu-Natal, Gauteng and Mpumalanga being wider than national range ratios of household incomes whilst the Eastern Cape shows relatively narrow range ratios.
- The relative mean deviation is around 0.8 across provinces as not much difference can be detected using the relative mean deviation as there is a hidden disparity due to the contribution of poor households toward the affluent households in economic transactions.
- The coefficient of variation in Eastern Cape is among those provinces with low disparities. High income disparities are detected in North West, Mpumalanga, Limpopo, Gauteng and Free States.
- The standard deviation of logarithms confirms the Eastern Cape as a province with settlements displaying extreme disparities.

- The Gini coefficient for the Eastern Cape displays low reading of equality because of low revenue in rural areas. The average of the rest of the provinces displays Gini coefficients around the national level of 0.60.
- The mean logarithmic deviation is high for Eastern Cape.
- The Theil inequality index also considers the information shared by the population as contributing to differentials between places. Eastern Cape and Northern Cape seems not to have many elements differentiating the population with their information shared.
- The Atkinson index is not able to discern the extreme spatial difference between places.

Using the measure of spatial dependence developed by MORAN (1948), the relative measure of spatial dependence on main place locations was computed for each of the disparity indicators on 14030 main places in the 2011 Census. For each disparity indicator the spatial pattern was computed using main place data. The results allow the classification of main places as clustered, dispersed or random, using the test with null hypothesis checking if the measures of the spatial disparities for each main place location are randomly distributed. Those main places having neighbouring places with similarly high or low measures correspond to a positive Local Moran as part of the same cluster with small p-values statistically significant. Therefore, a positive Moran's I index tends toward clustering as opposed to a negative Moran's I index point out dispersion. Table 4 shows the usual permutations – high-high as HH; high-low as HL; low-high as LH; low-low as LL; and not significant as NS.

Table 4: Results of Local Moran on disparities indicators

Disparity indicators	Clustering categories	Average Local Moran I Index	Average Local Moran I Z-Score	Average Local Moran I p-value	Number of places clustered or not
Multidimensional composite index	HH	0.4241	12.0900	0.0019	5 289
	HL	-0.4824	-5.6526	0.0082	179
	LH	-0.2756	-7.9273	0.0038	1 157
	LL	1.4573	13.9753	0.0022	2 750
	NS	0.0239	0.2239	0.5519	4 655
Mean range ratio (standardised mean ratio)	HH	0.9723	11.0344	0.0026	1 631
	HL	-0.2878	-8.5307	0.0032	722
	LH	-0.2701	-3.4574	0.0114	797
	LL	0.1788	5.5495	0.0030	4 848
	NS	0.0055	0.0617	0.4822	6 032
Relative mean deviation	HH	0.8368	19.9962	0.0046	8
	HL	-0.3426	-8.0349	0.0058	7

	LH	-0.0989	-3.4568	0.0090	33
	LL	0.3359	5.3837	0.0061	121
	NS	0.0021	0.0681	0.9189	13 861
Coefficient of variation	HH	0.4023	6.0255	0.0047	1 441
	HL	-0.3164	-9.5487	0.0032	935
	LH	-0.1997	-3.2600	0.0120	698
	LL	0.1821	5.4217	0.0042	3 282
	NS	0.0025	0.0253	0.5116	7 674
Mean logarithmic deviation	HH	0.4268	12.5241	0.0026	2 846
	HL	-0.6109	-7.6524	0.0054	438
	LH	-0.1446	-4.6716	0.0051	1 678
	LL	0.2785	3.7874	0.0068	3 206
	NS	0.0376	0.2609	0.4798	5 862
GINI coefficient	HH	0.2341	3.3557	0.0104	1 233
	HL	-0.2035	-5.9652	0.0055	1 211
	LH	-0.3264	-4.7517	0.0091	409
	LL	0.2901	7.2800	0.0040	2 251
	NS	0.0164	0.1547	0.5489	8 926
Kuznets ratio	HH	5.0715	62.6041	0.0033	215
	HL	-0.1582	-3.9999	0.0017	3
	LH	-0.1893	-2.6927	0.0099	8
	NS	-0.1893	-2.6927	0.0099	13 804
Standard deviation of logarithms	HH	0.4696	14.2629	0.0015	4 458
	HL	-0.5484	-6.5364	0.0056	678
	LH	-0.2615	-8.0276	0.0037	1 089
	LL	0.8189	9.0767	0.0028	3 279
	NS	0.0252	0.1947	0.4974	4 526
Theil inequality index	HH	0.3134	4.9001	0.0088	1 202
	HL	-0.0810	-3.5119	0.0138	250
	LH	-0.2663	-4.7357	0.0083	214
	LL	0.0940	3.7824	0.0077	2 322
	NS	0.0152	0.2923	0.5791	10 042
Atkinson index	HL	-0.2602	-2.0961	0.0366	4
	LH	-0.2602	-2.0961	0.0366	83
	LL	-0.4802	-10.1345	0.0014	111
	NS	-0.4802	-10.1345	0.0014	13 832

The classification of concentrations of high and low values allows for the rejection of the null hypothesis where the p-value is less than 0.05. The number of places that are outliers – i.e. part of the not significant count – and that have not been classified is significant. It shows the relative sensitivity of individual indices to identify clusters of places displaying disparity. Amongst these results the relative mean deviation with a count 98.8 per cent, the Kuznets ratio with 98.4 per cent and the Adkinson index with

98.8 per cent stand out. This also indicates that the computed indicators of spatial disparities are not similar.

Particular emphasis is placed on extreme measures of disparities where urgent policy intervention is required. HH clusters represent main places displaying extreme levels of spatial disparity. Different types of disparity measures can be used individually or collectively as indicators of the most appropriate policy measures for a particular or collection of neighbouring main places.

With the exception of the Atkinson index, all others measure of spatial disparities in table 7 have detected places where the spatial disparity is considered as extreme (HH). The combined Moran map of the HH cluster shows a clear pattern of the type of spatial disparities appropriate for particular main places (Figure 1). The Kuznet index (dark blue colour in Figure 1) confirmed the widening disparities between high and low earning households within the CBD and suburbs of the metropolitan municipalities of Johannesburg, Ekurhuleni, Tshwane, eThekweni, Mangaung, Cape Town and Nelson Mandela Bay. Extreme disparities are found in the surrounding places of the core metropolitan areas using the Theil inequality index (pink colour in Figure 1). Areas displaying extreme disparities are also visible around the mining and industrial areas using the coefficient of variation (orange colour in Figure 1) in North West, Mpumalanga and Free State. The Kuznet index was helpful in indicating similar areas plagued by extreme disparities in Rustenburg.

The Gini coefficient is used to identify traditional rural areas where subsistence forms of existence are predominant in North West, Free State, Eastern Cape, KwaZulu-Natal, Northern Cape and Limpopo (see the light blue colour in Figure 1). The range ratio better displays spatial disparity amongst commercial farms in the Western Cape, Free State, North West, Mpumalanga, Northern Cape and Eastern Cape (see the areas that are coloured blue in Figure 1). The extreme disparities measured using standard deviation of logarithms (purple colour in Figure 1) isolate settlements in the Eastern Cape that are trapped in poverty. Similarly, the multidimensional composite index was used to identify settlements that are lagging behind in socio-economic development in Eastern Cape and KwaZulu-Natal (see the yellow coloured areas in Figure 1).

Comparison of Spatial Disparities measure of Local Moran High-High cluster on 2011 Census Main Places

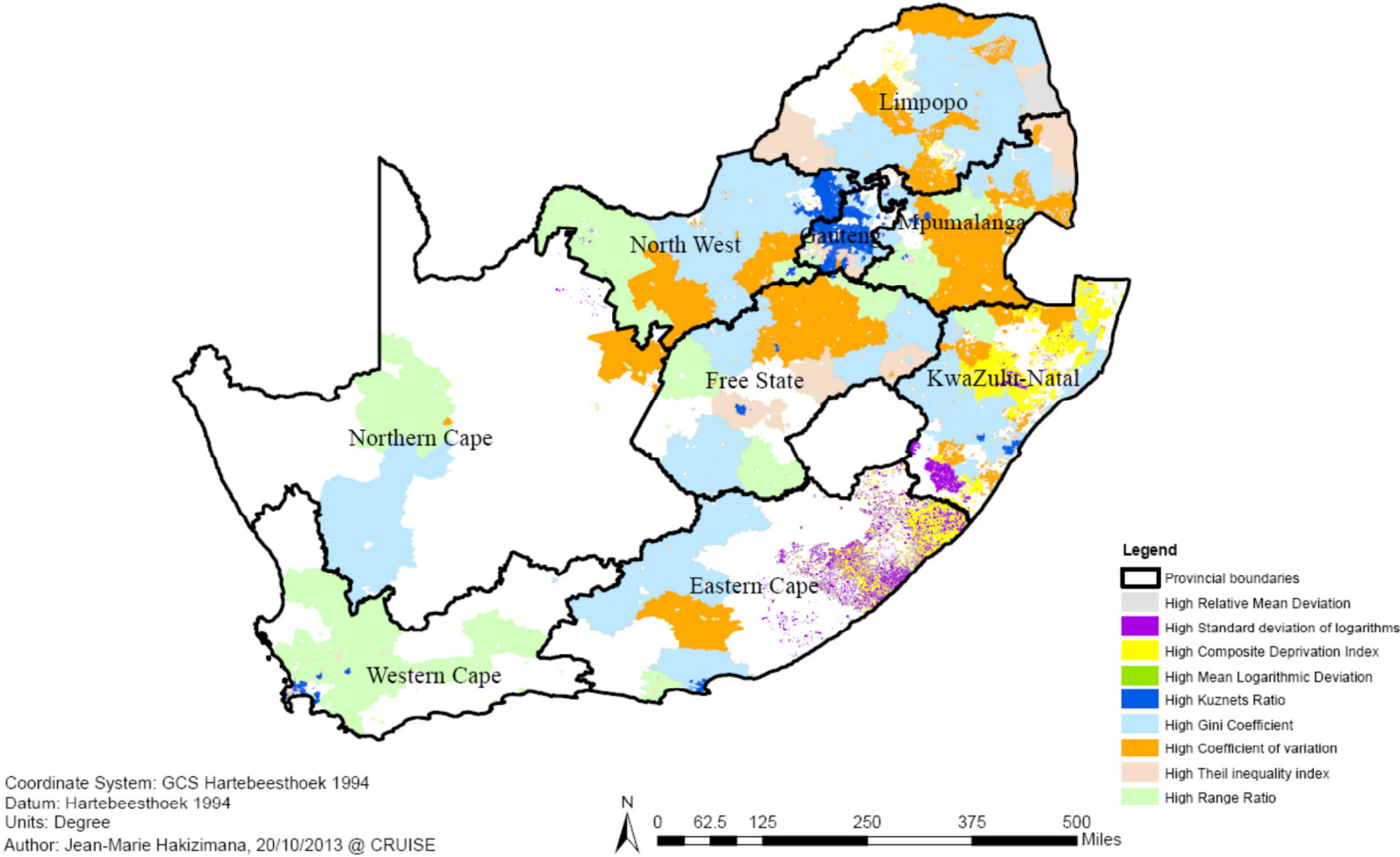


Figure 1: Comparison of spatial disparity measures of Local Moran High-High clusters on main places, 2011

## Conclusion

Income inequality in South Africa remains one of the highest in the world (WORLD BANK, 2009). Inequality is also demonstrated through lack of access to natural resources; a two-tiered educational system; a dual health system; and other socio-economic challenges. These areas of inequality are an issue of concern to policy makers and social scientists (CORNELL, 2012). To address these issues the government of South Africa has been implementing a series of pro-poor policies aimed at reducing poverty since the end of apartheid. This has culminated in the recently announced National Development Plan for 2030.

The focus in pro poor policies has all along been on the expansion of the social security system, especially vulnerable groups such as children, the disabled and the elderly. The question is whether the current width of the social security system – elements of which are leaning towards the welfare state concept – would remain sustainable. Another question is whether the provision of those services to vulnerable groups is sufficiently and equitably distributed (VAN DER BERG *et al.*, 2007).

To make a contribution in this area, this paper has attempted to quantify inequality and to identify areas that are plagued by inequalities in South Africa. A set of indices was used to measure inequality and deprivation at the main place level. However, the main findings are that there is no single measure that can be used to measure all causes of disparity adequately. Rather, the appropriate measure of spatial disparities is different and specific to particular places. Therefore, the kinds of spatial disparities that are associated with specific groups of places make it necessary to develop specific policy interventions.

The measured spatial disparities seem to follow the patterns of differential urbanisation separating the core cities and the fringes (GEYER *et al.*, 2012). For instance the Kuznet index has shown that some metropolitan areas are plagued by extreme inequality whereas the Theil index showed inequalities in peripheral areas of metropolitan areas, especially historically black residential areas. More broadly, the range ratio is effective in identify disparity within farming areas, while the coefficient of variation detected disparities in mining areas. Although it is a generally useful tool to measure disparity, the Gini coefficient is particularly suitable for the identification of inequality in tribal areas. Strong similarities have been found in the results from the composite index, the standard deviation of logarithms and the mean logarithmic deviation. This study has identified clusters of areas where the source of disparities are likely to be similar and thereby could contribute to the development of more situation-specific pro poor policies.

## References

- ALIKE S., ROCHE J. M., SANTOS M. E. And SETH S. (2011) Multidimensional poverty index 2011: Brief methodological note, Oxford Poverty & Human Development Initiative (OPHI), Oxford University.
- ANSELIN L. (1995) Local indicators of spatial association-LISA, *Geographical Analysis* **27**(2), 93-115.
- BARNES H., WRIGHT G., NOBLE M. and DAWES A. (2007) The South African index of multiple deprivation for children. Census 2001. HSRC Press: Cape Town.
- CHAKRAVORTY S. (1996) A measurement of spatial disparity: the case of income inequality, *Urban Studies* **33**(9), 1671-1686.
- Cornell V. (2012) Report on an international Symposium: The politics of poverty research and pro-poor policy making: Learning from the practice of policy dialogue. PLAAS, University of the Western Cape.
- D'SA E. R. (1986) Social indicators and spatial disparities in Papua New Guinea. *Social Indicators Research* **18**, 285-319
- EUROPEAN UNION (2007) Regional disparities and cohesion: What strategies for the future. European Parliament, Directorate-General for Internal Policies of the Union. IP/B/REGI/IC/2006\_201 14/05/2007. EU: Brussels.
- GEYER Jr.H. S., Geyer H. S., du Plessis D. J. and van Eeden A. (2012) Differential urbanisation trends in South Africa—regional and local equivalents *Environment and Planning A* **44**(12), 2940–2956
- GILTHORPE M. S. (1995) The importance of normalization in the construction of deprivation indices. *Journal of Epidemiology and Community Health* **49**(2), S45-S50.
- GOSCHIN Z., CONSTANTIN D-L., ROMAN M, ILEANU B (2008) The current state and dynamics of regional disparities in Romania, Academy of Economic Studies of Bucharest.
- HERRMANN-PILLATH C., KIRCHERT D., PAN J. (2002) Disparities in Chinese economic development: approaches on different levels of aggregation, *Economic Systems* **26**, 31–54.
- KANBUR R. and VENABLES A. J. (2005) Spatial inequality and development. Overview of the UNU-WIDER Project. September. [www.arts.cornell.edu/poverty/kanbur/WIDERProjectOverview.pdf](http://www.arts.cornell.edu/poverty/kanbur/WIDERProjectOverview.pdf).
- KLEINMAN M. (1999) There goes the neighbourhood: area policies and social exclusion, *New Economy*, **6**, 188-192
- KRIAA M., DRISS S. and KARRA Z. (2011): Multidimensional inequality and spatial disparities in Tunisia, Conference Paper, Politics and Economic Development, March 2011, Antalya, Turkey.

- KRISHNAN V. (2010): Constructing an area-based socioeconomic index: A principal components analysis approach, Early child development mapping project (ECMap), Community-University Partnership (CUP), Faculty of Extension, University of Alberta, Edmonton, Alberta T5J 4P6, CANADA.
- KUTSCHERAUER A., FACHINELLI H., HUČKA M., SKOKAN K., SUCHÁČEK J., TOMÁNEK P. and P. TULEJA (2010): Regional disparities: Disparities in country regional development - concept, theory, identification and assessment, WD-55-07-1.  
[http://disparity.idealnihosting.cz/edice\\_cd/cd11\\_regdis\\_mono\\_angl/pdf/Regional%20disparities.pdf](http://disparity.idealnihosting.cz/edice_cd/cd11_regdis_mono_angl/pdf/Regional%20disparities.pdf).
- LENGYEL, I. (2009) Bottom-up regional economic development: competition, competitiveness and clusters, in BAJMÓCY, Z. LENGYEL, I. (eds) Regional competitiveness, innovation and environment, JATE Press, Szeged, pp. 39-53.
- LI S., and XU Z. (2008) The trend of regional income disparity in The People's Republic of China, ADB Institute Discussion Paper No. 85, January, P.R.China, 310018.
- LITCHFIELD J.A. (1999) Inequality: methods and tools. Inequality, Poverty and Socio-economic Performance, World Bank. <http://www.worldbank.org/poverty/inequal/index.htm>.
- MOHIUDDIN S., HASHIA H. (2012) Regional socio-economic disparities in the Kashmir Valley (India) –A Geographical Approach, Bulletin of Geography. Socio–Economic Series No. 18, 85–98.
- MOLLE W. (2007) European cohesion policy. Routledge: London.
- MONFORT P. (2008) Working papers, A series of short papers on regional research and indicators, Convergence of EU regions, Measures and evolution, Directorate-General for Regional Policy, European Commission, Regional Policy, 01/2008.
- MORAN P. A. P. (1948) The interpretation of statistical maps, *Journal of the Royal Statistical Society*, B, **10**(2): 243-251.
- NOBLE M., DIBBEN C. and WRIGHT G. (2010) : The South African index of multiple deprivation 2007 at Data zone Level (modelled), Department of Social Development, Pretoria.
- NPC National Planning Commission (2011) National development plan, Vision for 2030. Ministry of the Presidency.  
<http://www.npconline.co.za/medialib/downloads/home/NPC%20National%20Development%20Plan%20Vision%202030%20-lo-res.pdf>.
- OECD (2002) Geographic concentration and territorial disparity in OECD countries. Paris: OECD Publications Service.
- OECD. (2003) Geographic concentration and territorial disparity in OECD countries. Paris: OECD Publications Service.



- PAUW K. (2003): Technical paper 2003: 4, Measures of poverty and inequality: A reference paper, PROVIDE Project. Elsenburg, South Africa.
- PENG J., GROENEWOLD N., HE J., LI Z. and YI Y. (2006) Regional finance and regional disparities in China, Hunan University, DISCUSSION PAPER 08.02.
- REY S. J. and JANIKAS M. V. (2005) Regional convergence, inequality, and space, *Journal of Economic Geography* **5**, 155-176.
- SAHN D. and STIFEL D. (2003) Urban-rural inequality in living standards in Africa. *Journal of African Economies* **12**(1), 564-597.
- SEN A. (1997) *On income inequality*, Clarendon Press, Oxford.
- SMITH, G.A.N., NOBLE M. and WRIGHT G. (2001) Do we care about area effects? *Environment and Planning A* **33**, 1341-1344.
- SMITH, G. R. (1999) Area-based initiatives: the rationale and options for area targeting, CASE Paper 25.
- STATISTICS SOUTH AFRICA (2013) : Census 2011 Household questionnaire A, November 2010
- STILLWELL J, NORMAN P., THOMAS C. and SURRIDGE P. (2010) *Spatial and social disparities: Understanding population trends and processes – volume 2*, Springer, Dordrecht.
- TOWNSEND, P. (1987) Deprivation, *Journal of Social Policy*, 16 (1), 125-146.
- TRIEGAARDT J. D. (2007) Poverty and inequality in South Africa: Policy considerations in an emerging democracy, Policy Analysis, ADRU.  
<http://www.dbsa.org/Research/Documents/Poverty%20and%20inequality%20in%20South%20Africa%20Policy%20considerations%20in%20an%20emerging%20democracy.pdf?AspxAutoDetectCookieSupport=1>
- UNDP (2013) Human Development Report. UNDP: Washington.
- UNITED NATIONS (2005) World population policies. United Nations, Washington.
- UNITED NATIONS (2006) World population prospects: The Revision, vol. I.  
<http://www.un.org/esa/population/publications/wpp2006/wpp2006.htm>
- VAN DER BERG S., BURGER R., BURGER R., LOUW M. and YU D. (2007) A series of national accounts-consistent estimates of poverty and inequality in South Africa, Department of Economics, Stellenbosch University, Economic Working Paper: 09/07.
- VENKATANARAYANA M. (2004): *Educational deprivation of children in Andhra Pradesh, Levels and trends, disparities and associative factors*, Working Paper 362.
- VITURKA, M. (2006) Innovative potential regional evaluation in the Czech Republic, Proceedings of the international conference on regional and urban modeling, p. 1-12.

- VITURKA, M. (2010) Regionální disparity a jejich hodnocení v kontextu regionální politiky, *Geografie*, **115**(2), 131–143.
- WILKINSON R. and PICKET K. (2009) *The spirit level: Why more equal societies almost always do better.* Allen Lane, London.
- WISHLADE F. and YUILL D. (1997) Measuring disparities for area designation purposes: Issues for the European Union, Regional and industrial policy research paper, No. 24 June.
- WORLD BANK (2009) World development report. Reshaping economic geography. World Bank, Washington.
- XIAOLI H. (2010) Comparisons of income distribution disparity measure methods, School of Economics and Trade, Hangzhou Dianzi University.
- ZAREDA H. (2013) South Africa's national development plan and its implications for regional development, TRALAC Trade Law Centre. <http://www.tralac.org/files/2013/07/D13WP012013-Zarenda-South-Africas-NDP-and-implications-for-regional-development-20130612-fin.pdf>.