Economic and demographic trends of municipalities in South Africa: An application of Zipf's rule

Hlabi Morudu & Danie du Plessis

Abstract

There are vast differences among South African local municipalities, with a limited number of large municipalities (both in terms of population size and economic activity) and a seemingly disproportionate number of intermediate-sized and small municipalities. No clear systematic national approach has yet been adopted to assess the distribution of core variables at municipal level in South Africa. Zipf's rule, which postulates a consistent regularity in the size and rank of cities, is applied to disaggregate the performance of South African local municipalities in terms of three variables (population, Gross Value Added and municipal income) within the overall national settlement pattern. The results indicate that the Zipf rank size rule distribution is applicable to municipal level population data in South Africa, but less so for Gross Value Added and municipal income. The position and relative changes of municipalities along the Zipf curve between 2001 and 2011 also provide plausible indications of potential future trajectories of the three variables classified according to the dominant settlement typology within each municipality. The results also emphasise the significant conceptual limitations when using only legally defined administrative municipal boundaries for analysis purposes without also considering economically functional boundaries.

EKONOMIESE EN DEMOGRAFIESE TENDENSE VAN MUNISIPALITEITE IN SUID-AFRIKA: 'N TOEPASSING VAN ZIPF SE REËL

Daar is wesenlike verskille tussen Suid-Afrikaanse munisipaliteite wat gekenmerk word deur 'n beperkte aantal groot munisipaliteite (in terme van bevolking en ekonomiese aktiwiteite) en 'n skynbaar buite verhouding groot aantal middelgroot en klein munisipaliteite. Op nasionale vlak is daar tans geen sistematiese benadering om die verspreiding van sleutelveranderlikes op munisipale vlak te beoordeel nie. Zipf se reël veronderstel 'n konsekwente reëlmatigheid in die grootte en rangorde van stede en word toegepas om die tendense van Suid-Afrikaanse munisipaliteite in terme van bevolking, ekonomiese produksie en munisipale inkomste binne die nasionale vestigingspatroon te beoordeel. Die resultate toon dat Zipf se rangorde-grootte reël van toepassing is op die verspreiding van bevolking op munisipale vlak, maar nie tot dieselfde mate vir ekonomiese produksie en munisipale inkomste nie. Die posisie en relatiewe veranderings van munisipaliteite op die Zipf-kurwe tussen 2001 en 2011 kan aanduidings verskaf van moontlike toekomstige neigings van die drie veranderlikes op munisipale vlak. Die resultate beklemtoon die moontlik beperkende invloed van die gebruik van administratiewe grense sonder inagneming van ekonomies funksionele grense.

TSELA EE THUTO EA MERUO LE PALO EA SECHABA E TSAMAEANG KA HARA BO MASEPALA BA AFRIKA BORWA: TSHEBELISO EA MOLAO OA ZIPF

Ho na le phapang tse kholo hara bo masepala a Afrika Borwa (ka moruo le palo ea batho). Ho bonahala ebile bo masepala ba sa lekane ka matla. Ha ho na tshebeliso e tloaelehileng e sebelisong ka mehla ka hara masepala naheng ea Afrika Borwa. Molao oa Zipf o bontsha hore tshebeliso ena e tlameha hjo ba teng ka hara masepala hore masepala u tle u khone hore u atlehe dintlheng tsena tse tharo e leng (palo ea batho, moruo le mokholo oa masepala). Ditlamorao tsa molao ona di bontsha molao ona u ka sebedisoa masepaleng o mong le o mong naheng ea Afrika Borwa. Diphetoho tsa bo masepala ho molao oa Zipf mahareng a selemo sa 2001 ho isa ho 2011, di bontsha molao ona u fana ka bokamoso bo phethahetseng oa bo masepala. Dipheto di bontsa hore molao oa Zipf u thusa le ho feta ha ho sebetsoa ho sheboe nthla e le ngoe fela ho sa chejoe tse ding.

D.H. Morudu, Executive Manager, Statistics South Africa, 170 Andries Street, Pretoria, 0002, South Africa. Phone: +27 12 3108911, email: <HlabiM@statssa.gov.za>

D.J. du Plessis, Centre for Regional and Urban Innovation and Statistical Exploration (CRUISE), Department of Geography and Environmental Studies, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa. Phone: +27 21 8083126, email: <ddp@sun.ac.za>

1. INTRODUCTION

The development planning policy framework and associated implementing tools in South Africa have undergone fundamental changes since the onset of the democratic era in 1994. At a national level, the main strands of these changes commenced with the Reconstruction and Development Programme (RDP) in 1994 that focused on investment in infrastructure and basic services, soon followed by the promulgation of the Development Facilitation Act (DFA) in 1995, representing the first step towards a new planning framework. The period from 1996 saw a change in focus to a competitive and rapidly growing economy and included the introduction of policies such as the Growth, Employment and Redistribution (GEAR) programme (DF, 1996), and later the announcement of the accelerated and shared growth initiative for South Africa (ASGISA) (The Presidency, 2006). The most prominent initiative in the period after 2009 has been the establishment of the National Planning Commission in 2010 as an advisory body tasked with preparing recommendations to the cabinet on issues affecting South Africa's long-term development and resulting in the subsequent National Development Plan 2030 (NDP) in 2012. At municipal level, the planning process has been governed by the introduction of a system of Integrated Development Planning (IDP) supported by a number of sectoral plans such as the Spatial Development Frameworks (SDF). It is, however, not always clear how nationally aggregated data used in national growth programmes such as the NDP, ASGISA and GEAR actually articulate themselves at the local municipality level and how these programmes link to municipal Integrated Development Plans. There are vast differences between South African local municipalities, with a limited number of large municipalities (both in terms of population size and economic activity) and a seemingly disproportionate number of intermediate-sized and small municipalities. No clear systematic national approach has yet been adopted to assess the distribution of core variables such as population size, Gross Value Added (GVA) and municipal income at municipal level in South Africa.

Rank-size distribution refers to the regularity in distribution of phenomena such as city sizes within a specified country or region. The distribution of city population in a country or a region is characterised by a few large cities and a much larger number of cities of a smaller magnitude. Zipf's rule posits a consistent regularity in the size and rank of cities. That is, the population size of any city ranked x is equal to (a/x) of the largest city with a = 1. Several empirical studies have applied the Zipf rule to population size distributions in a number of countries (Rosen & Resnick, 1980: 184; Gabaix, 1999: 740; Gabaix & Ioannides, 2004; Soo, 2005: 246; Fragkias & Seto, 2009: 189). Unlike earlier applications of Zipf's rule in South Africa which focused only on the distribution of the population among South African cities, this study also introduced GVA and municipal income as variables (Naude & Krugell, 2003: 178; Krugell, 2005: 82; Soo, 2005: 246). The research also departs from the narrow focus on cities in prior studies to include more extended observational units such as municipalities and potential broader economic functional regions as suggested by Berry & Okulicz-Kozaryn (2012: S21). The significance of this broadened approach is threefold: it allows the disaggregation of nationally aggregated estimates of identified variables to better comprehend their distribution among local municipalities; it illustrates the potential for disaggregating national variables to more rigorously assess the scope of economic performance at local municipal levels, and it illustrates how variables may have varying effects on different municipalities, and hence guide local policy formulation.

In this article, Zipf's rule is applied to disaggregate the performance of South African local municipalities within the overall national urban system. The objective of this article is twofold. Firstly, to investigate the hierarchical structure of the settlement pattern in South Africa at municipal level and to determine whether the Zipf rank-size rule is applicable to the distribution of population and economic variables at municipal level and, secondly, to explore the hierarchical distribution of the variables across municipalities according to the dominant settlement typology (National Urban Development Framework classification) within each municipality. The value of this type of empirical analysis is that it provides

insights into the spread of national aggregates at municipal level. It also provides the bases for defining plausible future trajectories of the population and economic activity at municipal level, disaggregated according to the dominant settlement typology within each municipality.

2. A REVIEW OF THE APPLICATION OF ZIPF'S RULE

The seemingly skewed spread of cities in terms of population size and economies of scale is a global phenomenon and has historically attracted significant research. The rationale for the skewness, drawn from the classic urban economic studies of Lösch (1954), Philbrick (1957), Berry (1964), Christaller (1966), and Von Thünen (1966) on the distribution and hierarchy of cities, postulates that a large number of small towns perform central place functions in their respective locations and fit into a system of higher order cities that progressively decrease in numbers the larger the scale of economies and population sizes. Emanating from the earlier empirical work by Auerbach (1913) and Gibrat (1931), Zipf's rule asserts that the population size of a city i is inversely proportional to its rank. That is, the population size of any city ranked x is equal to (a/x) of the largest city with a = 1. If Zipf's rule does not hold (either larger or smaller than 1), it provides an indication of the evenness of the distribution of the population among cities compared to the Zipf predictions. In his hypothesis, Hsu (2008: 1) argues that the distribution of cities by rank in central place theory is indeed consistent with Zipf's rule. A parallel development of Zipf's rule is Pareto's distribution function. Unlike Zipf's rule and its condition of a = 1, Pareto's distribution is based on a cumulative distribution function and presents the rank of a city i as an inverse power of the city's population, i.e. $R_i = CS_i^{-1/a}$ (Newman, 2006: 1). Nitsch (2005: 93) reports no significant difference in the results of the two equations, viz. a = 1.096, when population size is the dependent variable in the study and a = 1.082, when rank is the dependent variable.

A substantial body of research has been produced to empirically validate Zipf's rule that a = 1 and different findings have been reported. Using population data from 44 countries, Rosen & Resnick (1980:166) presented estimates ranging

from 0.81 to 1.96. Soo (2005: 239) used data from 73 countries and found that Zipf's rule was rejected in 53 out of 73 countries using ordinary least squares. Even with the use of the more accommodating maximum likelihood Hill estimator, based on maximum likelihood regressions on changing data categories in the data set, Zipf's rule was still rejected in 30 out of 73 countries (Soo, 2005: 239). Nitsch (2005: 90) reviewed 515 estimates from 29 studies and found that as much as two thirds of the estimates lay between 0.8 and 1.2. Berry & Okulicz-Kozaryn (2012: S17) argue that deviations from Zipf's rule are a result of using different urban-regional units. Mainly due to the lack of adequately defined official functional economic urban-regional units, empirical studies on Zipf's rule have resorted to the use of legal administratively defined urban-regional units. Using redefined US urban regional units along economic functional units, Berry & Okulicz-Kozaryn (2012: S17) established that Zipf's rule holds, i.e. a = 1. This confirms the findings of Nitsch (2005: 93) who observed that the estimated a became smaller when the urban-regional unit was redefined from a region (a = 1.29) to a city (a = 1.11) to an agglomeration (a = 1.02).

Another important strand of application of Zipf's rule is to understand the changes to the hierarchical distribution of cities and the factors that affect their development (Gao & Wu, 2008: 148; Fragkias & Seto, 2009: 197). The approach provides an avenue for exploring the effects over time of industrialisation policies or changes in city form. Nitsch's (2005: 93) meta-analysis revealed a clear decline in a estimates over extended periods, from an average of 1.35 prior to 1801 to 1.07 since 1950. Estimates made in South African studies vary from 1.36 using ordinary least squares and 1.27 using the Hill estimator based on 1991 population census data (Soo, 2005: 246), to 0.75 by Krugell (2005: 82) and Naude & Krugell (2003: 179) with the 2001-census data.

The power laws, depicted in Zipf's rule and the Pareto distribution, are understood to be general and also apply to phenomena other than population size. Li (2002: 16-17) noted studies that use Zipf's rule in word usage in languages; population sizes of ranked cities; web page visits and traffic; company sizes; science citations, and scaling in natural physical phenomena

Table 1: Summarised findings of studies on the rank-size rule

Source	Variable	Countries	a estimates
Rosen & Resnick (1980)	Population	44 countries	0.81-1.96
Soo (2005)	Population	73 countries	0.73-1.72
Nitsch (2005)	Population	515 estimates	0.80-1.20
Krugell (2005)	Population	South Africa	0.75
Naude & Krugell (2003)	Population	South Africa	0.75
Fujiwara, Di Guilmi, Aoyama, Gallegati & Souma (2003)	Company assets	France	0.88-0.89
	Company sales	France	0.89-0.91
	Company employees	France	0.98-1.01
Tanaka & Hatsukano (2011)	Employees	Cambodia	1.33
Rossi-Hansberg & Wright (2007)	Employees	USA	$\simeq 1$
Knudsen (2001)	Population	Denmark	$\simeq 1$
	Employees	Denmark	$\simeq 1$
Okuyama, Takayasu, M. & Takayasu, H. (1998)	Company income in:		
	Construction	Japan	1.13
	Electrical products	Japan	0.72
Okuyama, Takayasu, M. & Takayasu, H. (1998)	Employees in construction & electrical products	Japan	1.2-0.7
Hinloopen & Van Marrewijk (2006)	Ballasa index:		
	1970-1997	166 countries	0.85-1.03
	Per sector	166 countries	0.39-3.42
	Per country	166 countries	0.37-3.71

such as earthquakes. Kawamura & Hatano (2002: 2) review abstracts from real-life phenomena and used random numbers to explore the validity of Zipf's rule.

A number of empirical studies have explored variables other than city population sizes. Fujiwara *et al.* (2003: 197) tested Zipf's rule using data from a list of firms in European countries, by focusing on the number of employees, sales and total assets of each company and found that the results approximate the rule. For instance, in France, the estimates were 0.88 to 0.89 for total assets, 0.89 to 0.91 for sales, and 0.98 to 1.01 for employees. Tanaka & Hatsukano (2011: 2133) examined Zipf's rule, by concentrating on employment in Cambodian businesses. They found the value of a = 1.33, mainly because Cambodian industry is characterised by a disproportionately large number of small establishments compared to developed countries and as envisaged by Zipf's rule. Rossi-Hansberg & Wright (2007: 1657) used United States data and found the size distribution of enterprises to be very close to Zipf's

rule among enterprises with 50 to 10.000 employees. Knudsen (2001: 142) established evidence of Zipf's rule in Danish population data for cities and for firm size by employment. Okuyama, Takayasu, M. & Takayasu, H. (1998: 128) focused on the income size of companies in Japan and tested Zipf's rule for construction companies (a = 1.13) and electrical product companies (a = 0.72). They also appraised Zipf's rule using company employee data and yielded estimates between 1.2 and 0.7 for construction, electrical products and power companies. Hinloopen & Van Marrewijk (2006: 1) studied bilateral trade flows for 747 sectors in 166 countries from 1970 to 1997, and found Zipf's rule to generally hold true whenever the Balassa index revealed comparative advantage.

3. APPLICATION OF ZIPF'S RULE TO SOUTH AFRICAN MUNICIPALITIES

3.1 Methodology

The analysis aims to assess whether the differences observed among South African municipalities with respect to

population size, Gross Value Added (GVA) and municipal income are consistent with Zipf's rule and research findings on the application of Zipf's rule in other countries. Equation [1] captures the Zipf's rule as expressed in related research:

$$S_i = CR_i^{-a} \qquad [1]$$

where S_i denotes the population, GVA or income size of municipality *i*, C is a constant term, R_i is the rank of municipality *i*, and a is an exponential coefficient. In logarithmic form, [1] becomes:

$$\log(S_i) = C - a\log(R_i) + \varepsilon_i \qquad [2]$$

where ε_i is an independent random error term for municipality *i*. The unknown coefficient a in [2] is estimated through ordinary least squares. Should the resulting a estimate equal 1, Zipf's rule holds.

The literature provides Zipf distributions that are either produced through an iterative process of truncating the least essential settlements that fall below the Zipf estimate at the bottom tail of a Zipf curve, or through extending the unit of observation from individual cities or municipalities to include conglomerations of municipalities as economic functional regions (Perline, 2005: 68; Berry & Okulicz-Kozaryn, 2012: S21). In each instance, the analysis consisted of two options. Option 1 involved the use of all municipalities in terms of administratively defined boundaries only. Option 2 adopted the approach of considering an extended economic functional unit at the top end of the hierarchy, using a spatial entity referred to as the Greater Johannesburg Tshwane Functional Region (GJTFR), a conglomeration consisting of the adjoining municipalities of the City of Johannesburg, Tshwane, Ekurhuleni, Mogale City, Merafong City, Randfontein and Westonaria.

The study uses the notion of a Zipf estimation range bounded by the standard error of regression as defined in basic ordinary least squares estimation. All values within the Zipf estimation range are regarded as conforming to Zipf's prediction. The population analysis is based on municipal population statistics from the 2001 and 2011 population censuses. In the absence of official surveys to collect economic data at municipal level, Quantec provides estimates of GVA at municipal level. These estimates, though not based on standard statistical surveys or administrative data

but on modelled simulations, were required for comparative economic analysis and the Quantec estimates for 2001and 2011 were used. Concerning income at municipal level, Statistics South Africa has, since 2002, annually conducted a Financial Census of Municipalities. This census is essentially a compilation of administrative data from municipalities and is published with a one-year lag. Municipal income, in this instance, refers to income generated by local municipalities derived from rates and general service charges usually imposed as taxes or levies to cover services such as fire control or health-care facilities and from housing and trading services paid for directly by households on demand. The income also includes licences; permits; interest; dividends; parking fees; rent on buildings and equipment; sale of goods and services; fines; income from the disposal of property, plant and equipment; public contributions and donations, and recovered bad debts. Subsidies and grants from national and provincial governments, and other similar budgetary transfers, are excluded from the income variable to isolate the income-generation capabilities of each municipality. The analysis used the Financial Census of Municipalities data for 2006 (when the series appeared to have stabilised with a consistent set of variables) and 2011.

A further element of the analysis process was to disaggregate the hierarchical distribution of the variables across municipalities in terms of dominant settlement typologies (as per the suggested NUDF categories) within municipalities and to analyse the spatial patterns of the hierarchical distribution of the variables. Local municipalities were categorised based on the dominant urban settlement type in each local municipality according to the categories and nomenclature proposed in the NUDF (DCGTA, 2009: 31). Each municipality was classified by the highest order or dominant form of settlement or settlements in the municipality. For example, if a municipality contains a city region, the entire municipality is considered a city region municipality. Similarly, if a municipality contains more than one town classified in different categories such as regional service centre, niche town and dense rural centres, the municipality would be classified as "regional service centre municipality", as the highest order of settlement present within the municipality. In

municipalities with a fairly equal distribution of lower order settlements (e.g. niche towns and dense rural settlements), the classification is based on the settlement type covering the largest land area.

3.2 Zipf rule applied to municipal population data

Tables 2, 3 and 4 present a summary of the application of Zipf's rule. These results reveal two overall findings that are applicable to all three variables analysed (population figures, municipal GVA, and municipal income). First, the coefficient a has significant t-statistic values and high values of R^2 adjusted for the data sets analysed. Secondly, the a values tend to be closer to 1 when extending administrative municipal boundaries in Gauteng to include conglomerations of municipalities as a single economic functional region (Option 2) compared to the alternative analysis based on the individual administratively defined municipal entities only (Option 1).

A comparison of the 2001and 2011 population data shows that Zipf's

2011 municipal population figures, respectively. The overall patterns for 2001 and 2011 are very similar and reveal a number of very distinct characteristics. In 2001, the second- and third-ranked municipalities (eThekwini and Cape Town) are slightly below the Zipf estimate, but remain within the standard error of regression (0.41). However, by 2011, eThekwini and Cape Town swopped positions at the top end of the hierarchy underneath the GJTFR, both within the standard error of regression of 0.39. The population figures of the next seven highest ranked municipalities are below the expected Zipf standard error of regression for both 2001 and 2011. These include two municipalities classified as city region municipalities (Nelson Mandela Bay ranked 4th in 2001 and 2011, and Emfuleni ranked 6th in 2001 and 7th in 2011); four classified as city municipalities (Buffalo City ranked 5th in 2001 and 2011, Mangaung ranked 7th in 2001, and 6th in 2011, Msunduzi ranked 9th in 2001 and 2011, and Polokwane ranked 10th in 2001 and 8th in 2011), and a regional service centre municipality (Thulamela ranked 8th in 2001, and

Table 2:   Results of Zipf's rule based on municipal population data

Data	Option 1 (All municipalities) (Equation [2])		
	a	t-value	R^2 adjusted
Population 2001	1.06	38.46	0.86
Population 2011	1.07	40.16	0.87
GJTFR	Option 2 (GJTFR) (Equation [2])		
	a	t-value	R^2 adjusted
Population 2001	1.04	37.03	0.86
Population 2011	1.04	38.25	0.87

rule does apply to South African municipalities and reflects stability in a coefficient of about 1.07 in the case of administratively defined municipalities, and 1.04 in the case of GJTFR (Option 2). The slight increase between 2001 and 2011 in Option1 may be indicative of the delayed effects of normalisation of population movements following the removal of race-based restrictive policies with relative increases in the movement of the population towards the larger cities forming part of the upper end of the hierarchy. However, when applying Option 2 using the GJTFR at the top end of the hierarchy, the value remained constant between 2001 and 2011.

Figure 1 depicts the results of the Zipf rule applied to the 2001and

10th in 2011).

The middle section of the Zipf curve with actual population sizes fitting within the standard error of regression represents the majority of the municipalities (198 in 2001and 195 in 2011). At the lower tail of the Zipf curve, population figures are both below the estimated Zipf sizes and below the standard error of regression. In 2011, this category included the dense rural municipality of Richtersveld and 22 niche town municipalities.

Figure 2 provides a spatial representation of this information and depicts the geographical spread of municipalities according to three categories: municipalities with population figures higher than the Zipf estimate, but within the standard error of regression range (Bin 2);
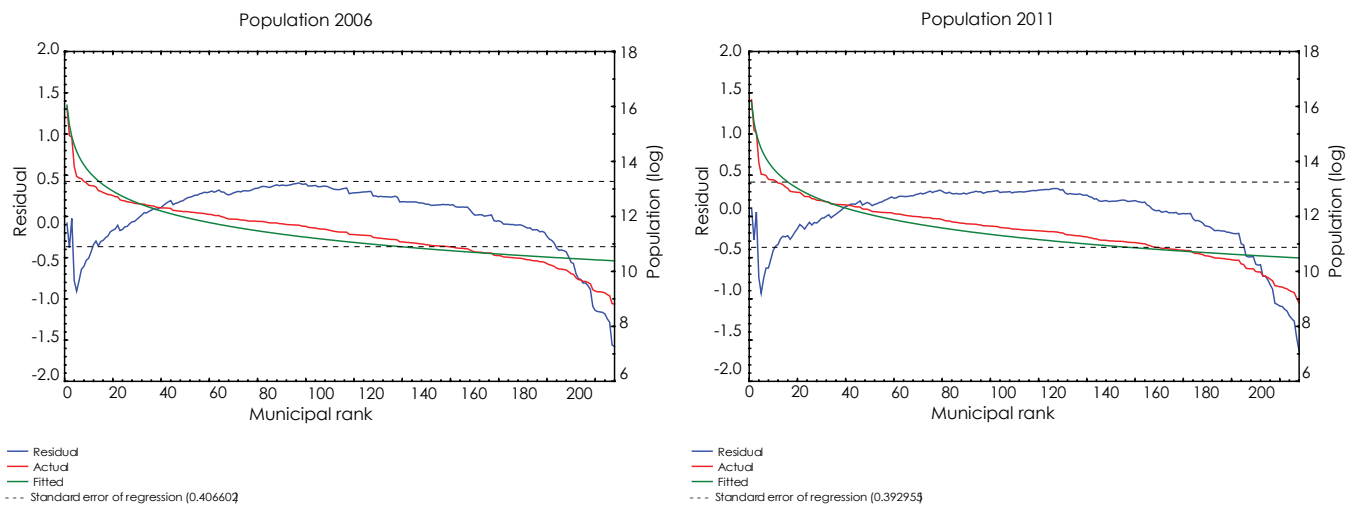
Figure 1: Actual, fitted and residual Zipf curves based on 2001 and 2011 population (Option 2)

municipalities with population figures lower than the Zipf estimate, but within the standard error of regression range (Bin 3), and municipalities with population figures lower than the Zipf estimate and outside the standard error of regression range (Bin 4). A number of characteristics are evident from this spatial analysis. Municipalities in Bin 2 are mainly concentrated in three areas. First, in most municipalities of the Eastern Cape, Kwazulu-Natal, North-West and Limpopo classified as dense rural

ideologies relocating large segments of the South African population to these areas along ethnic lines and limiting their ability to respond to normal push-and-pull factors that may have influenced their migration to larger urban centres. The second concentration in the north-eastern parts of the Free State, large parts of the Mpumalanga, and the southern parts of Limpopo represents areas mostly classified as regional service centres. The third concentration is located

linkages with GJTFR and, in the case of the latter, to the City of Cape Town.

Municipalities in Bin 3 are fairly randomly distributed across the spatial economic landscape and mostly consist of niche town municipalities. The Bin 4 municipalities consist of two distinct groups. The first group consists of city region and city municipalities following the three highest rank municipalities and include areas such as Nelson Mandela Bay, Buffalo City, Mangaung, Msunduzi and Polokwane. The second group mainly consists of niche town municipalities located in the Northern Cape and a limited number of these types of municipalities scattered across the other provinces. The position of these two groups of municipalities within the overall Zipf-based distribution is likely to imply distinct implications for possible future population trends. The first group is likely to experience increased population growth and a tendency to move closer to the Zipf-based population postulations for these areas, as these intermediate cities mature and their growth begins to exceed that of the primate city region. Conversely, the second group is likely to experience population decreases and a tendency to remain at the bottom part of the tail of the Zipf curve.

Figures 3a to 3f present a further analysis of the rank spread of municipalities according to settlement type classification and reveal a number of important characteristics. The city region and city municipalities are almost exclusively located at the top end of the Zipf curve, with population figures generally well below the estimates based on the Zipf rank size rule for both 2001 and 2011. In terms of the overall municipal hierarchy and their position
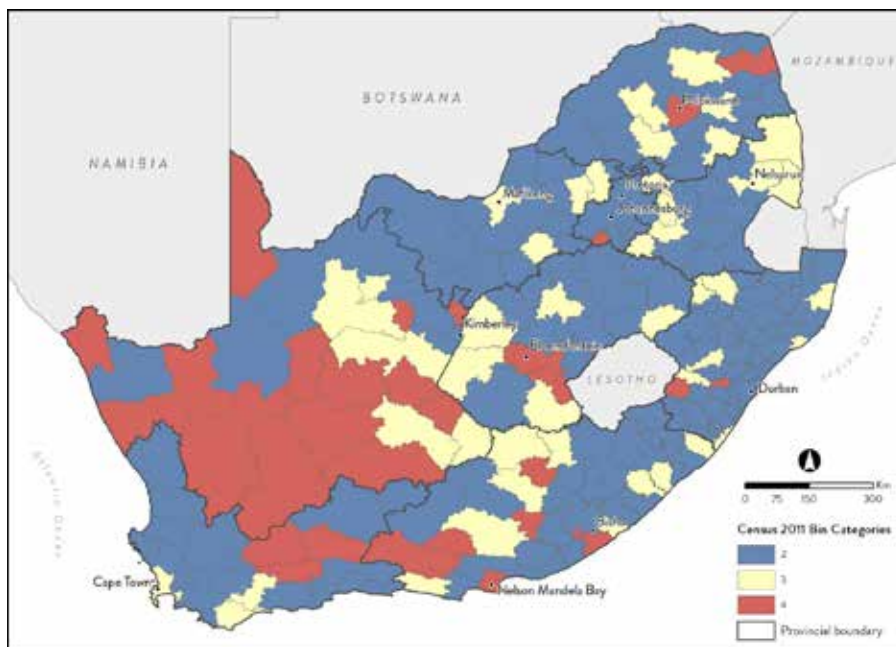

Figure 2: Spatial spread of municipalities according to 2011 population figures

municipalities and generally located within ex-Bantustan areas.

The larger than expected population in these municipalities (expressed in terms of Zipf's rule) can be ascribed to apartheid-based spatial development

in a band around the Cape Town Metropolitan municipality, consisting of regional service centre and service town municipalities. The former concentration can partially be ascribed to strong physical and functional

relative to the Zipf curve, it would thus be plausible to expect that the majority of these municipalities are likely to move closer to Zipf estimates in the future and experience increased rates of population growth and in-migration. This is consistent with the theory of differential urbanisation, postulating that the urban system enters the polarisation reversal phase when the urban system has reached the level of primacy where the populations of the largest urban agglomerations in the system might still be growing overall, but forces of agglomeration are now giving way to forces of dispersion (Geyer, Du Plessis, Geyer & Van Eeden, 2012: 66). The regional service centre municipalities represent a broad spread across the upper and middle parts of the Zipf curve (see Figures 3a and 3c). Over the period between 2001 and 2011, the municipalities on the upper part of the curve fell further below the Zipf estimates and those at the bottom part showed close approximation with

the Zipf estimates. Regional service centre municipalities in the upper part of the curve (e.g. Thulamela, Rustenburg, Bushbuckridge, Madibeng) are likely to experience above average population growth rates in the short to medium term as population growth and migration gains momentum relative to the largest urban agglomerations, whereas the population of those at the bottom part of the curve is likely to remain relatively constant.

The service town and the dense rural municipalities are both concentrated in the middle and lower parts of the curve. The distribution of the dense rural municipalities relative to the Zipf estimates changed substantially between 2001 and 2011 (see Figure 3e). In 2001, the population sizes of the majority of these municipalities (41 out of a total of 51) were above the Zipf estimates. However, by 2001, only 16 of these municipalities had population sizes in excess of the Zipf estimates.

This trend can be ascribed to the fact that the majority of these dense rural municipalities are located in ex-Bantustan areas where the migration decision-making of the population residing in these areas has historically been severely limited by apartheid-based spatial development policies. The population figures of many of these municipalities were thus somewhat artificially inflated and those located above the Zipf estimates in 2001 moved closer to predicted figures by 2011 as a result of the ability of the population to now respond to normal push-and-pull factors influencing migration patterns to larger urban centres in many of these municipalities over the medium term.

Based on their position in the hierarchical distribution, it is expected that the service town municipalities will maintain their population figures or experience marginal population growth rates generally below the national average in the foreseeable
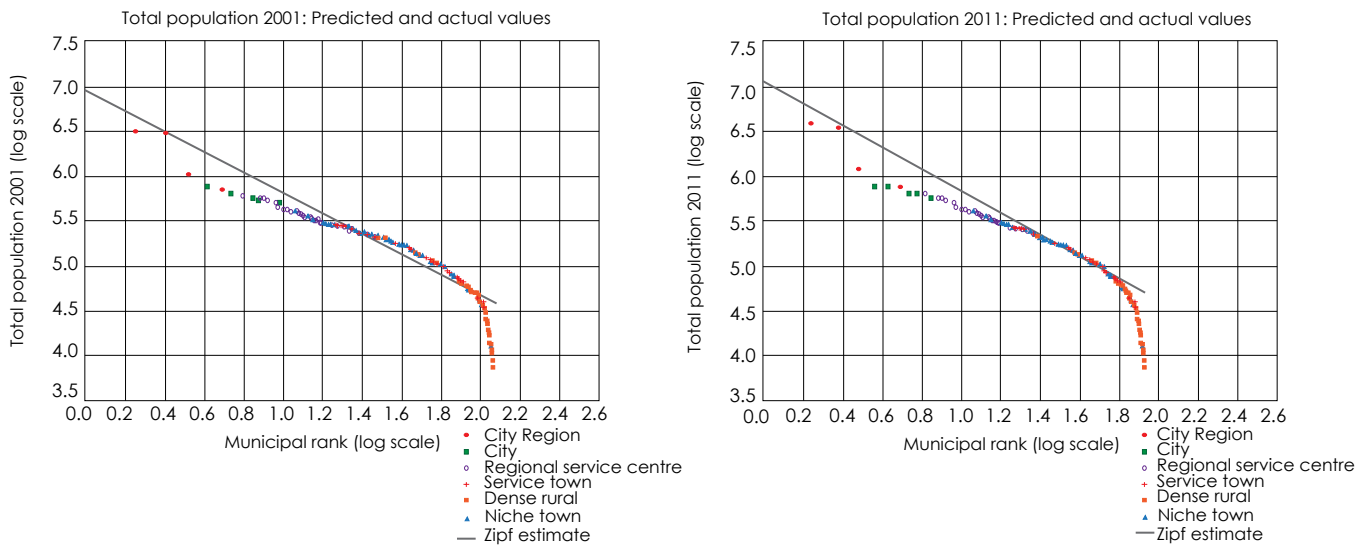

Figure 3a: Rank spread of all municipalities (2001 and 2011)

Figure 3b: Rank spread of city region and city municipalities (2001 and 2011)
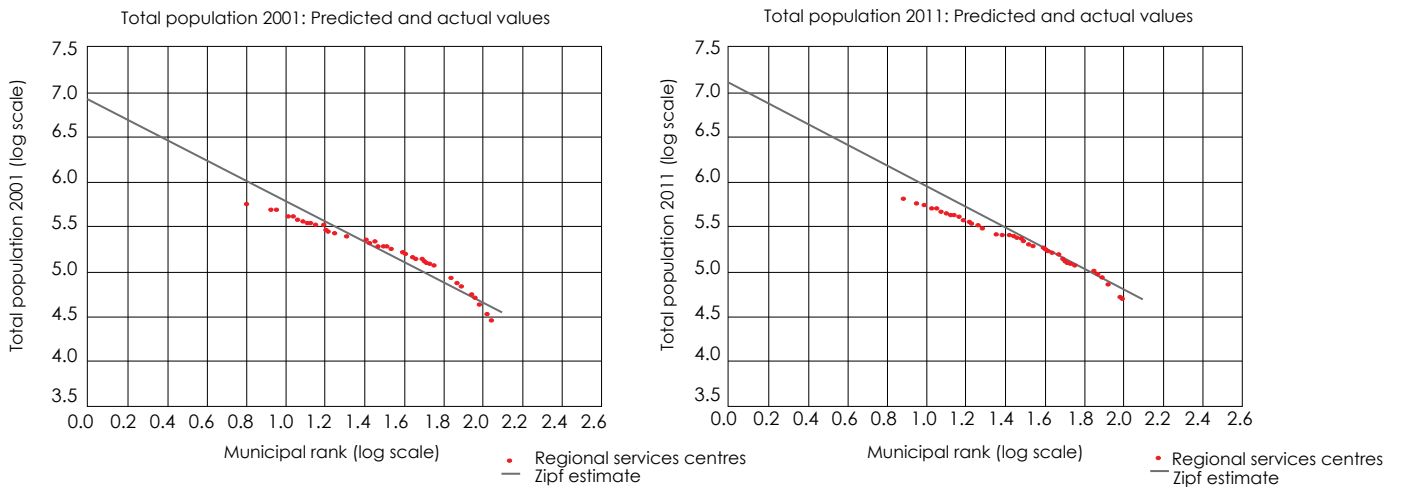
Figure 3c: Rank spread of regional service centre municipalities (2001 and 2011)
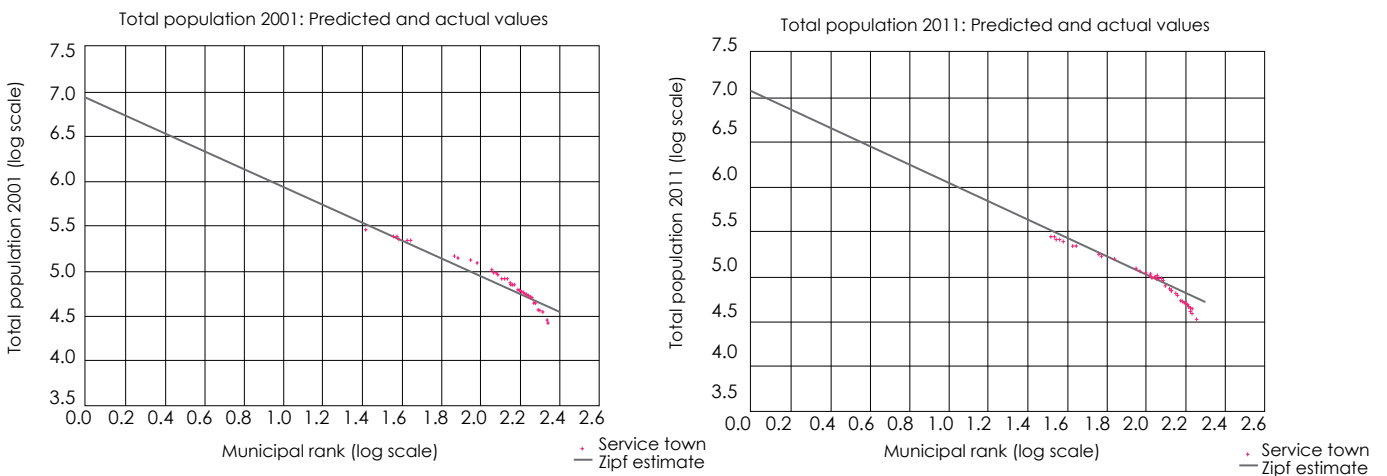

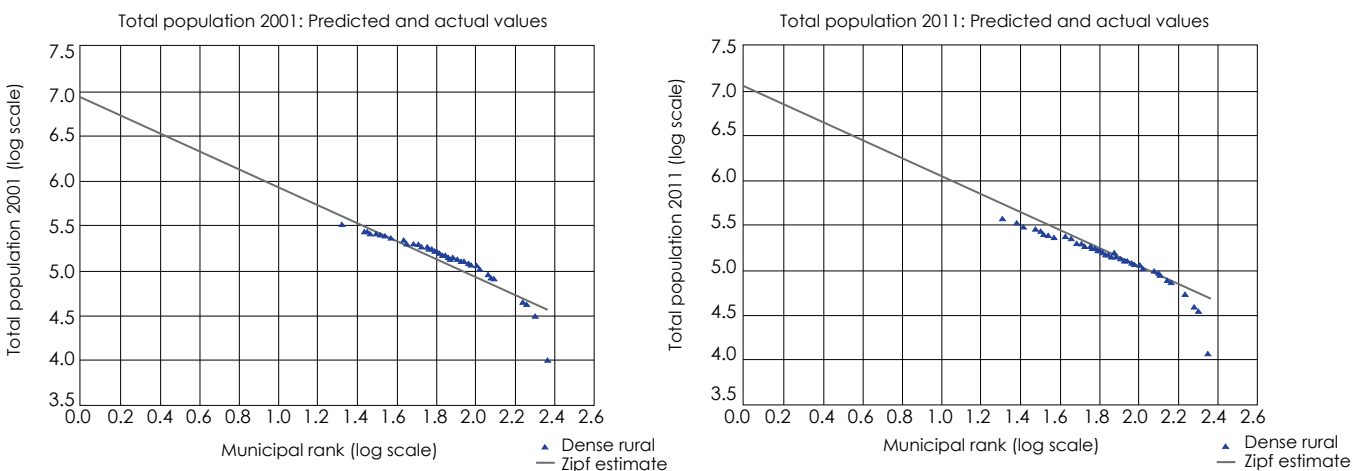Figure 3d: Rank spread of service town municipalities (2001 and 2011)

Figure 3e: Rank spread of dense rural municipalities (2001 and 2011)

future. The vast majority of niche town municipalities are located at the bottom end of the tail and, by 2011, mostly well below the Zipf estimates. It is expected that their population is likely to remain stagnant or decrease further in the foreseeable future.

3.3 Zipf rule applied to municipal economic (GVA) data

The analysis based on the municipal GVA and income data shows a significantly higher concentration in favour of the highest ranked municipalities with a values well in

excess of 1. The analysis based on the 2001 GVA data yields estimates of a at 1.31 for administratively defined municipalities and a marginally lower 1.26 for GJTFR. Marginal decreases are evident in the level of concentration of GVA among the highest ranked municipalities in 2011 with a = 1.27 for

Gross Value Added (GVA) 2001

Gross Value Added (GVA) 2011

Residual
Actual
Fitted
--- Standard error of regression (0.255963)

Residual
Actual
Fitted
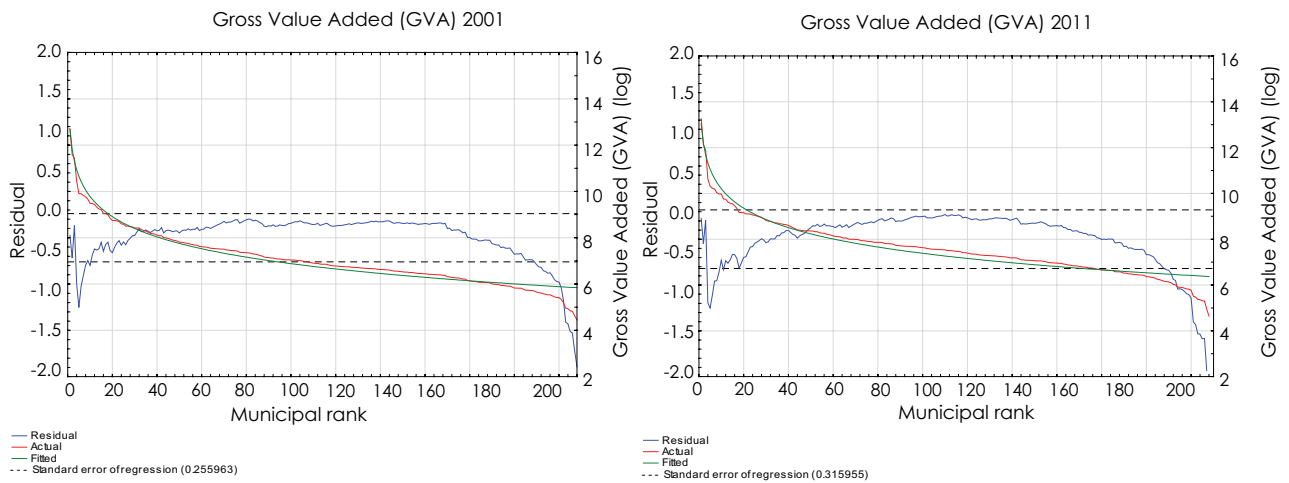--- Standard error of regression (0.315955)

Figure 4:   Actual, fitted and residual Zipf curves based on 2001 and 2011 GVA figures (Option 2)

administratively defined municipalities and a somewhat lower value of 1.23 for GJTFR. These results imply a marginally more even distribution of GVA outside the primate city region, examples being significant improvements in value added in mining areas in North-West and Limpopo, and municipalities surrounding the Cape Town metropolitan area and along the Cape South Coast during the period 2001 to 2011.

Following on the GJTFR at the top of the GVA hierarchy, the two city region municipalities of Cape Town (ranked 2nd) and eThekwini (ranked 3rd) are below the Zipf-based estimate, but within the standard error of regression of 0.26 in 2001 and 0.32 in 2011 (see Figure 4). No dramatic differences are apparent when comparing the

Table 3:   Results of Zipf's rule based on municipal GVA data

Data	Option 1 (All municipalities) (Equation [2])		
	α	t-value	R² adjusted
GVA 2001	1.31	70.75	0.95
GVA 2011	1.27	59.27	0.94
GJTFR	Option 2 (GJTFR) (Equation [2])		
	α	t-value	R² adjusted
GVA 2001	1.26	69.44	0.95
GVA 2011	1.23	56.05	0.93

2001 and 2011 graphs, although the discrepancies between the actual and estimated figures did become more pronounced. The position of the primate city region further strengthened against the other city and city region municipalities at the top end of the curve. A number of the municipalities

in the upper part of the hierarchy are characterised by GVA values well below the Zipf estimate and the standard error of regression. The middle section of the Zipf curve comprises the majority of municipalities with GVA values fitting within the standard error of regression. In 2011, it included nine city region and city municipalities, 32 regional service centres, 39 service towns, 57 niche towns, and 49 dense rural municipalities. The bottom tail of the Zipf curves for 2001 and 2011 constitutes municipalities whose actual GVA falls below the Zipf range and the standard error of regression and, in 2001, was comprised of 19 niche town municipalities and one dense rural municipality.

Figure 5 provides a spatial representation of this information and depicts the geographical spread of municipalities according to the same three categories as for the population estimates (labelled Bins 2, 3 and 4). It shows that the majority of the municipalities fall into Categories 2 and 3 and do not reveal any prominent spatial patterns or concentrations. Similar to the population estimates, the municipalities in Category 4 with
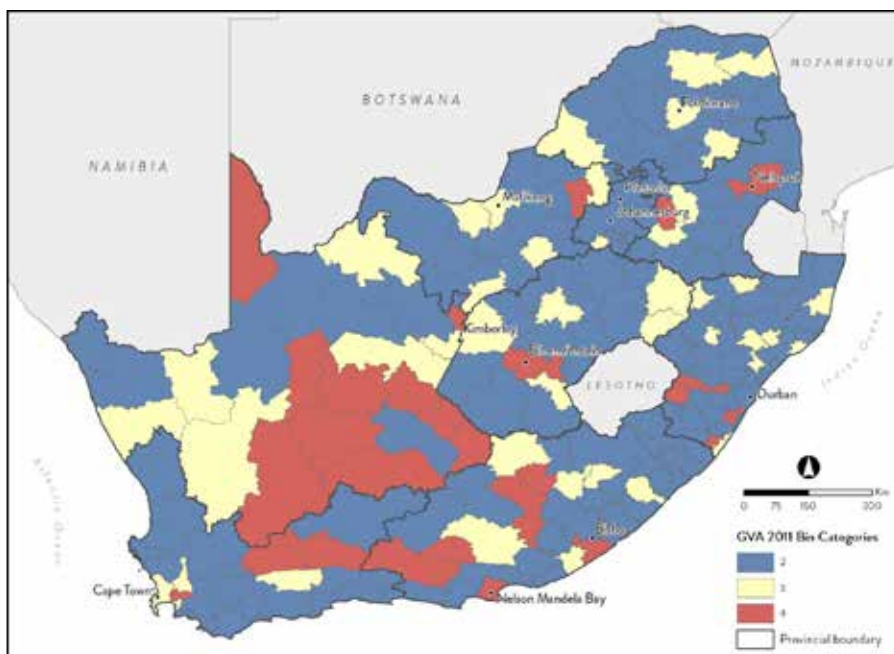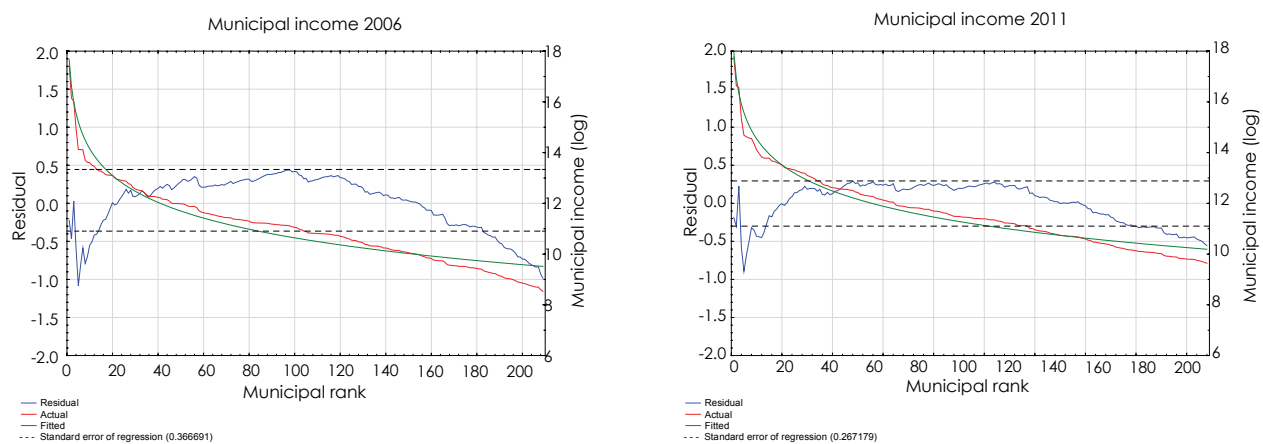

Figure 5:   Spatial spread of municipalities according to 2011 GVA

Figures 6a and 6b: Actual, fitted and residual Zipf curves based on 2006 and 2010 municipal income (GJTFR definition)

GVA estimates lower than the Zipf estimate and outside the standard error of regression range consist of two distinct groups. The first group is the highest ranked municipalities after the second and third municipality in the hierarchy (eThekwini and Cape Town) and includes municipalities such as Nelson Mandela Bay city region, and intermediate sized cities such as Mangaung, Buffalo City, Msunduzi, Mbombela and Rustenburg. The second group mainly consists of niche town municipalities located in the Northern Cape and a limited number of this type of municipalities scattered across the other provinces. From a future trajectory perspective, it is plausible to expect the economies (measured in GVA terms) of the municipalities forming part of the first group to expand substantially over the short to medium term to levels closer to the Zipf estimates. No significant improvement of the niche town municipalities at the bottom end of the tail is expected and the economic viability of these municipalities, which mostly remained unchanged in this position between 2001 and 2011, is questionable.

3.4 Zipf rule applied to municipal income

In line with the recommendations of Izsak (2006: 111) and Peng (2010: 3805), the analysis of municipal income based on the GJTFR as the highest order municipal entity necessitated the removal of the lowest ranked municipalities with insignificant values from the dataset (20 in the 2006-data and 40 in the 2011-data). The sample of municipalities in this analysis is thus smaller than for population figures and GVA. The results indicate that municipal income generation is highly skewed in favour of the highest ranked municipalities with an a value remaining at 1.77 between 2006 and 2011 for administratively defined municipalities and decreasing slightly from 1.53 to 1.48 for Option 2 (Table 4). This implies that the high-ranking municipalities generated a major share of municipal income during the period 2006 to 2011 and suggests little improvement of income-generating potential among lower ranked municipalities.

In 2011, the second- and third-ranked municipalities after GJTFR had values below the Zipf estimate and below the standard error of regression (Cape Town) and above the Zipf estimate and within the standard error of regression

range (eThekwini). The remainder of the municipalities in the upper tail performed not only below the expected Zipf estimate, but also below the Zipf range. These included two city region municipalities (Nelson Mandela Bay ranked 4th and Emfuleni ranked 5th); four city municipalities (Mangaung ranked 6th, Buffalo City ranked 7th, Msunduzi ranked 8th, and Polokwane ranked 13th), and four regional service centre municipalities (Rustenburg ranked 9th, uMhlathuze ranked 10th, Matlosana ranked 11th and Emalahleni ranked 12th).

The middle section of the Zipf curve (income above the Zipf estimate and mostly within the standard error of regression) has remained relatively unchanged since 2006 and, by 2011, comprised 115 municipalities, consisting mostly of regional service centre, service town, and niche town municipalities. The lower tail of the Zipf curve (municipalities that performed below the Zipf estimate and below the standard error of regression) consists of niche town and dense rural municipalities. The bottom part of the tail (below the standard error of regression) is also notably longer compared to the curves for population figures and GVA. This implies that the financial viability of these niche town and dense rural municipalities remains questionable.

3.5 Spatial relationships between rankings based on population, GVA and municipal income

Research has shown that Zipf's rule properties are retained despite general population movements over time (Eeckhout, 2004: 1448; Hsu, 2008: 1). To comprehend the complex changing patterns over time, it is necessary to simultaneously consider the relationship between population size, GVA and

Table 4: Results of Zipf's rule based on municipal income data

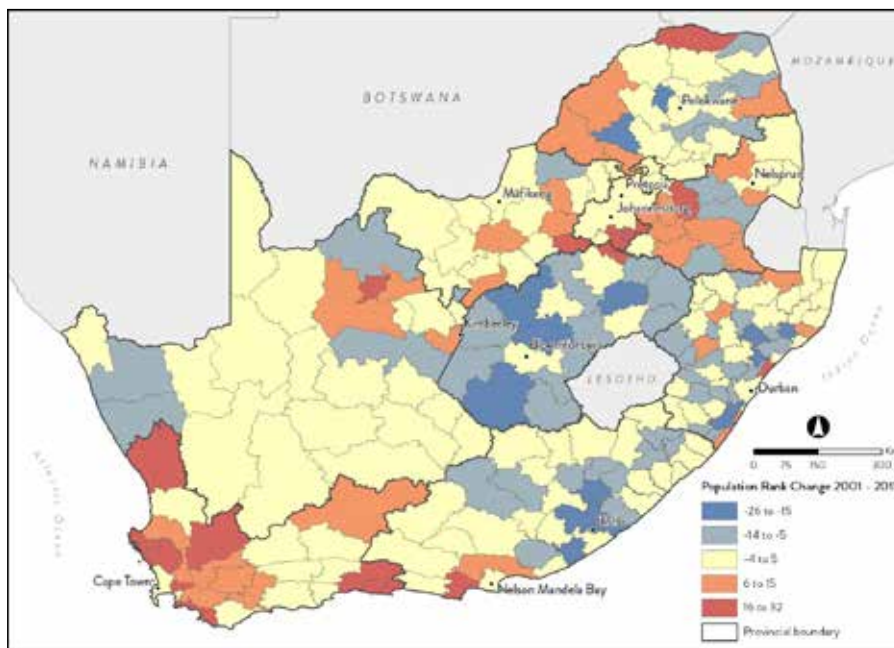| Data	Option 1 (All municipalities) (Equation [2])		
	a	t-value	R2 adjusted
Municipal income 2006	1.77	35.62	0.84
Municipal income 2011	1.77	44.00	0.89
GJTFR	Option 2 (GJTFR) (Equation [2])		
	a	t-value	R2 adjusted
Municipal income 2006	1.53	57.34	0.94
Municipal income 2011	1.48	72.35	0.97

Figure 7a: Municipal rank changes according to population size (2001 to 2011)

Figure 7b: Municipal rank changes according to GVA (2001 to 2011)

2001 and 2011. Municipalities with significant GVA rank decreases are concentrated in the Northern Cape and North-West provinces, as well as the north-eastern parts of the Free State and the southern parts of Mpumalanga (see Figure 7b). The most significant GVA rank increases are evident in municipalities along the Cape south coast, the northern parts of the Eastern Cape and Kwazulu-Natal, and parts of Limpopo.

Overall, the municipal rank pattern based on income size did not change significantly between 2006 and 2011. The spatial analysis of income rank changes (Figure 7c) indicates relative stability of municipalities at the upper end of the hierarchy and significant volatility of municipalities in the lower parts, with patterns of significant change scattered across the entire country.

The results reflected in Figures 7a to 7c thus reveal a significant number of potential permutations at municipal level. These permutations can be summarised in terms of six major categories of observed directions of changes in all three sets of variables (see Table 5). The first group of municipalities (Variety 1) consists of municipalities that remained relatively stable for all three variables over the study period. As much as 56% of city region municipalities fall within this category.

Variety 2 consists of municipalities that exhibit expected trends with rank improvements in population size, GVA and municipal income. A variant of the category (Variety 2b) consists of municipalities that were improving in population rank and GVA rank, but with declining municipal income rank. These category 2-type municipalities are most prevalent among city regions (33%) and niche town municipalities (22%). Variety 3, on the other hand, consists of municipalities that exhibit expected trends characterised by decreases in ranks of population size, GVA and municipal income. A variant of the category (Variety 3b) consists of declining municipalities in population rank and GVA rank, but with increasing municipal income rank. None of the city regions and city municipalities fall within this category which is dominated by regional service centre (25% of RSCs) and service town (29%) municipalities.

Variety 4 municipalities are characterised by a decline in population size rank despite

municipal income. Conceptually, municipalities with improving economies are intuitively expected to increase in terms of population ranking, and those with deteriorating economies to decline in terms of population ranking. In the context of this study, an increase in rank in terms of GVA between 2001 and 2011 is expected to be associated with an increase in population rank over the same period. The results, however, suggest a more complex relationship between population size and economic variables. Figures 7a, 7b and 7c illustrate rank changes in the geographical spread of population and GVA

between 2001 and 2011, and municipal income between 2006 and 2011.

The largest concentration of municipalities with significant population rank decreases is located in the Free State, the rural parts of Kwazulu-Natal and the Eastern Cape (see Figure 7a). A further very distinct pattern is the increase in population rank of municipalities bordering and in close proximity of the four largest metropolitan areas, especially around Cape Town and the Gauteng city region.

The GVA-based ranking of municipalities remained relatively stable in large parts of the country between

Figure 7c: Municipal rank changes according to income (2006 to 2011)

hierarchical distribution of municipalities remained stable over this period. There were, however, substantial changes in the relative position of individual municipalities and types of municipalities within this overall distribution. The delayed effects of the normalisation of population movements following the removal of race-based restrictive policies are borne out by the relative increases in the movement of the population towards the larger cities forming part of the upper end of the hierarchy and the changes occurring along the middle and lower part of the distribution curve.

The application of municipal economic data revealed a somewhat different picture with a = 1.26 and 1.23 for GVA in 2001 and 2011, respectively. This implies higher concentrations of economic activities in municipalities at the upper end of the hierarchy and indicates only a marginally more even distribution of GVA outside the major municipalities between 2001 and 2011. From a future trajectory perspective, it will be plausible to expect the economies of the municipalities at the top end of the curve (after the second and third municipality in the hierarchy), including municipalities such as Nelson Mandela Bay, Mangaung, Buffalo City, Msunduzi,

improvement in terms of GVA. Category 4 is especially prevalent among service town municipalities (32% of service town municipalities in this category) and niche town municipalities (33%). Variety 5 represents the opposite of variety 4 and is characterised by an increased population rank despite decreases in rank in GVA and/or municipal income.

category is fairly well represented in all municipal types (excluding city region municipalities), but is especially prevalent among city (40%) and dense rural municipalities (47%).

6.   CONCLUSIONS

The results presented in this article indicate that the Zipf rank size rule

Table 5:   Population, GVA and municipal income rank changes

Variety	1	2		3		4		5		6
										Changes within municipalities
Sub-varieties	a	a	b	a	b	a	b	a	b	a
ΔPopulation	Stable	↑	↑	↓	↓	↓	↓	↑	↑	0
ΔGVA	Stable	↑	↑	↓	↓	↑	↑	↓	↓	↑↓
ΔMunicipal income	Stable	↑	↓	↓	↑	↑	↓	↓	↑	↑↓

Table 6:   Frequency distribution by municipality type

Type of municipality	Change varieties (% of municipality type by category)										Total
	1a	2a	2b	3a	3b	4a	4b	5a	5b	6a	
City region	56	22	11	0	0	0	11	0	0	0	100
City	20	0	0	0	0	20	0	20	20	20	100
RSC	10	5	10	20	5	10	15	18	8	0	100
Service town	0	5	10	17	12	15	17	10	12	2	100
Niche town	0	9	13	9	4	15	18	24	4	6	100
Dense rural	0	6	11	9	4	8	8	45	2	8	100

Variety 5a consists of municipalities which, either through locational or cultural factors, improve in population rank, irrespective of a decline in GVA rank and municipal income rank. This

is applicable to municipal level population distribution in South Africa with a = 1.04. This value remained unchanged between 2001 and 2011, implying that the overall

Polokwane and Rustenburg, to expand substantially over the short- to medium-term future to values closer to those predicted by the Zipf curve. The marginal changes in a coefficients for

GVA between 2001 and 2011 hint at the limited impact of national policies such as GEAR, ASGISA and NGP since 2001, and the marginal spatial impact of local processes such as Local Economic Development Plans on economic variables at municipal level.

The a coefficient values for municipal income data were 1.53 and 1.48 in 2006 and 2011, respectively. This is indicative of a disproportionate distribution of municipal income from the primary municipalities in the hierarchy relative to the lower ranked municipalities. The decrease in the a value between 2006 and 2011 does, however, signify some tendency towards a more equal distribution of municipal income generation. The large-scale changes in ranking between 2001 and 2011 reflect some measure of instability and the financial viability of the niche town and dense rural municipalities at the bottom tail of the income-based Zipf curve thus remains questionable.

The results of this study also emphasise the significant conceptual limitations when using legally defined administrative municipal boundaries for analysis purposes. These administrative boundaries, especially in the case of larger cities and city regions, are apt to be inadequate to address factors such as extensive commuting or temporary migration from neighbouring municipalities, and exclude GVA and municipal income that should be covered within the boundaries of economically functional regions. There is thus a clear case to be made for providing official data not only along administrative boundary definitions, but also along economically functional boundaries.

The results not only confirmed the application of the Zipf rule to population figures at municipal level, but also demonstrated the joint effects of population size, GVA and income per municipality. It was found that municipalities are generally unique and may defy intuitively accepted conceptual expectations. For instance, some municipalities have shown substantial improvements in terms of GVA and municipal income rank, yet declined significantly in terms of population rank. In other instances, significant increases in population rank were revealed, despite decreases in both GVA and municipal income rank.

Although these results provide an interesting perspective on the South African urban system from a municipal perspective and are potentially valuable for regional level planning and policy-development purposes, a number of inherent limitations of rank-size type hierarchical distributions must be recognised. First, the impact of using different urban-regional units can be significant, and empirical analyses (such as presented in this article) often have no choice but to use administratively defined urban-regional units due to data limitations instead of more adequately defined functional economic urban-regional units. The application of a second analysis option, adopting an extended economic functional unit at the top end of the settlement hierarchy, and its impact on empirical results has been clearly illustrated. The impact of increasing suburbanisation on the growth of the larger metropolitan areas can thus also influence the results appreciably. Secondly, Zipf estimates from smaller samples tend to produce lower values. Analysis based on municipal spatial entities can thus result in values generally lower than would be the case for a larger number of individual cities and towns. Thirdly, the literature has indicated a relatively low correlation between rank-size measures and levels of primacy. This confirms the need for a variety of measures of city size distribution and not relying solely on the Zipf rank-size rule.

For more robust results, there is thus a need for further research on the validation of Zipf's rule using South African data. This could include more rigorous research on the application of the rank-size rule, extending the analysis from municipal level to also include individual cities and towns. This further application of Zipf's rule could provide an additional source of information to inform policy formulation in South Africa and fill some of the existing gaps between national and local municipality models. It could also provide institutions such as the National Planning Commission, the National Treasury, the Financial and Fiscal Commission and the Department of Cooperative Governance and Traditional Affairs with some spatial insights into likely future trends and trajectories of key variables at municipal level.

REFERENCES LIST

AUERBACH, F. 1913. Das gesetz der bevölkerungskonzentration. *Petermanns Geographische Mitteilungen*, 59(2004), pp. 74-76.

BERRY, B.L.J. 1964. Cities as systems within systems of cities. *Papers of the Regional Science Association*, 13(1), pp. 147-163.

BERRY, B.J.L. & OKULICZ-KOZARYN, A. 2012. The city size distribution debate: Resolution for US urban regions and megalopolitan areas. *Cities*, 29(Suppl. 1), pp. S17-S23.

CHRISTALLER, W. 1966. *Die Zentralen Orte in Süddeutschland*. Jena: Gustav Fischer.

DCGTA (DEPARTMENT OF COOPERATIVE GOVERNANCE AND TRADITIONAL AFFAIRS). 2009. National urban development framework: Harnessing a common vision for growth and development of South Africa's towns, cities and city-regions. Pretoria: Department of Cooperative Governance and Traditional Affairs and The Presidency and in partnership with the South African Cities Network. Working draft for consultation.

DF (DEPARTMENT OF FINANCE). 1996. *Growth, employment and redistribution: A macroeconomic strategy*. Pretoria: Department of Finance.

EECKHOUT, J. 2004. Gibrat's law for (all) cities. *The American Economic Review*, 94(5), pp. 1429-1451.

FRAGKIAS, M. & SETO, K.C. 2009. Evolving rank-size distributions of intra-metropolitan urban clusters in South China. *Computers, Environment and Urban Systems,* 33(3), pp. 189-199.

FUJIWARA, Y., DI GUILMI, C., AOYAMA, H., GALLEGATI, M. & SOUMA, W. 2004. Do Pareto-Zipf and Gibrat laws hold true? An analysis with European firms. *Physica A 335*: 197-216.

GABAIX, X. 1999. Zipf's law for cities: An explanation. *Quarterly Journal of Economics* 114(3), pp. 739-767.

GABAIX, X. & IOANNIDES, Y.M. 2004. The evolution of city size distributions. In: Henderson, J.V. & Thisse, J.F. (eds.), 1999. *Handbook of urban and regional economics*. Amsterdam: Elsevier, pp. 2342-2375.

GAO, H. & WU, K. 2008. Zipf's law and influential factors of the Pareto exponent of the city size distribution:

Evidence from China. *Frontiers of Economics in China* 3(1), pp. 137-149.

GEYER, H.S., DU PLESSIS, D.J., GEYER, H.S. & VAN EEDEN, A. 2012. Urbanisation and migration trends in South Africa: Theory and policy implications. In: OECD (Organisation for Economic Co-operation and Development), *Redefining "Urban": A new way to measure metropolitan areas*. s.l.:OECD, pp. 59-89.

GIBRAT, R. 1931. *Les inégalités* économiques. Paris: Librairie du Recueil Sirey.

HINLOOPEN, J. & VAN MARREWIJK, C. 2006. *Comparative advantage, Zipf's rule, and Zipf's law*. Tinbergen Institute Discussion Paper 06-100/1. Amsterdam: Tinbergen Institute.

HSU, W. 2008. *Central place theory and Zipf's law*. Ph.D. thesis. Ann Arbour, MN: Proquest.

IZSAK, J. 2006. Some practical aspects of fitting and testing the Zipf-Mandelbrot model: A short essay. *Scientometrics*, 67(1), pp. 107-120.

KAWAMURA, K. & HATANO, N. 2002. Universality of Zipf's law. *Journal of the Physical Society of Japan*, 71(5), pp. 1211-1213.

KNUDSEN, T. 2001. Zipf's law for cities and beyond: The case of Denmark. *American Journal of Economics and Sociology,* 60(1), pp. 123-146.

KRUGELL, W.F. 2005. *The geographical economy of South Africa*. Ph.D. thesis. Potchefstroom: North-West University.

LI, W. 2002. Zipf's law everywhere. *Glottometrics*, 5, pp. 14-21.

LÖSCH, A. 1954. *The economics of location*. 2nd edition. New Haven, CT: Yale University Press.

NAUDE, W.A. & KRUGELL, W.F. 2003. Are South Africa's cities too small? *Cities*, 20(3), pp. 175-180.

NEWMAN, M.E.J. 2006. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5), pp. 323-351.

NITSCH, V. 2005. Zipf zipped. *Journal of Urban Economics*, 57(1), pp. 86-100.

OKUYAMA, K., TAKAYASU, M. & TAKAYASU, H. 1998. Zipf's law in income distribution of companies. *Physica A*, 269, pp. 125-131.

PENG, G. 2010. Zipf's law for Chinese cities: Rolling sample regressions. *Physica A*, 389, pp. 3804-3813.

PERLINE, R. 2005. Strong, weak and false inverse power laws. *Statistical Science*, 20(1), pp. 68-88.

PHILBRICK, A.K. 1957. Principles of areal functional organisation in regional human geography. *Economic Geography*, 33(4), pp. 229-336.

QUANTEC. 2011. Quantec. [online]. Available from: <www.quantec.co.za> [Accessed: 14 November 2011].

ROSEN, K.T. & RESNICK, M. 1980. The size distribution of cities: An examination of the Pareto law and primacy. *Journal of Urban Economics*, 8(2), pp. 165-186.

ROSSI-HANSBERG, E. & WRIGHT, M.L.J. 2007. Establishment size dynamics in the aggregate economy. *The American Economic Review*, 97(5), pp. 1639-1666.

SOO, K.T. 2005. Zipf's law for cities: A cross-country investigation. *Regional Science and Urban Economics*, 35(3), pp. 239-263.

TANAKA, K. & HATSUKANO, N. 2011. The size distribution of all Cambodian establishments. *Economics Bulletin*, 31(3), pp. 2128-2137.

THE PRESIDENCY. 2006. *Background document: A catalyst for accelerated and shared growth - South Africa (ASGISA)*. Pretoria: Department of the Presidency.

VON THUNEN, J.H. 1966. *Von Thunen's isolated state*. Oxford: Pergamon Press.