07 OCCASIONAL PAPER

A composite index of quality of life for the Gauteng city-region:

a principal component analysis approach

Talita Greyling

Department of Economics and Econometrics, University of Johannesburg

Commissioned by the Gauteng City-Region Observatory

November 2013

Gauteng City-Region Observatory

UNIVERSITY OF JOHANNESBURG

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

ISBN Number: 978-0-620-59015-0

07

OCCASIONAL PAPER

A composite index of quality of life for the Gauteng city-region:

a principal component analysis approach

Talita Greyling

Department of Economics and Econometrics, University of Johannesburg

Commissioned by the Gauteng City-Region Observatory

November 2013

Abstract

Central to improving people's quality of life is the ability to measure this concept. This is, however, made difficult by the concept's multi-dimensional nature. The primary research objective of this paper was to construct a composite index to measure and compare the quality of life of different demographic and socio-economic groups across the Gauteng City-Region (GCR) in South Africa. The second research objective was to determine the dimensions that explain the most variance in the data set of each of the different demographic and socio-economic groups. A method introduced by Nicolette et al. (2000) that employs Principal Component Analysis (PCA) to weight the index was used, and this paper represents the first attempt in South Africa to apply this method. PCA was also used to analyse variance between the demographic and socio-economic groups. The paper found the quality of life scores of urban, high income, male, Asian and White, and younger respondents to be higher than those of the other groups. Furthermore, the quality of life scores of Africans, low income, female, older, and non-urban dwellers were relatively low. The dimension 'housing and infrastructure' explained the most variance for the groups with lower quality of life scores, while the dimension 'social relationships' explained the most variance in the data set for the groups with higher quality of life scores. Furthermore, the dimension 'socio-economic status' explained a high proportion of variance in all the groups. These dimensions give an indication of the areas to be addressed to improve quality of life.

Contents

List of Figures

List of Tables

1. Introduction

The improvement of the quality of life of all South Africans is high on the agenda at national (The National Planning Commission, 2012) and regional levels of government (The Gauteng Planning Commission, 2012) and it is therefore important to develop an instrument that can measure this multi-dimensional concept. The need therefore exists for a composite index of quality of life with the ability to both track the quality of life of people over time and compare it across different demographic and socio-economic groups. Such a measure could identify those demographic and socio-economic groups with low levels of quality of life and also highlight dimensions that need to be prioritised in order to improve the wellbeing of people.

In South Africa there are a limited number of quality of life indices and measures of wellbeing. Indices that measure wellbeing nationally include: the Quality of Life Index of Moller and Schlemmer (1983), the Living Standard Measure (LSM) Index produced by the South African Audience Research Foundation (SAARF) (2013), the South African Development Index of the South African Institute of Race Relations (2011), and the Everyday Quality of Life Index (Higgs, 2007). The following indices measure wellbeing at a regional level: the Quality of Metropolitan City Life in South Africa Index (Naude, et al., 2009), the Non-Economic Quality of Life Index at Sub-National Levels (Rossouw & Naude, 2008) and the Quality of Life Index of the Gauteng City-Region Observatory (GCRO, 2011).

Although these quality of life indices make distinctive contributions to the study field, the focus of these studies is often to measure only objective or subjective quality of life or only economic or non-economic quality of life, rather than all of the above. Furthermore, many of the indices use equal weighting, which does not necessarily reflect the priorities of the communities.

The primary objective of this paper is therefore to construct a composite index of quality of life to measure wellbeing in the Gauteng City-Region (GCR)[1] that will take into account objective, subjective, economic and non-economic dimensions and that will be objectively weighted. To this end, the paper employs a novel method introduced by Nicoletti et al. (2000) drawing on Principal Component Analysis (PCA), and used for the first time in South Africa in this paper to construct a composite index of quality of life. Furthermore, the quality of life of different demographic and socio-economic groups in the GCR is then compared using this newly constructed composite index.

The second objective of this paper is to compare the components that explain the most variance in the data set of the different demographic and socio-economic groups. Here a data set collected by GCRO in 2009 on quality of life in the GCR was used and separate PCAs were conducted for each group.

The remainder of this paper is structured as follows: Section 2 provides a brief literature review on the development and measurement of quality of life. This is followed by an explanation in Section 3 of the method used to construct and weight composite indices, the practical application of PCA and the selection of the indicator variables included in the analysis. Section 4 discusses the results of the PCA analysis and Section 5 explains the construction of the composite index of quality of life. In Section 6 the constructed

[1] The GCR is an area in South Africa which includes the Gauteng Province and the cities (municipalities) in adjacent provinces that are economically and functionally linked to the province and are within a 175 km radius of central Johannesburg. Gauteng is the smallest province in South Africa but contributes almost 33% to the country's Gross Domestic Product (GDP) (Gauteng Provincial Government, 2010).

index is used to calculate and compare the quality of life scores for the different demographic and socio–economic groups in the GCR. Section 7 reviews the results of the components, explaining the reason(s) for the variance between the different demographic and socio-economic groups. The paper ends with a series of conclusions based on the application of the PCA and the drawing up of the composite quality of life index.

2. Literature review on quality of life

This section briefly discusses the concept of quality of life and the theories and approaches that have developed over the years. The section also reviews the literature on the measurement of quality of life as well as current quality of life measures and quality of life studies in South Africa.

2.1 Quality of life: definitions and theoretical approaches

The literature does not contain a standard definition of quality of life. According to Sumner (2004) the definition of quality of life has evolved as the theory developed from a pure economic to a multi-disciplinary multi-dimensional approach. In the 1950s, quality of life was perceived as an economic phenomenon that was measured by a single indicator – GDP per capita. This perception changed, and in the 1970s it was recognised that quality of life is a much broader, multi-dimensional concept. The pure economic approach to quality of life was based on the theory of utility (Samuelson, 1948). Utility theory assumes that as income increases the consumption of goods increases, which in turn leads to higher levels of utility and wellbeing. In the 1970s the utility theory was replaced by the basic needs approach, which is much broader. Basic needs theory evolved to include a wider range of needs described as the human development theory. This theory assumes that people experiencing a high quality of life have significantly satisfied their developmental needs. The theory is based on Maslow's (1943) hierarchy of needs and includes lower-order needs, such as health, safety and economic factors, and higher-order needs such as social factors, self-esteem, actualisation, knowledge and aesthetics.

In the 1980s, one of the most influential theories of quality of life was developed by Nobel Laureate Amartya Sen (Costanza et al., 2007). According to this theory a person's life can be viewed in terms of functionings and capabilities. Functionings are the activities and situations that people consider as important in their lives. These can be captured through observable achievements such as health status, level of education and current employment status. Capabilities are an assessment of a person's life according to the degree of freedom a person has to choose among various functionings. The functionings or achievements are the end goals of human life; while capabilities are the 'freedoms of choice' a person has to experience the functionings (Costanza et al., 2007). This theory strongly influenced the Human Development Index, which is one of the most used developement indices and has been published since 1990 (United Nations Development Progamme, 2010).

In the late 1990s, the tradition of measuring wellbeing subjectively that originated in the behavioural sciences was increasingly accepted as an alternative approach to explain quality of life (Diener & Suh, 1997). The theory identified by Sirgy as "personal utility" is explained as the personal evaluation of community members' satisfaction with their "[o]verall life, social life, family life and spiritual life" (Sirgy, 2011: 6). According to Sirgy, many theories were developed under the umbrella of the subjective approach to quality of life. These theories include the hedonic psychology theory, the social judgement theory, the positive/negative affect theory, the human flourishing, flow and engagement theories and the purpose and meaning in life theories.[2]

[2] For a more detailed discussion of these theories see Sirgy (2011).

A recent development in the quality of life theoretical approaches is the concept of sustainability, which can be described as the effort to meet the needs of the present generation without compromising those of future generations. This theory considers both environmental and human wellbeing concepts and implies that people and the ecosystem must be treated equally. According to Sirgy (2011: 15) "[o]ne cannot have a good human condition in a bad environment".

As there is no standard definition of quality of life, how it is defined for the purposes of this paper is influenced by the theoretical framework adopted. Stiglitz et al. (2009: 58)[3] state that the information relevant to evaluate quality of life includes both "[p]eople's self-report and perceptions" as well as "[m]easures of their functionings and freedoms", and this paper therefore adopted the capabilities and functionings and the subjective wellbeing theories to describe the quality of life of people in the GCR. It follows that the definition of quality of life used in the paper is therefore fairly comprehensive and includes functionings and capabilities as explained by Sen (1985), and subjective measures as presented in the personal utility approach explained by Sirgy (2011). Quality of life is therefore defined as the extent to which objective functionings are fulfilled considering peoples' capabilities in relation to personal or group perceptions of subjective wellbeing (definition adapted from Costanza et al., 2007). According to Costanza et al. the perceived satisfaction can be affected by mental capacity, cultural context, information, education and temperament and the next section of the paper discusses how quality of life can be measured.

2.2 The measurement of quality of life

As the theoretical approaches to quality of life evolved, the measures of the concept also developed. Quality of life was originally measured by a single indicator – GDP. This approach was challenged by Easterlin (1974) who found that growth in GDP and subjective wellbeing correlate poorly. Therefore, GDP is not an adequate measure of the quality of life, and broader measures of wellbeing are needed (McGillivray, 2005).

One of the approaches to improve the measures of quality of life has been to adjust GDP by monetising different aspects that are not counted in GDP; for example, social and environmental factors. Other adjustments that have been made are to incorporate purchasing power parity and to include differences in income distribution among countries. The problem with many of these adjustments is the difficulty in quantifying and monetising these different aspects (Conceicao & Bandura, 2008).

Attempts to measure quality of life going beyond GDP have traditionally included two approaches. The first method used since the 1970s is an objective approach in which various social indicators are used to complement GDP. In this case, quantitative statistics which are readily observed are used, such as crime rates, life expectancy, years of education, disease rate and housing standards (Moller & Schlemmer, 1989). The second approach is subjective in nature and uses subjective measures that are personal judgments of objective conditions (Moller & Schlemmer, 1983) expressed as satisfaction or happiness. In recent years there has also been a movement to construct composite indices of quality of life that are multi-dimensional and include either objective, subjective or both types of indicators (Cummins, 2000). The composite indices can be constructed at different geographical levels, such as international, national, regional or community levels. Sharpe (1999: 47), however, argues that "[t]he geographical dimension has little effect on the construction of an index, as the only difference between constructing indices at

3 In the paper, "Stiglitz et al.", refers to the Report by the Commission on the Measurement of Economic Performance and Social Progress and is used interchangeably with "Stiglitz Report".

different geographical levels is the availability of data for the different geographical demarcations". In this research the main focus is on the construction of composite indices of quality of life that are multi-dimensional and include both objective and subjective measures.

2.3 A review of current quality of life indices: dimensions and indicators

This section reviews select quality of life indices which include both objective and subjective measurements. In addition, it reviews recommendations made in seminal reports on the construction of quality of life measures. These measures and reports have been developed or published recently and include: Your Better Life Index by the Organisation for Economic Co-operation and Development (OECD, 2011), Beyond GDP (European Commission, 2007), the Happiness Index of Bhutan (The Centre for Bhutan Studies, 2008), the Canadian Wellbeing Index (Canadian Index of Well-being, 2009), the recommendations of the Stiglitz Report (2009), the Happiness Index of the United Kingdom (UK) (Office for National Statistics of the United Kingdom, 2012) and the recommendations of the World Happiness Report (United Nations, 2012). The quality of life dimensions included in the measuring instruments and highlighted in the recommendations of the reports are shown in Table 1 below.

Table 1: The dimensions of quality of life measures

Dimensions of QoL* included in indices	Your Better Life Index	Beyond GDP	Happiness Index of Bhutan	Canadian WB** index	Stiglitz Report	Happiness Index of the UK	World Happiness Report
Housing	×			×			
Income/GDP	×	×	×	×	×	×	×
Jobs	×	×			×	×	×
Community involvement	×		×	×	×	×	×
Education	×	×	×	×	×	×	×
Environment	×	×	×	×	×	×	×
Civic engagement/ good governance	×		×	×	×	×	×
Health	×	×	×	×	×	×	×
Life satisfaction	×	×	×		×	×	×
Safety	×				×	×	
Work-life balance	×		×	×	×	×	
Culture			×	×			
Age, gender							×

***QoL=Quality of life**
****WB=Wellbeing**

Source: Authors' selection of indices

According to Table 1, the following dimensions of quality of life are included in the reviewed quality of life measures and are also recommended in the seminal reports on quality of life: housing, income, employment, community involvement, education, civic engagement and good governance, health, life satisfaction, safety, culture, work-life balance, a dimension which includes a selection of demographic characteristics, and an environmental dimension. As can be seen from Table 1 the same dimension often recurs in the different measuring instruments.

Each of these dimensions is measured by either observable, objective measuring items or by measuring items that assess the subjective responses of respondents. Table 2 provides a number of examples of the measuring items used in the selected quality of life measures.

Table 2: The measuring items used in contemporary measures of quality of life

Dimension	Objective measuring items	Subjective measuring items
Housing	Type of housing Piped water in the house Electricity for lighting	Satisfaction with dwelling
Income/GDP per person	Income per month	Satisfaction with money available Income relative to neighbours Satisfaction with standard of living Perceived socio-economic status
Jobs/employment	Being employed in the previous seven days Type of employment	Satisfaction with working conditions
Community/social connections	Membership of civic organisations	Satisfaction with time spent with family Satisfaction with time spent with friends Satisfaction with marriage
Education	Years of education Literacy rate School enrolment rate	
Environment	Emission indicators Access to clean water Deforestation	
Civic engagement/ good governance	Participation in elections	Opinion of political freedom Satisfaction with the performance of government
Health	Life expectancy Mortality rates HIV infection rates	Self-reported health
Life satisfaction		Perceived life satisfaction Perceived happiness
Safety	Crime rate	Self-reported safety
Work-life balance	Hours spent at work Number of hours available for leisure	
Culture	Language Socio-cultural participation	
Age, gender	Respondent's age Respondent's gender	

Table 2 illustrates that the majority of the dimensions of quality of life can be measured by objective as well as subjective measuring items.

In this research the selection of the measuring items to be included in a composite index of quality of life was guided by the reviewed theoretical approaches and the reviewed current quality of life measures.

In the next section literature on quality of life in South Africa is discussed.

2.4 Quality of life research in South Africa

Since the 1980s there have been a number of initiatives to collect quality of life data in South Africa. An important early survey was the South African Quality of Life Trends Project of the Institute of Social and Economic Research, which tracked the subjective wellbeing in South Africa from 1983 to 2010 (Moller, 2012). A set of 35 subjective indicators was developed and the full set of indicators was used to collect data at three points over the time period. In addition, two longer trend lines for global indicators measuring satisfaction with life as a whole and happiness were collected more frequently.

Other surveys that specifically focus on the quality of life indicators of South Africa are: the Afrobarometer, which reports on the social, political, and economic atmosphere in African countries; the Human Sciences Research Council's (HSRC) Social Attitudes Survey (SASAS) that includes two global subjective wellbeing questions as well as a subjective question on respondents' future outlook; the AMPS survey, conducted by the SAARF, and which includes LSM; and since 2008 the Southern Africa Labour and Development Research Unit (SALDRU) has been collecting quality of life data containing objective and subjective measures through the National Income Dynamics Study (NIDS).

In Gauteng, the Community Agency for Social Enquiry (CASE) conducted a quality of life survey in 2006 that focussed on three areas of Johannesburg. However, the most comprehensive regional survey that includes the Gauteng Province was conducted by the GCRO for the GCR, and it is this data that are utilised in this research.

Although quality of life surveys have been conducted in South Africa, there have been limited attempts to construct composite indices. Moller and Schlemmer (1983) found the use of a selection of indicators to be more adequate to describe quality of life in South Africa than synthesizing the indicators into a composite index. However, in recent years the demand for a composite index that can reflect the multi-dimensional concept of quality of life has increased. Such an index is used by policy-makers to gauge the overall progress of a society (OECD, 2008) and guide policy decisions.

Possibly the best-known quality of life index used in South Africa is the SAARF's Development Index that they have published since 1994. It includes 13 variables that have been collected in the AMPS data set relating to the provision of services and changes in wealth. The South African Institute of Race Relations (SAIRR) has also constructed a development index (2011). The index includes 27 variables that assess South Africa's performance across six areas including the economy, education, health, living conditions, gender, and crime. The choice of these six areas is based on SAIRR's assessment of the importance of these factors for development in South Africa. Furthermore, a Mercer index also exists which measures the quality of life of expatriates in cities across the world, including in Johannesburg and Cape Town. It is based on 39 criteria grouped into ten categories: the political-social environment, economic environment, socio-cultural, medical and health, schools and education, public services and transport, recreation, natural environment, consumer goods and housing.

Higgs (2007) developed the Everyday Quality of Life Index (EQoLI) to measure the development progress of South Africans. The measure includes the following areas of wellbeing: socio-economic status, urbanisation, health, stress/pressure, quality of the environment, satisfaction of human needs, connectivity, optimism, subjective wellbeing and an overall measure of wellbeing. The index is constructed by making use of correspondence analysis to determine the relative continuum of variables in which the weighting of the first factor, and in certain accounts the second factor, are used to weight the index. The study compiled intermediate composite indices which were then combined to form the EQoLI.

Furthermore, a composite index of quality of life was constructed for the GCR by the GCRO. This index includes 56 subjectively and objectively measured indicators, covering ten dimensions of quality of life: work, socio-political factors, security, life satisfaction, education/connectivity, community, family, housing, infrastructure and health. The index is constructed by calculating an average score for each dimension and adding the scores linearly.

On a sub-national level (metropolitan cities and municipalities), Rossouw and Naudé (2008) and Naudé et al. (2009) constructed non-economic quality of life measures that use PCA in the initial phases of the construction of composite indices.

Except for those focussing racial differences, studies highlighting the differences in the quality of life of different demographic and socio-economic groups in South Africa are scarce. In general, the studies focussing on racial differences find African households worse off than other race groups, as would be expected (see Moller & Schlemmer, 1983; Klasen, 2000; Moller and Dickhow, 2002; Higgs, 2007; Posel & Casale, 2011; GCRO, 2011; Moller, 2012;and the SASAS by the HSRC, 2013). These findings reflect the apartheid history of South Africa and its ongoing legacy.

The studies that compare other demographic and socio-economic groups found women, people in rural areas and people over the age of 46 to have poorer levels of quality of life than others (see Klasen, 2000; Higgs, 2007; Rossouw & Naudé, 2008 and GCRO, 2011). Rossouw and Naudé (2008) found that if geographical quality of life is measured, non-urban areas have a higher score than urban areas.

Furthermore, studies on quality of life and subjective wellbeing show similar results regarding the factors that influence wellbeing the most. These factors are: housing, basic services, social relationships, education, employment and safety (Kingdon & Knight, 2003; Powdthavee, 2003; Bookwalter, et al., 2006; CASE, 2006; Higgs, 2007; Hinks & Gruen, 2007; GCRO, 2011; Moller, 2012 and HSRC, 2013). Amongst these factors, those that appear to be the most influential are housing and basic services, employment and education.

3. Methodology

This section reviews the methodology used to construct and weight composite indices. It describes PCA and the practical application of this method. Furthermore, the GCRO Quality of Life Survey data set is described, as well as the selection of the indicator variables.

3.1 The construction of composite indices

According to Sharpe and Smith (2005: 7) a composite index "[i]s the aggregation of individual indicators into a single index or bottom line using a certain weighting scheme".[4]

McGranahan *et al.* (1972) set out certain steps that need to be followed in the construction of a composite index. First, the dimensions need to be selected based on a theoretical framework. This is known as the top-down approach, although in the construction of composite indices of quality of life the bottom-up approach is often used (Sirgy, 2011). The bottom-up approach uses survey methods to determine which dimensions should be included in a composite index without adopting a certain overarching theory that guides the selection of the dimensions (Dluhy & Swartz, 2006). According to Dluhy and Swartz (2006) the selection of the appropriate theory should be guided by the fitness-for-purpose principle and it should clearly define the concept that needs to be measured as well as the sub-components, the selection of the individual indicators and an indication of which components are important to guide the weighting of the indicators.

The second step according to McGranahan et al. (1972) is to acquire good quality data as the quality of the data determines the soundness of the composite index. The third step is to select a method to treat missing data. According to the *Handbook on the Construction of Composite Indices* (OECD, 2008) there are three ways of dealing with missing data: deleting the case listwise or pairwise if any of the variables are missing, single imputation, or multiple imputation (OECD, 2008). The fourth step is the exploration of the data using multivariate analysis techniques to determine the underlying structure and constructs in the data.

McGranahan et al. (1972) describe the fifth phase in the construction of a composite index as the weighting and aggregation of the index. The selection of the weighting method is one of the main challenges in the construction of composite indices as the method should address the research question, it should be acceptable to policy-makers and the people whose quality of life is assessed, and it should be a true reflection of the measured quality of life. The final stage is to assess the robustness of the composite index.

Various advantages and disadvantages of composite indices are debated in the literature. One of the advantages mentioned is that the indices make it easier for the general public and policy-makers to follow trends in these indices rather than to identify common movements across many separate indicators (OECD, 2008). Furthermore, the indices are useful tools in policy analysis as they help set policy priorities and benchmark performance (Nardo et al., 2005) and provide a means to compare different measuring units of analysis in which the different indicators are measured.

Disadvantages mentioned are that composite indices can communicate misleading information if the index is poorly constructed and the selection and weighting of the indicators is not transparent (Sharpe,

[4] In the literature a composite index is also referred to as a synthetic index.

2004). These indices can also contribute to users or policy-makers reaching the wrong conclusions and consequently making incorrect policy decisions. In addition, composite indices can disguise poor performance in certain dimensions and therefore confuse remedial action (OECD, 2008).

Judged by the increasing demand for composite indices by policy-makers and the growing number of composite indices that have been developed across the world over the last few years (Bandura, 2006), it appears that composite indices have advantages.

According to Sumner (2004), a constructed composite index should be: relevant to the policy-makers, measurable, cost-effective, user-friendly, a direct measure of progress, specific to the phenomena, valid at all times, reliable, not be easily manipulated, and up to date at all times.

3.2 The weighting of composite indices

As mentioned in the previous section, one of the main challenges in the construction of composite indices is the selection of an appropriate weighting method. This section will therefore discuss the weighting of indices in more detail.

According to the *Handbook on Construction of Composite Indices* (OECD, 2008) there are participatory methods which can be described as subjective methods of weighting, and statistical methods which can be described as objective methods of weighting.

One of the first subjective methods mentioned by the *Handbook on the Construction of Composite Indicators* (OECD, 2008) is to weight an index according to the imposition of a researcher's own judgement on the priority of the quality of life dimensions. A second group of subjective weighting methods are participatory methods incorporating the opinions of stakeholders (this may include experts, citizens and politicians) in order to determine the weights of the dimensions.

An example of a participatory method is the budget allocation approach. Experts are given a 'budget' of a certain number of points that can be distributed over different individual indicators. The experts can then give higher weights to those indicators they wish to emphasise (OECD, 2008).

The main disadvantage of subjective measures of weighting is that it reflects the judgement of the researcher or the consulted experts and not necessarily the communities' opinion on the priorities of the quality of life dimensions (Ram, 1982). Furthermore, if the opinions of experts are used to determine the weights of the different dimensions, for example the budget allocation method, only a limited number of variables can be included, as too many indicators complicate the reasoning to establish which indicator is more important than the other. Subjective measures for the weighting of indices are also not conducive to the construction of comparable synthetic indices, as the weighting of different indices for different groups varies. And it is difficult to find a weighting system that is acceptable and endorsed by the majority of the assessed groups and policy-makers. To overcome the shortcomings associated with subjective measures of weighting indices, objective or statistical weighting measures may be used. With statistical methods the weights are not arbitrarily assigned but based on information contained in the data set.

There is an array of statistical methods which have been used in the construction of composite indices including equal weighting, principal components analysis, factor analysis, Data Envelopment Analysis (DEA), the 'benefit-of-the-doubt' approach, multiple regression, Structural Equation Modelling (SEM), and various combinations of these methods.

According to Hagerty and Land (2007), equal weighting is the most frequently used method to weight composite indices. Equal weighting does not imply no weighting, but rather that all the dimensions are equally important, even though this is seldom the weighting bestowed by community on each dimension. Equal weighting can also lead to double counting in the index (OECD, 2008) if highly correlated variables describing the same dimension are included in the index with each having the same weight. This skews the value of the synthetic indicator towards the dimension with the highly correlated variables. The main reason for employing equal weighting in the construction of indices is the simplicity of the method (OECD, 2008).

Multiple regression analysis is an alternative method to weight indices in which the regression coefficients are used as the weights. A benefit of using the regression coefficients is that the coefficients can to some degree explain the relationships between a large number of indicators (independent variables) and the dependent variable. One of the drawbacks of the method is that the variables included in the analyses need to conform to the assumptions of multiple regression analysis which includes linearity and independence of the explanatory variables (OECD, 2008). These assumptions are restrictive as they may limit the choice of variables that can be included in the synthetic indicator.

Recently, researchers have also used the benefit-of-the-doubt approach which uses DEA to weight synthetic indicators (Cherchye et al., 2007). This method uses linear programming to project an 'efficiency frontier' that is used as a benchmark to measure the performance of entities such as countries, regions or municipalities. A set of weights is then derived from the comparison of the distance between the entity's measure and the benchmark measure. One of the major advantages of this methodology is its flexible weighting as each decision-making unit has a unique set of weights. These weights are used to highlight the comparative advantages of the different decision-making units and to evaluate the performance of the units. This is an ideal method to use if the research objective is to evaluate and compare the quality of life of communities, while recognising the dimensions in which the community performs well.

PCA and Factor Analysis (FA) are of the most often used multivariate statistical techniques to weight composite indices (Booysen, 2002). To derive the weights using PCA or FA the factor loadings of the variables on the first component are used as the variation explained by the first component is generally sufficient to adequately represent the original variables (Ram, 1982). However, if the explanatory value of the first component is less than 55%, subsequent components should be included to derive the weights to prevent useful non-redundant information being lost (Aivazian, 2005: 74).

Often PCA and FA are used interchangeably as both are variable reduction methods, although they do differ in a number of ways. The purpose of PCA is to extract the maximum variance from a data set with a few orthogonal components, whereas the main purpose of FA is to analyse the covariance or communality in a data set. If FA is used as an analytical technique, only the variance that each observed variable shares with other observed variables is available for analysis.

If a choice needs to be made as to which method is most appropriate for constructing composite indices, Tabachnick and Fidell (2007: 635) recommend the use of FA "if the researcher is interested in a theoretical solution without error variability or without a unique mathematical solution". However, if the researcher needs an empirical summary of the data set that explains the maximum variance with a unique mathematical solution, then PCA is preferred.[5] The *Handbook on Constructing Composite Indices* recommends the use of PCA in the development of composite indices "as it has the virtue of simplicity and allows for weights

[5] For discussions on PCA versus FA see Tabachnick and Fidell (2007: 634-640).

representing the information content of individual indicators" (OECD, 2008: 69). Furthermore, using PCA avoids duplication of information as the components are orthogonal and it is a useful method to select the variables that should be included in a synthetic index (Somarriba & Pena, 2009).

The weights derived from the PCA are determined by the factor loadings on the components and are fixed across all groups. Therefore, an instrument is derived which compares the quality of life of different groups. But PCA also has certain weaknesses, for example the weights can only be estimated if there is sufficient correlation between the selected variables. In addition, the units of measurement change the principal components and it is therefore important to standardise the units of measurement before analysis (Brooks, 2008).

Guided by the primary research objective of this paper and considering the strengths and the weaknesses of the different weighting methods, PCA is considered to be the appropriate method to weight the composite index of quality of life in the GCR and to compare the wellbeing across different demographic and socio-economic groups. PCA objectively weights the dimensions included in the composite index, therefore limiting *a priori* information needed on the weighting. It is a basic method to determine the weights and it contributes to the selection of a limited number of indicator variables which represent the data and the relevant dimensions of quality of life in a region. Moreover, both objectively and subjectively measured indicators, as recommended by the Stiglitz Report (Stiglitz et al., 2009), can be included in the composite index.

3.3 A description of PCA

PCA is a multivariate statistical technique that, when applied to a data set, reveals which variables in the set form coherent subsets that are relatively independent of one another.[6] The variables that are highly correlated are combined into components. The components are expected to reveal the underlying processes that have created the correlation among the variables (Tabachnick & Fidell, 2007).

PCA aims to extract the maximum variance from a data set with each component (Tabachnick & Fidell, 2007). "The first principal component is the linear combination of observed variables that maximally separate subjects by maximising the variance of their component scores" (Tabachnick & Fidel, 2007: 635). The second component is computed from the residual correlations. It is the linear combination of observed variables that extracts maximum variability. This variability is uncorrelated to the first component. The subsequent components also extract maximum variability from the residual correlations and are independent from all the other components (Tabachnick & Fidell, 2007). The extracted components represent most of the variance of the original data set and can be used in further analysis.

In mathematical terms, PCA can be explained as follows:

From a set of variables X_1, X_2 to X_m, the principal components PC_1 to PC_m are extracted:

$$PC_1 = a_{11}X_1 + a_{12}X_2 + \ldots\ldots + a_{1n}X_n$$

$$\vdots$$

$$PC_m = a_{m1}X_1 + a_{m2}X_2 + \ldots\ldots + a_{mn}X_n$$

[6] See Tabachnick and Fidell (2007) for a comprehensive explanation of PCA.

where a_{mn} represents the weight for the mth principal component and the nth variable. The weights of each principal component are given by the eigenvectors of the correlation matrix or the co-variance matrix. The variance (λ) for each principal component is given by the eigenvalue of the corresponding eigenvector. The analysis conducted for this paper relied on the eigenvalues and their corresponding eigenvectors, as they summarise the variance in a correlation or covariance matrix.[7]

3.4 The practical application of PCA

The first step in conducting a PCA is to verify that there is sufficient correlation between the variables in order to perform the analysis. According to Tabachnick and Fidell (2007) the tests that should be used to determine the correlation are the Kaiser-Meyer-Olkin (KMO) measure and the Bartlett test of sphericity.

The KMO statistic is a ratio of the sum of squared correlations to the sum of squared correlations plus the sum of squared partial correlations (Tabachnick & Fidell, 2007). The KMO statistic should be at least 0.6 in order to proceed with factor analysis (Kaiser & Rice, 1974).

The Bartlett test of sphericity is used to test the null hypothesis that variables in the population correlation matrix are uncorrelated; that is that the correlation matrix is an identity matrix. If the associated probability is less than 0.05, the null hypothesis is rejected.

The second step is to identify a certain number of latent components (fewer than the initially selected variables included in the analysis) that represent the data. As different types of variables were included in the PCA with divergent units, a correlation matrix rather than a covariance matrix was used to extract the components. The use of a correlation matrix ensured that all the data had equal weighting (Vyas & Kumaranayake, 2006). When deciding on the number of components to extract, two factors were considered: the parsimony concept and the need to explain as much of the variance in the original data set as possible.

There are three techniques that guide the decision as to the number of components that needs to be extracted. The first is the Kaiser's criterion or the eigenvalue rule. The eigenvalue (variance) for each principal component indicates the percentage of variation explained in the total data set. Using this rule, components with an eigenvalue of 1.0 or more are extracted; an alternative measure is to extract the number of components that cumulatively contribute to the explanation of the overall variance by more than 50%. Second, the scree test can be used. The scree plot shows each of the eigenvalues of the components. The number of components to extract can be determined by identifying the point on the scree plot at which the shape of the curve changes direction and becomes horizontal. All the components above this point can be retained. Third, the interpretability of the rotated components is considered (Tabachnick & Fidell, 2007).

After the components have been extracted the factor loadings of each of the variables on the components are calculated. The factor loadings are the correlations between the latent components and a variable.

The next step in the PCA process (step three), deals with the rotation of the data, which is used to minimise the number of individual variables that have a high loading on a specific component. When the data are rotated it transforms the factorial axes to obtain a simpler structure of the components, which reveals the differences between the extracted components. The ideal is that each variable loads on only one of

[7] For an illustration of the calculation of eigenvalues and eigenvectors see Tabachnick and Fidell (2007).

the components, though the variables are often complex and load on more than one component (OECD, 2008).

To decide which type of rotation to use it first had to be determined whether the data were oblique or orthogonal by performing different tests. To test the data it was assumed that the data were oblique and therefore selected the oblimin method of rotation. As part of the output of this method, a correlation matrix indicating the correlations between the individual components is produced. If the correlation is less than 10% between the components it can be assumed that the data are orthogonal and varimax, and a method suitable for orthogonal data, is used as rotation method (Tabachnick & Fidell, 2007). After the rotation, a matrix of factor loadings is produced which was used in further analyses and in the construction of the composite index.

3.5 The GCRO Quality of Life Survey data set

The data set analysed in this paper is the GCRO Quality of Life Survey data set. It was collected by the GCRO in 2009 and released for analysis in May 2010. The data were collected in order to measure the quality of life and customer satisfaction of people living in the GCR.

The GCRO used a stratified sampling method to collect the data. The sampling frame was based on a population universe as at Census 2001. The sample was stratified by municipality to ensure significant coverage and a total of 6 639 respondents in 602 wards in 17 different municipalities were interviewed (GCRO, 2011).

3.6 Selection of the indicator variables

In order to measure the quality of life of the respondents in the sample, the initial selection of indicator variables of quality of life was based on the adopted theoretical approaches, the dimensions and indicator variables included in the reviewed indices (see Section 2.3), and the availability of indicator variables in the data set. The initially selected variables are discussed in Section 3.6.1 below and represent the following dimensions of quality of life: 'housing and infrastructure (basic services)', 'social relationships', an 'economic dimension', 'education', 'health', 'governance', 'safety' and 'global satisfaction with life'.[8] Section 3.6.2 describes the selection of a concise set of indicator variables from the initially selected variables that offer a good representation of the data set and explains an acceptable proportion of the variance in the data. These indicator variables will be used in further analysis.

3.6.1 The initially selected indicator variables

To measure 'housing and infrastructure' the following indicator variables were included in the data set, and selected for inclusion in the analysis: 'type of dwelling', 'type of sanitation', 'electricity used for lighting', 'satisfaction with dwelling', and 'piped water on the premises'. All these variables are nominal and were recoded into dichotomous variables with one indicating a positive influence on quality of life, for example 'living in a formal dwelling', and zero indicating a negative effect on quality of life, for example 'not living in a formal dwelling' (see Appendix A for a description of the coding of the variables).

To measure 'social relationships' four indicator variables were selected for inclusion in the data set: 'satisfaction with time with family', 'satisfaction with time with friends', 'satisfaction with time to do

[8] Global satisfaction with life measures the satisfaction of a person with all spheres of his/her life.

own things' and 'satisfaction with marriage'. The respondents were requested to rate their perceptions on a five-point scale ranging from 'very dissatisfied' to 'very satisfied'.

To measure the dimension related to 'economic factors' the following variables were selected: 'work conditions', 'income', 'standard of living', 'satisfaction with amount of money available', 'satisfaction with life' and 'perceived social status'. The income variable was recoded to reflect different income categories. The other indicators were rated on a five – point scale ranging from 'very dissatisfied' to 'very satisfied'.

To measure 'education' the only indicator variable available in the data set was 'years of education', and this variable was consequently selected for inclusion.

Based on the recommendations of the Stiglitz Report (2009) and the UN's Happiness Report (2012) it was decided to also include a global satisfaction indicator, namely 'how satisfied are you with your life'.

Three indicator variables were selected to measure 'health'. Two of the three indicators determined if a person's health kept him/her from doing their daily work or taking part in social activities, and the third indicator measured if the respondent was 'satisfied with their health'. Each of these indicators was measured on a four-point scale ranging from 'never' to 'always'.

In order to measure 'safety' three indicators were selected: 'how safe do you feel in the area where you live during the day', 'how safe do you feel in the area where you live after dark', and 'how safe do you feel'. All three indicators were categorical indicators with one indicating 'feeling unsafe' and five indicating 'feeling very safe'.

The final dimension included in the analysis was 'governance', and we selected five variables were selected: 'satisfaction with local government', 'politics is not a waste of time', 'the country is going in the right direction', the 'judiciary is free from influence' and the 'election was free and fair'. These indicators were measured on a five-point scale ranging from 'very dissatisfied' to 'very satisfied'.

3.6.2 The final selection of the indicator variables

To select the final set of indicator variables it was necessary to conduct successive rounds of PCA during which different combinations of the initially selected variables was used to eliminate those variables with the lowest communalities.[9] This was done in order to arrive at a set of indicator variables that, compared to other sets of indicator variables, explains the most variance in the data set. The final selection is shown in Table 3.

[9] Amount of variance in the indicator explained by the component.

Table 3: The final selection of indicator variables

Variable	Type of data	Min	Max	Mean	SD*
Type of dwelling	Dichotomous	0	1	0.8505	0.3566
Piped water on premises	Dichotomous	0	1	0.9160	0.2774
Electricity used for lighting	Dichotomous	0	1	0.9050	0.0036
Satisfied with time available for family	Ordinal	1	5	3.9985	0.8878
Satisfied with time available for friends	Ordinal	1	5	3.7534	1.0157
Satisfied with time available for own things	Ordinal	1	5	3.5010	1.0707
Work conditions	Ordinal	1	5	3.5117	1.1588
Level of education	Ordinal	1	5	3.6589	0.0127
Income category	Ordinal	1	5	2.3907	0.9315
Satisfaction with life	Ordinal	1	5	3.1718	0.0153
Satisfaction with health	Ordinal	1	4	3.2115	0.7039
How often does health prevent you from taking part in social activities	Ordinal	1	4	3.1900	0.8720
Feel safe at home	Ordinal	1	5	4.2874	0.2125
Satisfied with local government	Ordinal	1	5	2.9142	0.0144
Judiciary is free	Ordinal	1	5	3.1100	1.0900

***Standard deviation**

Source: Authors' own calculations using GCRO's Quality of Life Survey data (GCRO, 2009)

The selected variables set out in Table 3 were used in the empirical analysis to address the primary and secondary research questions.

4. Results

This section reviews the descriptive statistics of the GCRO data set and discusses the results of the PCA.

4.1 Descriptive statistics of the GCRO Quality of Life Survey data set

Table 4 below summarises the descriptive statistics on the demographic and socio-economic characteristics of the sample population.

Table 4: Descriptive statistics of the Quality of Life Survey data set

Geo type	Frequency (N)	% of Sample
Urban formal (built up town or city area)	4 156	62.6
Urban informal	1 654	24.9
Peri-urban (mostly informal/smallholding)	609	9.2
Tribal settlement	82	1.2
Farming	92	1.4
Income groups		
R0-R1600	2 157	33
R1 601-R12 800	3 065	46
R12 801 – R102 400	720	11
R102 401 – more	65	1
Race		
African	5 452	82.2
Asian/Indian	79	1.2
Coloured	246	3.7
White	859	13.0
Gender		
Male	2 708	40.8
Female	3 928	59.2
Age		
18-20	460	7
21-30	1 971	30
31-40	1 604	24
41-50	1 120	17
51-65	934	14
66-75	347	5
75+	133	2

Source: Authors' calculations using GCRO's Quality of Life Survey data (GCRO, 2009)

Table 4 illustrates that the majority of respondents in the sample reside in urban formal or urban informal areas, which indicates the degree of urbanisation of the GCR; while only 2.6% of the respondents live in tribal or rural areas.[10]

The average income of the respondents was in the income bracket of R1 601 to R12 800 per month, with only 12% of the sample indicating an income in excess of R12 800 per month. The share of income of the lowest decile of the respondents was merely 2% of the total income earned by households in the sample, compared to 68% earned by the highest decile of income earners. This reflects the considerable skewness of the income distribution in the GCR (GCRO, 2011).

The sample reflects the race distribution of South Africa as approximately 80% of respondents were African. More than half of the respondents were female and almost 93% of the respondents were of working age between 18 and 65 years.

4.2 The results of the PCA

The KMO measure of sampling adequacy was 0.745, which indicated a high sampling adequacy for FA. The Bartlett test of sphericity associated probability was less than 0.05 at 0.000, which was small enough to reject the null hypothesis of no correlation in the data set. Therefore both these tests indicated that there was sufficient correlation in the selected variables to perform PCA.

To decide on the number of components to extract the Kaiser rule, the scree plot and the interpretability of the rotated component matrix were used.

To consider the Kaiser rule the results shown in Table 5 below were examined. In the first column showing the initial eigenvalues without restricting the number of extracted components, it was noted that component six had an eigenvalue of 0.858 which is less than one and, according to the Kaiser rule, should then not be considered for extraction. Based on the Kaiser rule five components should therefore have been extracted. To verify this decision the scree plot and the interpretability of the rotated components were also considered.

[10] The majority of Gauteng's residents stay in urban formal and urban informal areas, and very few stay in tribal and rural areas. Therefore, for the purposes of this analysis, only respondents in urban formal and urban informal geographical types were considered.

Table 5: Total variance explained by the eigenvalues of the extracted components

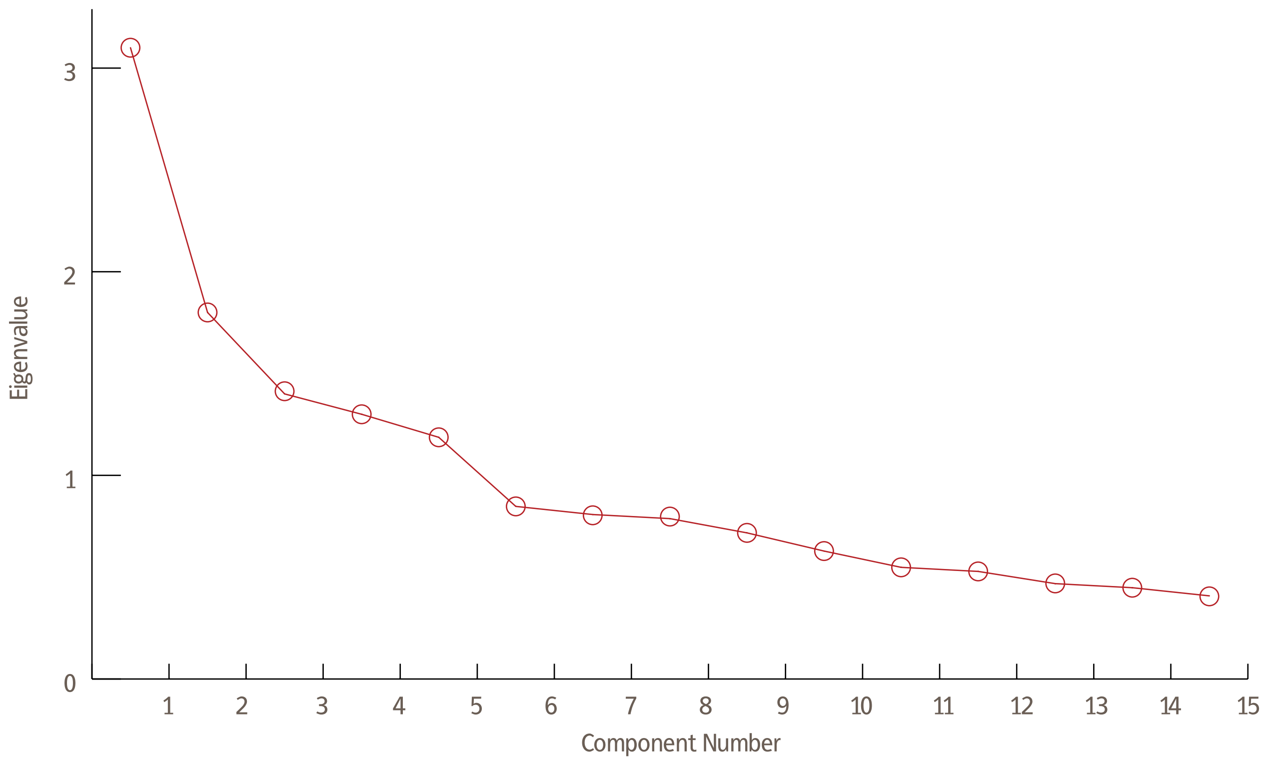
	Initial eigenvalues			Eigenvalues of extracted components			Eigenvalues of rotated extracted components		
Com-ponent	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.985	19.899	19.899	2.985	19.899	19.899	2.118	14.119	14.119
2	1.827	12.180	32.079	1.827	12.180	32.079	1.842	12.280	26.399
3	1.410	9.400	41.479	1.410	9.400	41.479	1.669	11.129	37.528
4	1.277	8.511	49.990	1.277	8.511	49.990	1.585	10.570	48.097
5	1.066	7.107	57.097	1.066	7.107	57.097	1.350	9.000	57.097
6	0.858	5.719	62.816						

Extraction method: Principal Component Analysis

Source: Authors' calculations using GCRO's Quality of Life Survey data (GCRO, 2009)

In Figure 1 below the scree plot shows an elbow at the fifth and the sixth extracted components, indicating that either five or six components should be extracted.

Figure 1: Scree plot indicating the quality of life components



The last guideline for the selection of the number of components to be extracted was based on the interpretability of the rotated components. The interpretability was determined by analysing the highest factor loadings of the indicators on each component. If these indicators with the highest factor loadings on a specific component were highly correlated and related to a common latent variable – which is consistent with the quality of life theory and the literature – then the extracted rotated components are interpretable. It was found that if five components were extracted, the components were interpretable, whereas if six components were extracted, the interpretability was obscured. According to the interpretability guideline it was concluded that five components should be extracted. Therefore, considering the Kaiser rule, the scree plot and the test of interpretability, it was decided that five components should be extracted.

The results of the eigenvalues of the extracted five components are shown in the second column and the eigenvalues of the rotated extracted components are shown in the third column of Table 5. The five extracted components explain more than 57% of the variation in the data set, which was an acceptable value of explained variance to be used in further analysis (for comparative studies see Vyas and Kumaranayake, 2006; Rossouw and Naudé, 2008; Naudé et al., 2009; Lewer et al., 2009; and Rossouw and Pacheco, 2012). In further analyses we used the results derived from the five extracted components with rotation.

The pattern and structure matrix of the five extracted components with varimax rotation using PCA is shown in Table 6 below. The table shows each indicator variable with its corresponding factor loadings on each component. The factor loadings are the correlation coefficients between the variables and the factors. The component was given a descriptive name according to the highest factor loadings on a component. The remainder of this section highlights those indicator variables which had the highest factor loadings on each component and the labels that were assigned. In addition, the percentage of variance explained by each component is considered.

Table 6: Pattern and structure matrix for PCA with varimax rotation

Variable	Factor loadings on components				
	1	2	3	4	5
Piped water on premises	**.837**	.025	.035	.025	.051
Satisfied with local government	.102	-.025	.149	-.066	**.694**
Satisfaction with life	.212	.460	**.504**	-.053	.252
Satisfied with time available for friends	.021	**.692**	–.033	.190	-.050
Work	**-.079**	-.047	.705	.047	.036
Electricity used for lighting	.831	.050	.069	.007	.026
Type of dwelling	**.768**	.085	.120	-.024	.029
Income category	.197	.101	**.678**	.096	-.078
Satisfaction with health	.008	.068	.195	**.793**	.095
How often does health prevent you from taking part in social activities	-.039	.096	.020	**.835**	.056
Feel safe at home	.044	.203	.028	.187	**.568**
Judiciary is free	-.051	-.064	-.210	.035	**.646**
Satisfied with time available for own things	.081	**.699**	.202	-.021	.172
Satisfied with time available for family	.029	**.758**	-.017	.046	-.042
Level of education	.153	.077	**.540**	**.407**	-.088
Percentage explained variance	14.12%	12.28%	11.13%	10.57%	9.00%

Note: Major factor loadings for each item are in bold

Source: Authors' calculations using GCRO's Quality of Life Survey data (GCRO, 2009)

Table 6 shows that type of dwelling (0.768), water on the premises (0.837) and electricity for lighting (0.831) have the highest factor loadings on the first component. This component was labelled housing and infrastructure. This dimension represents one of the basic needs of people and according to Maslow's (1943) hierarchy of needs, is one of the first needs people fulfil. This dimension explains the most variance in the data set, 14.12%. The values on the indicator variables measuring the component vary greatly among the included respondents and have an important role in explaining the different levels of quality of life. This finding was consistent with previous research done by Bookwalter et al. (2006), CASE (2006). Richards et al. (2007), Hinks and Gruen (2007), GCRO (2011) and Moller (2012).[11] These studies illustrate that the fulfilment of basic needs, such as living in a house and having access to basic services, are some of the biggest contributors to the quality of life and subjective wellbeing of the people living in South Africa.

In the second component, 'time with family' (0.758) and 'time with friends' (0.692) and 'sufficient time for leisure' (0.699) loaded the highest on the component and this component was labelled as 'social relationships'. Social relationships is also a need described by Maslow (1943), who explained that people need to be loved and cherished and feel part of a community. This dimension explained the second most variance in the data set (12.28%), and these indicator variables therefore also showed considerable variance between the respondents.

In the third component the variables with the highest factor loadings were 'income' (0.678), 'employment' (0.705), 'education' (0.540) and 'subjective wellbeing' (0.504). These variables are related to economic issues in which education plays an important role in both employment and earning higher levels of income. Furthermore, being satisfied with life is highly correlated with economic variables especially at lower levels of income. This component was described as 'socio-economic status' and explained the third most variance in the data set at 11.13%.

'How often does health prevent you from taking part in social activities' (0.835) and 'satisfaction with health' (0.793) were the two variables with the highest factor loadings on the fourth component and were labelled as 'health'. This component is related to the functionings of humans, as health is essential to fulfil the end goals of human life. This dimension explained 10.57% of the variance in the data.

In the fifth component the variables with the highest factor loadings were 'feel safe at home' (0.568), 'satisfaction with local government' (0.694) and 'freedom of the judiciary' (0.646); these indicators reflect functionings of a higher order. 'Feel safe at home' clusters with 'satisfaction with local government' and 'freedom of the judiciary'. This could possibly be explained by the high correlation and close relationships between safety, governance and the judiciary system. People are safer in an environment in which the government and the judiciary system function well and laws are enforced to prevent crime. This component was labelled 'governance and safety'. This dimension explained 9% of the variance in the data.

To summarise, the patterns in the data revealed by the PCA indicated that there were five components or dimensions of quality of life captured in the selected data set. Each component explained progressively less of the variance in the data set. Furthermore, the PCA analysis showed that the dimensions revealed in the data corresponded to those described in other studies of quality of life, and were included in the construction of quality of life composite indices. The finding that 'housing and infrastructure' explained the most variance in the data set accords with previous research done in South Africa. The results of the PCA were supported by the reviewed theory and literature on quality of life, which underscored the robustness of the selection of the indicator variables and the decision to extract five components.

[11] See also other works of Moller and co-authors such as Moller and Schlemmer (1983), Moller and Schlemmer (1989), and Moller and Dickhow (2002).

5. The construction of a Quality of Life index for the GCR

In Section 3.1 the steps for the construction of a composite index were explained. In this section we proceeded with the construction of the composite index by following the steps explained in Section 3.1, using the selected indicator variables.

5.1 The construction of the index

The method introduced by Nicoletti et al. (2000) was used to weight and aggregate the composite index of quality of life. This method differs from other standard methods found in the literature to weight composite indices using PCA as it does not only consider the first principal component to weight the index, but also the factor loadings of the consecutive extracted components. The benefit of this method is that a bigger proportion of the variance in the data set is explained.

According to this method, the individual indicators are grouped into intermediate composite indices based on the highest factor loadings of the indicators on a specific component. Table 7 again shows the results of the PCA (also illustrated in Table 6) as well as further analysis (see the explanation below the Table 7) of the data with five extracted components. Based on these results, five intermediate composite indices were constructed.

Table 7: Factor loadings based on PCA

	Factor loadings on components					Squared factor loadings (scaled to unity sum)				
	1	2	3	4	5	1	2	3	4	5
Piped water on premises	.837	.025	.035	.025	.051	.350	.000	.000	.000	.000
Satisfied with local government	.102	–.025	.149	–.066	.694	.000	.000	.000	.000	.390
Satisfaction with life	.212	.460	.504	–.053	.252	.000	.000	.170	.000	.000
Satisfied with time available for friends	.021	.692	–.033	.190	–.050	.000	.310	.000	.000	.000
Work	–.079	–.047	.705	.047	.036	.000	.000	.330	.000	.000
Electricity used for lighting	.831	.050	.069	.007	.026	.350	.000	.000	.000	.000
Type of dwelling	.768	.085	.120	–.024	.029	.300	.000	.000	.000	.000
Income category	.197	.101	.678	.096	–.078	.000	.000	.310	.000	.000
Satisfaction with health	.008	.068	.195	.793	.095	.000	.000	.000	.470	.000
How often does health prevent you from taking part in social activities	–.039	.096	.020	.835	.056	.000	.000	.000	.530	.000
Feel safe at home	.044	.203	.028	.187	.568	.000	.000	.000	.000	.260
Judiciary is free	–.051	–.064	–.210	.035	.646	.000	.000	.000	.000	.340

	Factor loadings on components					Squared factor loadings (scaled to unity sum)				
	1	2	3	4	5	1	2	3	4	5
Satisfied with time available for own things	.081	.699	.202	–.021	.172	.000	.320	.000	.000	.000
Satisfied with time available for family	.029	.758	–.017	.046	–.042	.000	.370	.000	.000	.000
Level of education	.153	.077	.540	.407	–.088	.000	.000	.190	.000	.000
Explained variance (Eigenvalue)	2.118	1.842	1.669	1.585	1.350					
Total percentage of explained variance	.247	.215	.195	.185	.158					

Note: Factor loadings in bold indicate the highest factor loadings on a specific component

Note: The first five columns of Table 7 repeat the information of Table 6

Source: Authors' calculations using GCRO's Quality of Life Survey data (GCRO, 2009)

The same names were given to the constructed intermediate composite indices than were used in previous sections to label the extracted components. In the order of explaining the most variance in the data set the intermediate composite indices were 'housing and infrastructure', 'social relationship', 'socio-economic status', 'health' and 'safety and governance'. To construct the intermediate composite index 'housing and infrastructure', the variables with the highest factor loadings on this component, namely 'piped water on the premises' (0.837), 'electricity used for lighting' (0.815) and 'type of dwelling' (0.768) were included. The weighting of each of the variables was derived by squaring the factor loadings of the variables. The squared factor loadings represented the proportion of the total unit variance of the indicator, which was explained by the component. Furthermore, the weights were scaled to unity sum. Therefore, the weighting of 'piped water on the premises' was 0.35, 'electricity used for lighting' was also 0.35 and 'type of dwelling' was 0.30 in the intermediate composite index labelled 'housing and infrastructure'. In the same manner, the weights of the other variables were derived and included in the successive four intermediate composite indices.

Once the five intermediate composite indices had been constructed, they were aggregated by allocating a weight to each one of them equal to the proportion of the explained variance in the data set. For example, the weighting of the first intermediate composite index was 0.247 (24, 7%), calculated as follows:

$0.247 = (2.118 / (2.118+1.842+1.669+1.585+1.350))$

In the same manner the weights of each intermediate composite index in the total composite index were calculated (see Table 7 for the weighting of the intermediate composite index expressed as the total percentage of explained variance of each component). Note that the weighting of each consecutive intermediate composite index contributed less to explaining the variance in the data set, decreasing from 24.7% to 15.8%.

In summary, the method proposed by Nicoletti et al. (2000) was used to construct a composite index of quality of life which was objectively weighted according to the explained variance in the data set.

5.2 Validation of the robustness of the composite index of quality of life

Two methods were used to test the robustness of the newly created composite index of quality of life. The first was to employ different methods to treat the missing data. The second method was to correlate the values of the composite index to traditionally used single-dimensional measures of quality of life such as income or subjective wellbeing.

Using the first method, three different PCAs were run, using listwise and pairwise[12] deletion as well as single imputation based on the means of the variables. This was followed by a comparison of the derived rotated factor matrices and an inspection of the matrices for meaningful differences in the factor loadings of the variables. If meaningful difference were found it could indicate that the selection of the method to treat the missing data had an effect on the constructed composite index.

No meaningful differences were found between the results of the rotated matrices using different methods of treating missing data in the indicator variables (see Appendix B). Therefore, based on these results it was concluded that the newly constructed composite index was robust.

Using the second method the values of the composite index were correlated to income and subjective wellbeing using the income variable and the life satisfaction variable included in the GCRO Quality of Life Survey (GCRO, 2009).[13] The quality of life score positively and strongly correlated with income and subjective wellbeing (see Table 8), with correlation scores of 0.553 (satisfaction with life) and 0.659 (income). Based on these findings it was concluded that the newly created quality of life index was robust. Furthermore, although single-dimensional measures such as income and life satisfaction might possibly reflect the trend in quality of life, a multi-dimensional measure presents a more holistic view of the experienced quality of life of people.

Table 8: Correlation matrix

	QoL score	Satisfaction with life	Income
QoL score	1		
Satisfaction with life	.553***	1	
Income	.659***	.301***	1

Note: QoL = Quality of Life

*****Significant at the 1% level.**

Source: Authors' calculations using the GCRO's Quality of Life Survey data (GCRO, 2009)

Based on these findings it was concluded that the newly developed quality of life measure was robust and adequate for use in further analysis.

12 If listwise treatment of missing variables is used any missing value for any variable is omitted. If pairwise treatment of a case is used the respondent is only left out of the analysis for those variables that have missing values.

13 It should be acknowledged that comparing the newly constructed composite index to the income variable and the life satisfaction variable can raise the question of endogeneity. However, the lack of data leaves no other option and the results are therefore interpreted with care.

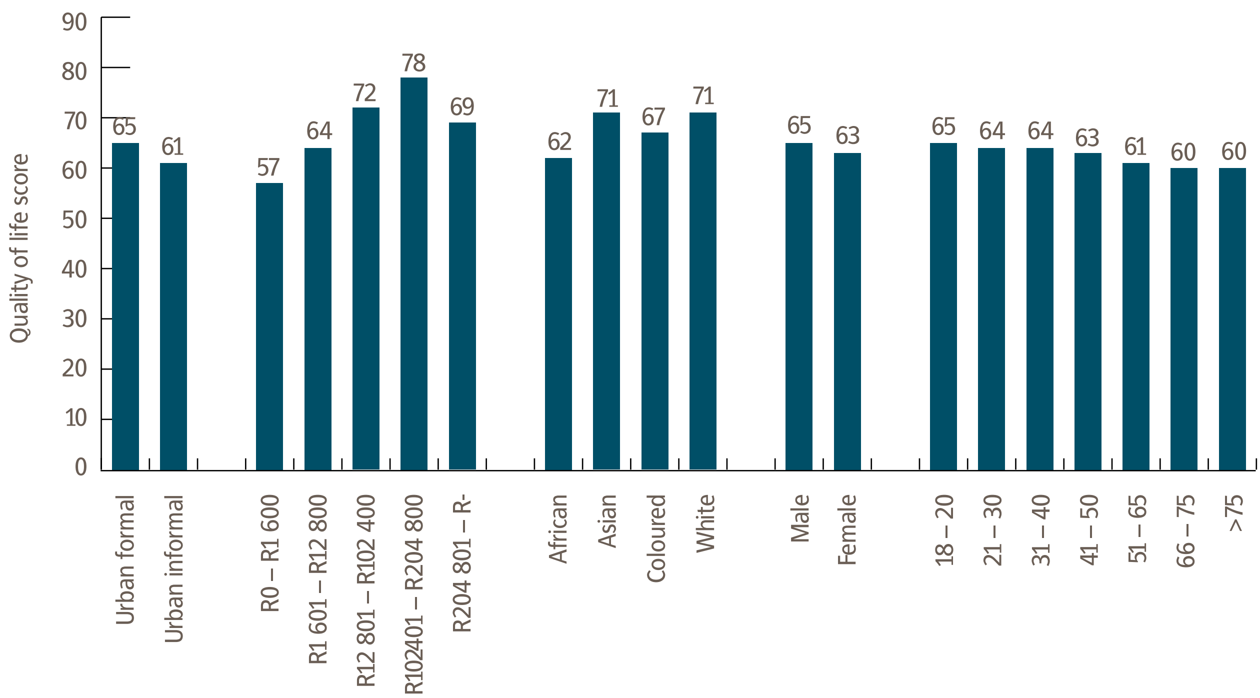
6. Quality of life scores of different demographic and socio-economic groups

This section explains how the newly constructed quality of life index was used to calculate and compare the quality of life scores of different demographic and socio-economic groups within the GCR.

In Figure 2 the calculated quality of life scores for the different groups are shown as a percentage, with 100% the maximum score. To attain a 100% score the respondent had to score the highest value for each indicator variable within each dimension of quality of life.

Comparing the life scores illustrated that the groups with the lowest life scores were respondents living in informal urban areas, respondents in a household earning less than R1 601 per month, African repondents, and people over the age of 65. This, naturally, did not imply that all the repondents within these groups had low quality of life scores, but rather that within these groups there were more individuals with lower levels of quality of life than within other groups.

Figure 2: Quality of life scores of different demographic and socio-economic groups



Source: Authors' calculations using the GCRO's Quality of Life Survey data (GCRO, 2009)

In the first instance the quality of life scores of respondents living in different geographical types were analysed and compared. Figure 2 shows that the quality of life of urban informal residents is lower than that of urban formal residents. This finding is supported by other literature, for example Klasen (2000), Higgs (2007), and Rossouw and Naudé (2008).

Based on the GCRO data set[14] the differences between urban informal and urban formal areas can perhaps be explained by the lower income levels reported by respondents in urban informal areas (R2 400 compared to R4 800 per month), lower employment rates (35% compared to 46%) and less access to formal housing and infrastructure.

In the urban formal areas almost all the respondents resided in formal dwellings, had electricity for lighting, and water on their premises. This compares with only 80% of the urban informal respondents that resided in formal dwellings, 10% did not have water on their premises and 12% did not have electricity for lighting.

Second, the findings were viewed in light of the differences in income between groups. The lowest quality of life score calculated based on income differences was that of households earning up to R1 600 per month. According to these calculations quality of life increases as income increases up to a level of R204 801 per household per month, at which point the quality of life score starts to decline. This finding corresponds with the findings of Easterlin (1974) based on subjective wellbeing where it was shown that as income increases subjective wellbeing increases up to a certain point, after which it starts to decline. It is notable that this paper's findings on quality of life, a multi-dimensional measure, confirmed Easterlin's findings on subjective wellbeing, a single-dimensional measure. The results presented in this paper are in line with the findings in the broader literature that subjective wellbeing might be a better proxy for quality of life than income. Secondly, the results of this paper show that the relationship between wellbeing and income in developing regions such as the GCR might follow a similar pattern as in developed countries.

Conceivable reasons for the lower quality of life of high income earners, reported by respondents in the GCR, can possibly be found in lower levels of 'satisfaction with time spent with family and friends', 'satisfaction with leisure time' and lower levels of overall 'life satisfaction'.

The lower levels of quality of life among lower income earners can partially be explained by a lack of employment. The main reason mentioned by the respondents for not being employed was a shortage of jobs. There was also a marked difference in the education level of the lower and higher income groups. Only 4% of the lowest income groups had tertiary training compared to almost 62% among the income group earning more than R51 000 per month (GCRO, 2009).

Comparing dwelling types and infrastructure, almost all the higher income group respondents lived in formal housing, had electricity for lighting and had water on the premises. In the income group earning less than R800, almost 30% lived in informal housing, 16% did not have water on their premises and approximately 4% did not have electricity for lighting.

Only 20% of the lowest income groups reported being satisfied or very satisfied with life. Those not satisfied with life mentioned a lack of income, high costs of living and a shortage of employment opportunities as the main reasons for being dissatisfied. In comparison, 75% of the wealthier respondents reported being either satisfied or very satisfied with life (GCRO, 2009).

Third, the findings were reviewed based on comparisons between different race groups. The results showed that Whites and Asians have higher levels of quality of life than Africans and Coloureds. This finding was expected based on South Africa's apartheid history. This result was furthermore supported by other studies (Klasen, 2000; Higgs, 2007; Posel & Casale, 2011; GCRO, 2011; Moller, 2012 and SASAS by the HSRC, 2013).

[14] All the statistics cited in this discussion are based on the authors' analyses of the GCRO's Quality of Life Survey data.

Analysis of the GCRO data showed that the income variable, the life satisfaction variable and the variables related to social relationships differed markedly between African and White respondents. The average income of African households is approximately R2 200 per month compared to R9 500 earned by Whites (GCRO, 2009). Furthermore, only 60% of the African respondents reported to be employed compared to approximately 90% of the White respondents. Many of the employed African respondents reported being employed in low-skilled jobs such as working in private households. The other race groups reported being employed in higher-skilled jobs.

The life satisfaction variable differed markedly between the race groups (GCRO, 2009). Almost 84% of Asians and Whites reported to be either satisfied or very satisfied with life compared to only 35% of Africans and 50% of Coloureds. Higgs (2007), Posel and Casale (2011), GCRO (2011) and Moller (2012) came to the same conclusions. Analysis of the GCRO data suggests that the main reasons mentioned by respondents across all race groups for not being satisfied with life were economic in nature.

The 'social relationships' variables of Africans were on average higher than that of Whites and Asians, implying that the former were more satisfied with the time they had available to spend with their family and friends, as well as for leisure.

Fourth, the quality of life scores of men and women were compared. The quality of life score of men was slightly higher than that of women. Analysis of the GCRO data illustrated that the variables 'employment' and 'income' were very different between the two groups. These results were similar to what Klasen (2000) and GCRO (2011) found. Furthermore, the GCRO data showed that men are 20% more likely to have worked in the past seven days than women. An explanation for this could be the lower levels of female education reported in the GCRO survey as well as reasons listed by female respondents for not finding employment such as "they have to look after children or family members" and that "they live too far away" from employment opportunities (GCRO, 2009). Adding to this could be gender segmentation and possible discrimination in the labour market. The average income earned by women is also slightly less than the average income earned by men (GCRO, 2009).

Lastly, the quality of life scores based on age differences were analysed. There was very little difference between the quality of life scores of different age groups, though the quality of life score of younger respondents was slightly higher than that of middle-aged and older respondents. This finding was similar to the findings of the GCRO (2011). The higher quality of life scores of younger respondents based on the GCRO data might be explained by their higher levels of education and being healthier than older respondents; 40% of the respondents between the ages of 18 and 35 indicated that they had matriculated compared to only 14% in the age group over 40. Even though younger respondents had higher levels of education and self-reported health, their level of employment was still relatively low (40%) compared to respondents between the ages of 36 to 45 (58%). The main reason stated by the young respondents for being unemployed was that there were 'no jobs' available (GCRO, 2009).

There was a negative relationship between the quality of life score of the different age groups and income, thus although income increases with age up to a turning point at 65, the quality of life score decreases. Reasons for the decrease in quality of life scores might possibly be ascribed to weaker health of respondents over the age of 40 (GCRO, 2009). Furthermore, from the age of 35 respondents indicated that they had less time to spend with friends and family and for leisure. Middle-aged and older respondents also seemed to be less satisfied with the performance of the local government.

To summarise, the results showed that the demographic and socio-economic groups with the lowest quality of life scores were respondents residing in informal urban areas, African respondents, respondents with low inome levels, older respondents and women. The groupings with higher quality of life scores were Whites, Asians, younger respondents, higher income earners and urban dwellers.

7. The components that explain the most variance in the data set for different groupings

The selected indicator variables were used and separate PCAs run for different demographic and socio-economic groups to determine which dimensions of quality of life explain the most variance in the data set for each group. This analysis was based on the descriptive label given to the first extracted components of each group. The first three extracted components for each group are shown in Table 9 below.[15]

Table 9: The first three components extracted for each grouping

Group demarcations	Components		
	1	2	3
Race			
Africans	Housing and infrastructure	Social relationships	Health
Coloureds	Social relationships	Housing and infrastructure	SES*
Asians and Whites	Social relationships	SES	Safety and governance
Income			
R0 – R800	Housing and infrastructure	Social relationships	Health
R801 – R6 400	Housing and infrastructure	Social relationships	SES
R 6 401 – R51 200	Housing and infrastructure	Social relationships	SES
R51 201 – higher	Social relationships	Housing and infrastructure	SES
Age			
18-35	Housing and infrastructure	Social relationships	SES
36-48	Housing and infrastructure	SES	Social relationships
48-65+	Housing and infrastructure	SES	Social relationships
Gender			
Male	Housing and infrastructure	Social relationships	SES
Female	Housing and infrastructure	Social relationships	Health
Geographical type			
Urban formal	Social relationships	SES	Health
Urban informal	Housing and infrastructure	Social relationships	Health

***SES: Socio-economic status**

Source: Authors' own analysis of GCRO data

15 As the objective was to compare the components that explain the most variance in the data set only the first three extracted components were compared.

For the majority of the groups, the biggest variance in the data was explained by the 'housing and infrastructure' component (see Table 9). This implied that the indicator variables 'type of dwelling', 'piped water on the premises' and 'electricity for lighting' vary considerably within these groups. Therefore, it could be argued that improved 'housing and infrastructure', which is a basic need, could improve the quality of life of these groups.

The dimension 'social relationships' explained the most variance in the data set for the following groups: Coloureds, Asians, Whites, respondents with a monthly income of more than R51 201 and urban formal dwellers. This implied that the indicator variables 'satisfaction with time to spend with friends', 'satisfaction with time to spend with family' and 'satisfaction with time available for own things' varied the most within these groups. It could therefore be argued that in order to improve the wellbeing of people in these groups the availability of time to spent with friends, family and for leisure should be addressed. Furthermore, as 'social relationships' explained most of the variance in the data set and the 'housing and infrastructure' dimension was more consistent amongst these respondents (which meant that the majority of them live in formal housing, and have access to electricity and water), it is possible that their basic needs had been met.

A component that ranked within the top three positions in most of the groups was 'socio-economic status'. For the majority of groups this component explained a considerable proportion of the variance. Therefore, this dimension may have had a notable effect on the quality of life of all demographic and socio-economic groups, and varied considerably between all respondents.

Based on the findings of Section 7 on the component which explained the largest proportion of the variance in the data set and the calculated quality of life scores of the different demographic and socio-economic groups (Section 6), it was found that the groups in which 'social relationships' explained the most variance in the data set largely corresponded with the groups with the highest quality of life scores. It can therefore be argued that to improve the wellbeing of people that experience relatively higher levels of quality of life the 'social relationship' component should be addressed, even though this component is to a significant extent beyond the reach of public policy interventions.

The groups experiencing lower levels of quality of life correponded with those groups in which 'housing and infrastructure' explained the most variance in the data set. A case could therefore be made that in order to improve the wellbeing of people experiencing lower levels of quality of life the component 'housing and infrastructure' should be addressed. This component included indicator variables such as 'housing type', 'electricity for lighting' and 'water on the premises'; although it is also possible that addressing any indicator variable closely related to these variables, such as 'sanitation' and 'electricity for cooking', could also improve quality of life.

Furthermore, as 'socio-economic status' explained a considerable portion of the variance in the data set of all the groups and was highlighted by the majority of groups as a concern which influences their quality of life, addressing this component could contribute to the wellbeing of all people in the GCR.

8. Conclusion

Quality of life is a multi-dimensional concept that needs to be measured by a composite index that is able to assess the quality of life in a region. Only once quality of life is measured can trends in quality of life be monitored and analysed in order to direct policy decisions.

This paper contributed to the literature on the measurement of quality of life by constructing an objectively weighted composite index of quality of life and applying the index to compare quality of life across the diverse population of the GCR. Furthermore, the paper also determined which components explained the most variance in the data set of the different groupings.

A novel method developed by Nicoletti et al. (2000) was used to construct the composite index of quality of life, which incorporated not only the first principal component to weight the index, but also additional components to achieve a better representation of the data. The constructed index included the relevant dimensions of quality of life for the region and included both objective and subjective indicator variables as well as economic and non-economic variables. This was the first measure of this type constructed for the GCR and used in South Africa.

Data from the 2009 Quality of Life Survey conducted by GCRO were used and the initial selection of the quality of life index measures were based on the capabilities and functionings and subjective wellbeing theoretical approaches. Through PCA a set of fifteen indicator variables was identified from the initially selected indicator variables. These indicator variables represented five dimensions or components of quality of life in the GCR and were labelled as: 'housing and infrastructure', 'social relationships', 'socio-economic status', 'health' and 'safety and governance'. Each of the consecutive extracted components or dimensions explained less of the variance in the data set.

Corresponding to the extracted components five intermediate composite indices were constructed. The intermediate composite indices were aggregated by weighting each one according to the percentage of variance explained by the component. The index was validated and found to be robust.

Using the newly constructed composite index the quality of life of the different demographic and socio-economic groups in the GCR was assessed. It was found that the quality of life scores of the following categories of demographic and socio-economic groups were relatively lower than the others: urban informal, low income, female, African and older respondents. The groupings with higher quality of life included Asians, Whites, high income earners, younger respondents and urban dwellers.

To address the secondary research objective a PCA was run for each different demographic and socio-economic group to determine which component explained the most variance in the data set for each group. It was found that the first extracted component for the different groupings varied. For African repondents, non-urban dwellers, the majority of age groups, all genders and low income earners, the 'housing and infrastrcuture' component, which represents a basic need, explained the most variance in the data set.

For White, Coloured and Asian respondents, as well as for high income earners and repondents in urban areas, the dimension described as 'social relationships' which is a higher order need, explained the most variance in the data set. Furthermore, it was found that the dimension labelled 'socio-economic status' represented a considerable proportion of the explained variance of all the groupings.

When the results of both the primary and the secondary research questions are considered, it is clear that the groups which attained lower quality of life scores largely coincided with the groups in which the dimension 'housing and infrastructure' explained the most variance in the data set. Therefore, it might follow that to improve the quality of life of these groups the indicator variables 'type of dwelling', 'electricity for lighting' and 'water on the premises' as well as the indicator variables closely associated with these variables such as 'sanitation' and 'electricity for cooking' should be addressed.

The groups with higher quality of life scores largely corresponded with the groups in which it was found that 'social relationships' explained the most variance in the data set. Therefore it might follow that to increase the quality of life of these groups the indicator variables 'time with family and with friends' as well as 'leisure time' should be addressed insofar as possible. Within all the groups the dimension 'socio-economic status' explained a considerable proportion of the variance in the data set. This dimension was also highlighted as contributing to lower levels of quality of life. Therefore, we could argue that improved 'socio-economic status' can directly improve the wellbeing of all groups of people in the GCR.

This paper analysed and measured quality of life and identified the demographic and socio-economic groups and the dimensions of quality of life which should be addressed and prioritised to improve the quality of life of all people in the GCR.

Acknowledgements We would like to thank the Gauteng City-Region Observatory for their financial support and for access to the GCRO Quality of Life Survey data set.

Appendix A

Recoding of variables

The recoding of the nominal variables was based on the guidelines provided in the Reconstruction and Development Programme (1996)[16] as well as the cut-off points used in the poverty index compiled by the GCRO in the Gauteng City-Region Review (GCRO, 2011).

Table A: Coding of variables

Variable	Type of variable	Coding	Description
Type of dwelling	Dichotomous	0	Informal dwelling
		0	Traditional dwelling
		1	House or formal structure
		1	Flat
		1	Town/cluster/semi-detached
		1	Unit in retirement village
		1	Room/flatlet in main dwelling
		1	Hostel
Piped water on premises	Dichotomous	0	Water not piped and not on premises
		1	Piped – in dwelling
		1	Piped – yard tap
Electricity used for lighting	Dichotomous	0	Gas/LPG
		0	Paraffin
		0	Wood
		0	Candles
		0	Solar energy
		1	Electricity

[16] After the democratic elections in South Africa the government implemented the Reconstruction and Development Programme to address the problems of poverty and inequality.

Variable	Type of variable	Coding	Description
Type of sanitation	Dichotomous	0	Septic tank
		0	Pit latrine
		0	Chemical toilet
		0	Communal toilet
		0	Neighbours
		0	Bush
		0	Bucket
		0	No toilet
		1	Full waterborne (flush toilet)
Satisfaction with dwelling	Dichotomous	0	Neither satisfied nor dissatisfied
		0	Dissatisfied
		0	Very dissatisfied
		1	Very satisfied
		1	Satisfied
Education	Ordinal variable	1	0 years
		2	1-4 years
		3	5-8 years
		4	9-12 years
		5	Tertiary training

Appendix B

Robustness test of the composite index

The robustness of the newly constructed composite index was tested by running three separate PCAs, using different methods to treat the missing variables. The first method used was listwise deletion (see Table B.1), the second was pairwise deletion (see Table B.2) and the third was imputing the missing variables by making use of single imputation (see Table B.3). The rotated matrices of each of the PCAs were compared to determine if there were meaningful differences between the results of the matrices. No meaningful differences were found and, based on the results shown in Table B.1, Table B.2, and TableB.3, the conclusion was drawn that the newly constructed composite index was robust.

Table B.1: Rotated component matrix: using listwise deletion of missing variables

	Component				
	1	2	3	4	5
Satisfaction with local government	.102	–.021	.145	–.064	.692
Satisfaction with life	.212	.453	.508	–.053	.252
Satisfaction with time available for friends	.028	.702	–.033	.189	–.054
Satisfaction with time available for own friends	.028	.783	–.013	.054	–.036
Work	–.081	–.040	.699	.047	.036
Electricity used for lighting	.832	.051	.062	.006	.032
Type of dwelling	.764	.084	.122	–.026	.028
Income category	.194	.087	.685	.094	–.075
How often does health prevent you from taking part in social activities	–.041	.097	.019	.836	.055
Satisfaction with health	.008	.068	.196	.792	.095
Feel safe at home	.045	.198	.033	.188	.569
Satisfied with time available for own things	.076	.708	.202	–.026	.170
Judiciary is free	–.050	–.063	–.206	.030	.649
Piped water on premises	.843	.026	.038	.025	.049
Level of education	.150	.083	.541	.404	–.090

Source: Authors' calculations using GCRO's Quality of Life Survey data (GCRO, 2009)

Table B.2: Rotated component matrix: using pairwise deletion of missing variables

	Component				
	1	2	3	4	5
Satisfaction with local government	.101	–.024	.148	–.067	.693
Satisfaction with life	.212	.452	.509	–.053	.254
Satisfaction with time available for friends	.023	.702	–.034	.187	–.053
Satisfaction with time available for own friends	.025	.784	–.012	.052	–.036
Work	–.079	–.039	.699	.048	.032
Electricity used for lighting	.833	.049	.068	.007	.026
Type of dwelling	.767	.083	.122	–.024	.029
Income category	.196	.086	.684	.097	–.074
How often does health prevent you from taking part in social activities	–.039	.098	.020	.835	.056
Satisfaction with health	.008	.066	.195	.793	.096
Feel safe at home	.044	.197	.030	.186	.570
Satisfied with time available for own things	.082	.704	.203	–.023	.171
Judiciary is free	–.051	–.062	–.211	.035	.645
Piped water on premises	.841	.026	.034	.025	.051
Level of education	.152	.082	.538	.407	–.089

Source: Authors' calculations using GCRO's Quality of Life Survey data (GCRO, 2009)

Table B.3: Rotated component matrix: using the mean to impute missing variables

	Component				
	1	2	3	4	5
Satisfaction with local government	.101	–.024	.148	–.067	.693
Satisfaction with life	.213	.452	.508	–.053	.254
Satisfaction with time available for friends	.023	.701	–.034	.188	–.053
Satisfaction with time available for own friends	.025	.784	–.012	.052	–.036
Work	–.080	–.039	.700	.047	.032
Electricity used for lighting	.831	.049	.069	.007	.026
Type of dwelling	.768	.083	.121	–.024	.029
Income category	.197	.086	.683	.097	–.074
How often does health prevent you from taking part in social activities	–.040	.098	.019	.835	.056
Satisfaction with health	.008	.066	.195	.793	.096
Feel safe at home	.045	.197	.030	.186	.570
Satisfied with time available for own things	.082	.704	.203	–.023	.171
Judiciary is free	–.051	–.062	–.211	.035	.645
Piped water on premises	.837	.026	.034	.025	.051
Level of education	.153	.082	.538	.407	–.090

Source: Authors' calculations using GCRO's Quality of Life Survey data (GCRO, 2009)

Reference list

- Aivazian, S., 2005. Synthetic Indicators of Quality of Life: Construction and Utilisation for Social-Economic Management and Comparative Analysis. *Austrian Journal of Statisics*, 34(2), pp. 69-77.
- Bandura, R., 2006. *A Survey of Composite Indices Measuring Country Performance; 2006 Update, United Nations Development Programme – Office of Development Studies.* [Online] Available at: http://web.undp.org/developmentstudies/docs/indices_2008_bandura.pdf [Accessed 10 November 2012].
- Bookwalter, J. T., Fuller, B. S. & Dalenberg, D. R., 2006. Do Household Heads Speak for the Household? A research note. *Social Indicators Research,* Volume 79, p. 405–419.
- Booysen, F., 2002. An Overview and Evaluation of Composite Indices of Development. *Social Indicators Research,* Issue 59, pp. 115-151.
- Brooks, C., 2008. *Introductory Econometrics for Finance.* Second ed. Cambridge: Cambridge University Press.
- Canadian Index of Well-being, 2009. *Canadian Index of Well-being.* [Online] Available at: http://www.ciw.ca/en/Home.aspx[Accessed 19 August 2011].
- Cherchye, L., Moesen, W., Rogge, N. & Van Puyenbroeck, T., 2007. An Introduction to 'Benefit of the doubt' Composite Indicators. *Social Indicators Research,* Issue 82, pp. 111-145.
- Community Agency for Social Enquiry (CASE), 2006. *A quality of life survey in low income areas in the City of Johannesburg,* Johannesburg: CASE.
- Conceicao, P. & Bandura, R., 2008. *Measuring Subjective Wellbeing: A summary review of the literature,* New York: Office of Development Studies, UNDP.
- Costanza, R; Fisher, B; Ali, S; Beer, C; Lynne, B; Boumans. An Integrative Approach to Quality of Life Measurement, Research, and Policy. *SAPIENS Journal,* 1(1), pp. 17-21.
- Cummins, R. A., 2000. Objective and Subjective Quality of Life: an Interactive Model. *Social Indicators Research,* 52(1), pp. 55-72.
- Diener, E. & Suh, E., 1997. Measuring Quality of Life: Economics, Social, and Subjective Indicators. *Social Indicators Research,* 40(1-2), pp. 189-216.
- Dluhy, M. & Swartz, N., 2006. Connecting Knowledge and Policy: The Promise of Community Indicators in the United States. *Social Indicators Research,* Volume 79, pp. 1-23.
- Easterlin, R. A., 1974. Does Economic Growth Improve the Human Lot?. In: David, P.A. & Reder, M.W., ed. *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz.* New York: Academic Press, Inc.
- European Commission, 2007. *Beyond GDP.* [Online]. Available at: http://www.beyond-gdp.eu/indicatorList.html?indicator=Well-being. [Accessed 30 July 2012].
- Gauteng City-Region Observatory (GCRO), 2009. *GCRO Quality of Life Survey (dataset),* Johannesburg: Global Print.
- Gauteng City-Region Observatory (GCRO), 2011. *The City-Region Review,* Johannesburg: Gauteng City-Region Observatory.

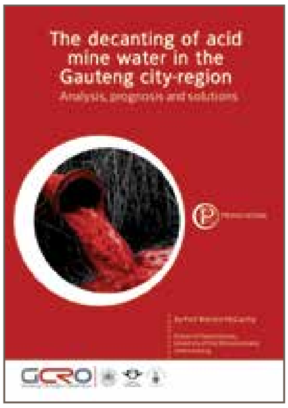
- Gauteng Provincial Government, 2010. *Socio-economic Review and Outlook.* [Online] Available at: http://www.gautengonline.gov.za/Publications%20and%20Reports/Socio_Economic_Review_and_Outlook_2010.pdf. [Accessed 7 October 2012].
- Hagerty, M. R. & Land, K. C., 2007. Constructing summary indices of quality of life, a model for the effect of heterogeneous importance weights. *Sociological Methods and Research,* Volume 35(40), pp. 455-496.
- Higgs, T., 2007. Measuring and Understanding the Wellbeing of South Africans: Everyday Quality of Life in South Africa. *Social Indicators Research,* Volume 81, pp. 331-356.
- Hinks, T. & Gruen, C., 2007. What is the Structure of South African Happiness Equations? Evidence from Quality of Life Surveys. *Social Indicators Research,* 82(2), pp. 311-336.
- Human Sciences Research Council, 2013. *South African Social Attitudes Survey (SASAS).* [Online] Available at: http://www.hsrc.ac.za/en/departments/sasas [Accessed 10 August 2012].
- Kaiser , H. F. & Rice, J., 1974. Little Jiffy, Mark IV. *Educational and Psychological Measurement,* Volume 34, pp. 111-117.
- Kingdon, G. G. & Knight, J., 2003. Well-being poverty versus income poverty and capabilities poverty. *Centre for the Studies of African Economies: Working Paper 2003-16.*
- Klasen, S., 2000. Measuring Poverty and Deprivation in South Africa. *Review of Income and Wealth,* Series 46, Number 1(March), pp. 33-58.
- Lewer, J. J., Pacheco, G. & Rossouw, S., 2009. Do Non-Economic Quality of Life Factors Drive Immigration?. *The Institute for the Study of Labor (IZA): Discussion Paper No. 4385.*
- Maslow, A. H., 1943. A Theory of Human Motivation. *Psychological Review,* 50(4), pp. 370-396.
- McGillivray, M., 2005. Measuring Non-economic Well-being Achievement. *Review of Income and Wealth,* 51(2), pp. 337-364.
- McGranahan, D. V., Richard-Proust, C., Sovani, N. V. & Subramanian, M., 1972. *A Staff Study of the United Nations Research Institute for Social Development (UNRISD).* New York: Praeger Publisher.
- Moller, V., 2012. South African Quality fo Life Trends Over Three Decades, 1980-2010. *Social Indicators Research,* Volume published online.
- Moller, V. & Dickhow, H., 2002. The Role of Quality of Life Surveys in Managing Change in Democratic Transitions: The South African Case. *Social Indicators Research,* 58(1/3), pp. 267-292.
- Moller, V. & Schlemmer, L., 1983. Quality of Life in South Africa: Towards an instrument for the Assessment of Quality of Life and Basic Needs. *Social Indicators Research,* Volume 12, pp. 225-279.
- Moller, V. & Schlemmer, L., 1989. South African Quality of Life: A Research Note. *Social Indicators Research,* Volume 21, pp. 279-291.
- Morris, D., 1979. *Measuring the Conditions of the World Poor, the Physical Quality of Life Index.* New York: Pergamon Press.
- Nardo, M., Saisana, M., Saltelli, A. & Tarantola, S., 2005. *Tools for Composite Indicator Building.* European Commission: Joint Research Centre.
- Naude, W., Rossouw, S. & Krugell, W., 2009. The Non-monetary Quality of City Life in South Africa. *Habitat International,* Volume 33, pp. 310-326.

- Nicoletti, G., Scarpetta, S. & Boylaud, O., 2000. *Summary indicators of product and market regulation with an extension to employment protection legislation.* [Online].Available at: ww.oecd.org/eco/eco [Accessed 4 October 2011].
- Organisation for Economic Co-operation and Development (OECD), 2008. *Handbook on constructing composite indicators: Methodology and user guide.* Paris: OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD), 2011. *Your Better Life Index.* [Online]. Available at: http://www.oecdbetterlifeindex.org [Accessed 19 August 2011].
- Office for National Statistics of the United Kingdom, 2012. *Office for National Statistics.* [Online] Available at: http://www.ons.gov.uk/ons/guide-method/user-guidance/well-being/index.html [Accessed 8 August 2012].
- Posel, D. R. & Casale, D. M., 2011. Relative Standing and Subjective Well-being in South Africa: The role of Perceptions, Expectations and Income Mobility. *Social Indicators Research,* Volume 104, pp. 195-223.
- Powdthavee, N., 2003. Is the structure of happiness equations the same in poor and rich countries: The case of South Africa. *Warwick Economic Research Papers.*
- Ram, R., 1982. Composite Indices of Physical Quality of Life, Basic Needs Fulfilment, and Income. A Principal Component Representation. *Journal of Development Economics,* Volume 11, pp. 227-247.
- Reconstruction and Development Programme, 1996. *African National Congress.* [Online] Available at: http://www.anc.org.za/show.php?id=234, [Accessed 20 March 2012].
- Richards, R., O'Leary, B. & Mutsonziwa, K., 2007. Measuring Quality of Life in Informal Settlements in South Africa. *Social Indicators Research,* Volume 81, pp. 375-388.
- Rossouw, S. & Naude, W., 2008. The Non-Economic Quality of Life on a Sub-National Level in South Africa. *Social Indicators Research,* Volume 86, pp. 433-452.
- Rossouw, S. & Pacheco, G., 2012. Measuring Non-Economic Quality of Life on a Sub-National Level; A Case Study of New Zealand. *Journal of Happiness Studies,* Volume 13, pp. 439-454.
- Samuelson, P. A., 1948. *Economics: An Introductory Analysis,* New York: McGraw-Hill.
- Sen, A., 1985. Well-Being, Agency and Freedom: The Dewey Lectures. *Journal of Philosophy,* 82(4), pp. 169-221.
- Sharpe, A., 1999. *A Survey of Indicators of Economic and Social Wellbeing,* Ottawa: Centre for Study of Living Standards.
- Sharpe, A., 2004. *Literature Review of Frameworks for Macro-indicators,* Ottawa, Canada: Centre for the Study of Living Standards.
- Sharpe, A. & Smith, J., 2005. *Measuring the impact of research on wellbeing,* Ottawa: Centre for the Study of Learning Standards.
- Sirgy, M. J., 2011. Theoretical Perspectives Guiding QOL Indicator Projects. *Social Indicators Research,* 103(1), pp. 1-22.
- Somarriba, N. & Pena, B., 2009. Synthetic Indicators of Quality of Life in Europe. *Social Indicators Research,* 94(1), pp. 115-133.
- South African Audiance Research Foundation, 1974-2013. *Development Index.* [Online]. Available at: http://www.saarf.co.za/amps/developmentindex.asp. [Accessed 15 February 2011].

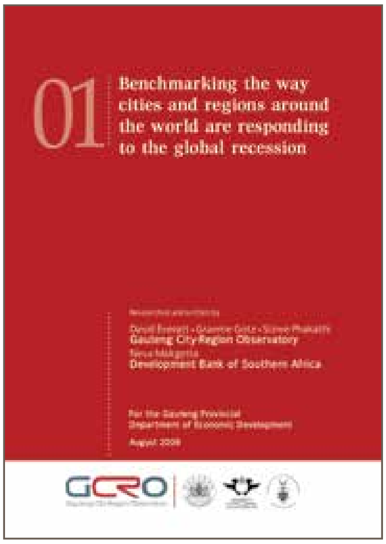
- The South African Institute of Race Relations (SAIRR), 2011. *The South African Development Index.* [Online]. Available at: www.sairr.org.za [Accessed 22 August 2011].
- Stiglitz, J., Sen, A. & Fitoussi, J. P., 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress.* [Online]. Available at: www.stiglitz-sen-fitoussi.fr/en/index/html [Accessed 20 February 2011].
- Sumner, A., 2004. Economic well-being and non-economic well-being: A review of the meaning and meaurement of poverty. *Research paper no 2004/3. WIDER(World Institute of Development Research),* pp. 1-21.
- Tabachnick, B. G. & Fidell, L. S., 2007. *Using multivariate statistics.* 5 ed. Boston: Pearson Education.
- The Centre for Bhutan Studies, 2008. *Gross National Happiness.* [Online] Available at: www.bhutanstudies.org.bt. [Accessed 20 August 2011].
- The Gauteng Planning Commission, 2012. *2055 Gauteng: Long Term Development Plan for Gauteng.* [Online]. Available at: http://gauteng2055.co.za/?page_id=26. [Accessed 5 January 2013].
- The National Planning Commission, 2012. *The National Development Plan.* [Online] Available at: http://www.npconline.co.za.[Accessed 5 January 2013].
- United Nations Development Progamme, 2010. *Human Development Report,* New York: UNDP.
- United Nations, 2012. *The World Happiness Report,* New York: The Earth Institute.
- Vyas, S. & Kumaranayake, L., 2006. *How to do (or not to do)....Constructing Socio-economic Status indices: how to use Principal Components Analysis.* Oxford: Oxford University Press.

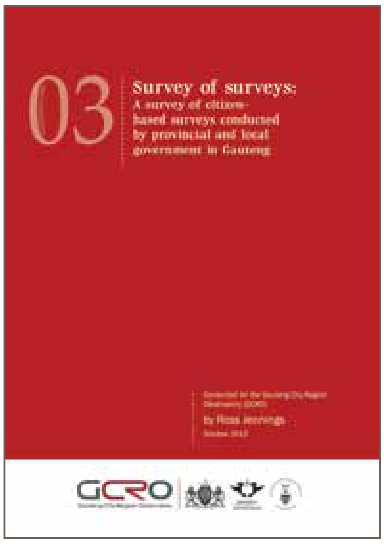
Other Publications

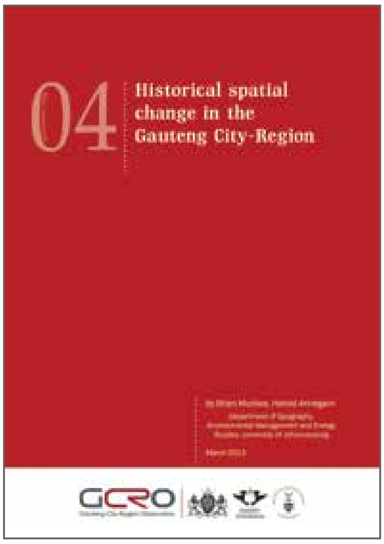
Provocation Series

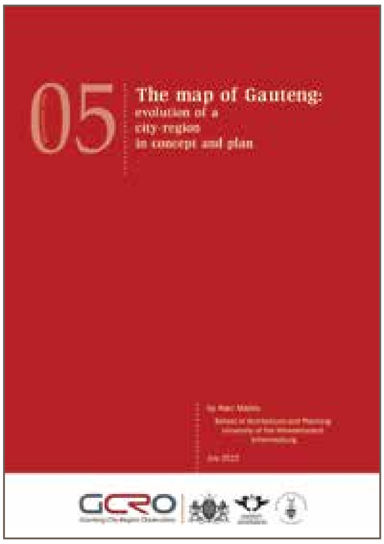


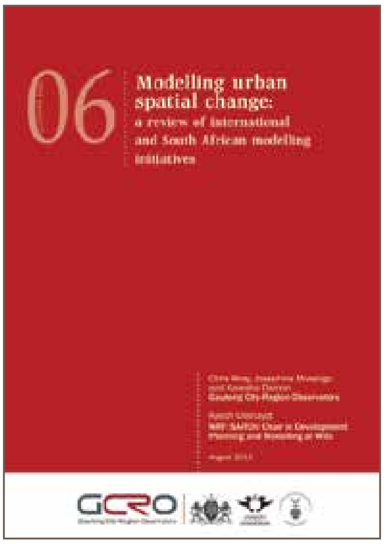
Occasional Paper Series



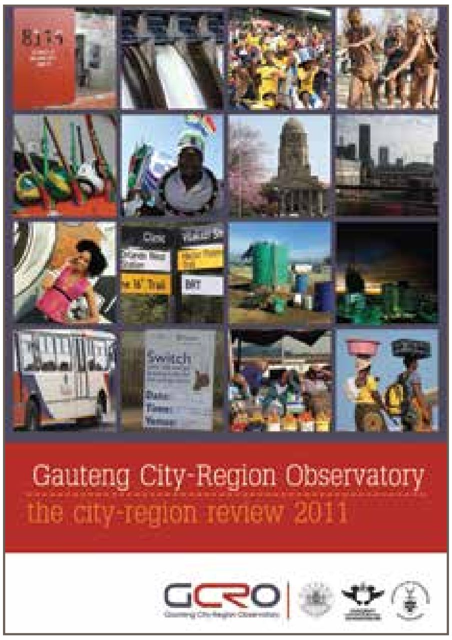




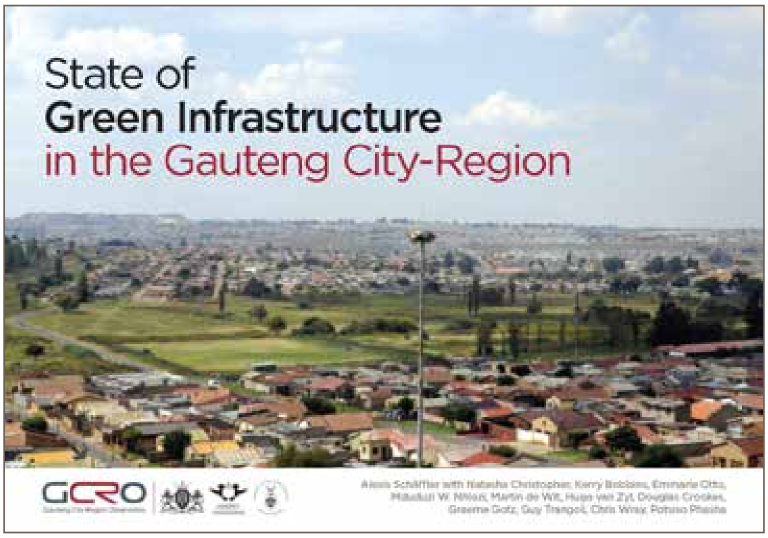


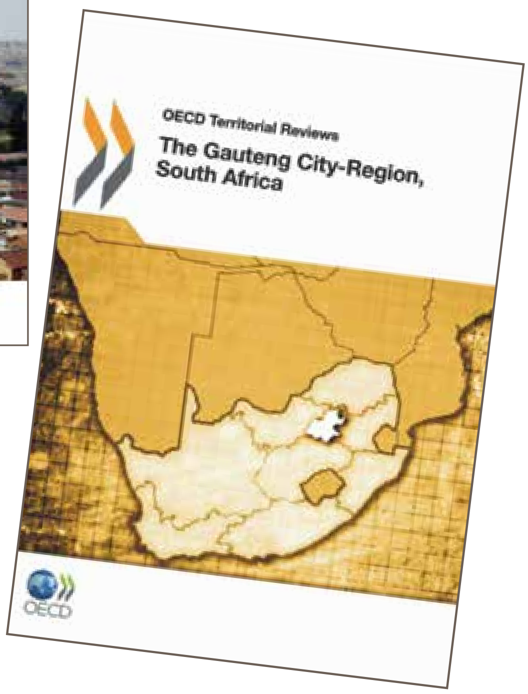


State of the GCR Review 2011 and 2013

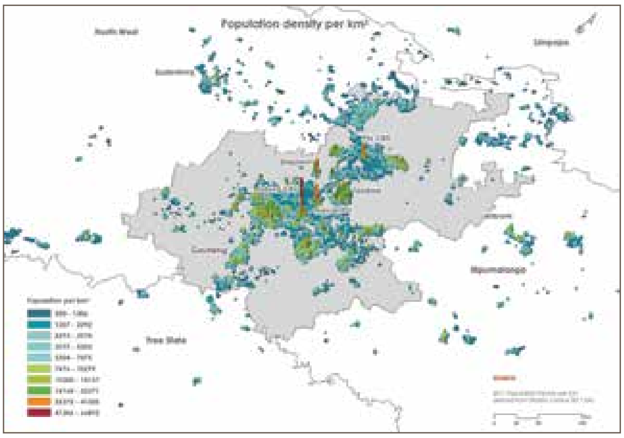


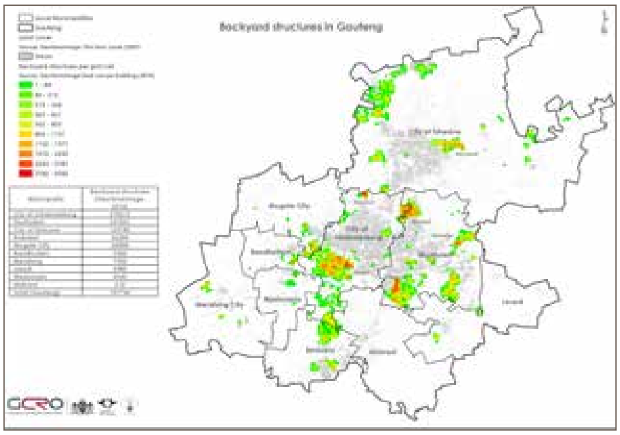
The State of Green Infrastructure in the Gauteng City-Region and the OECD Territorial review

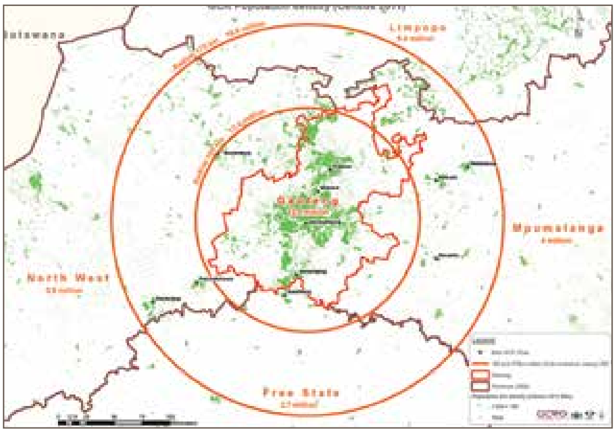


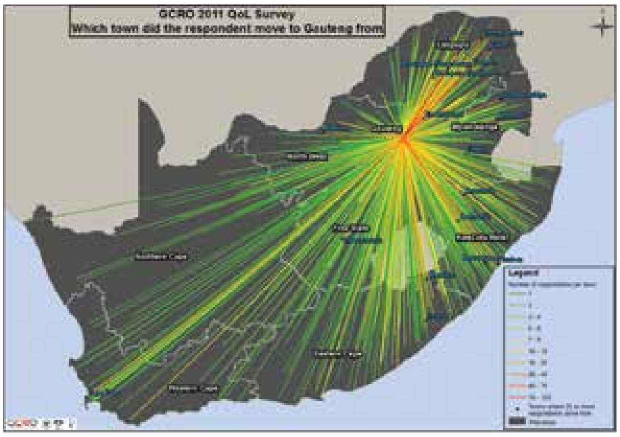


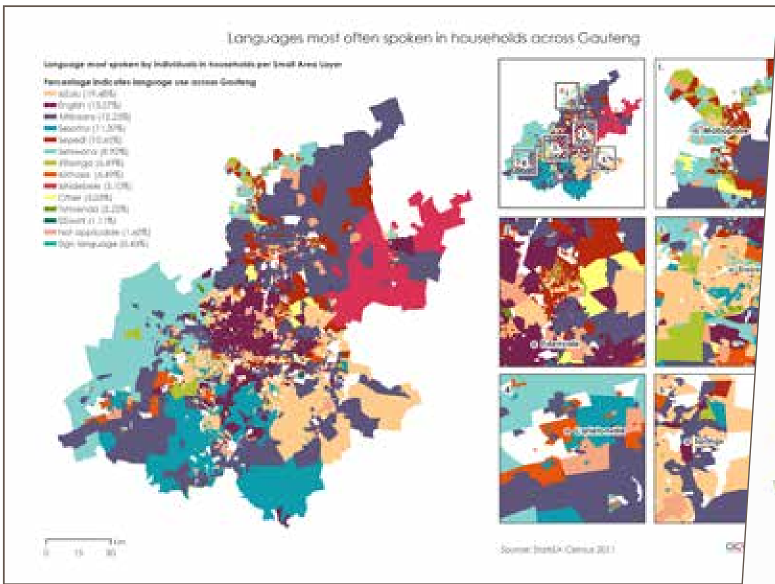
GCRO Maps

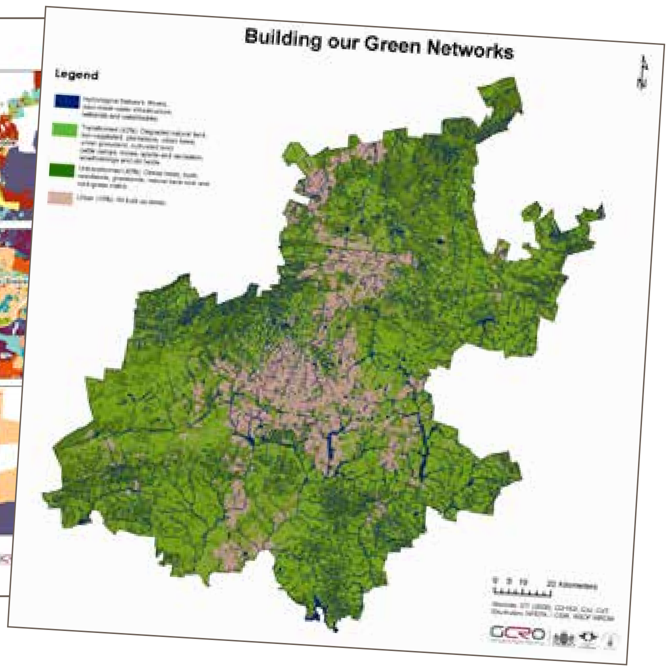

Population density per km²
North West
Limpopo
Gauteng
Mpumalanga
Free State

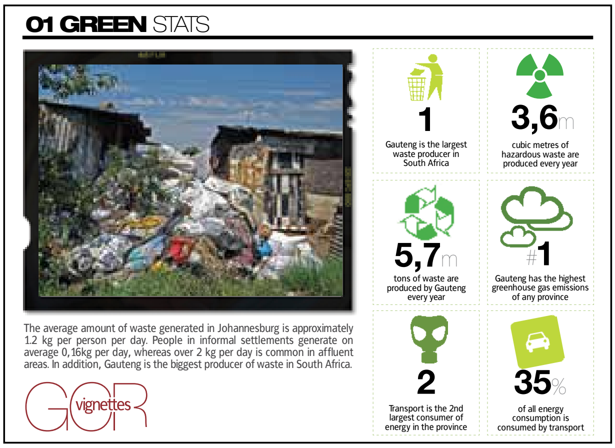

Backyard structures in Gauteng

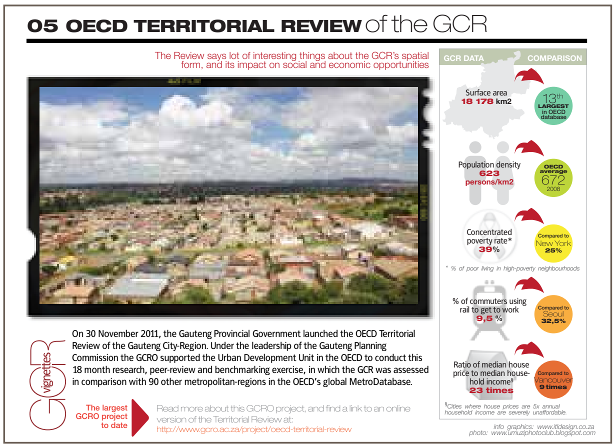

Limpopo
Mpumalanga
North West
Free State

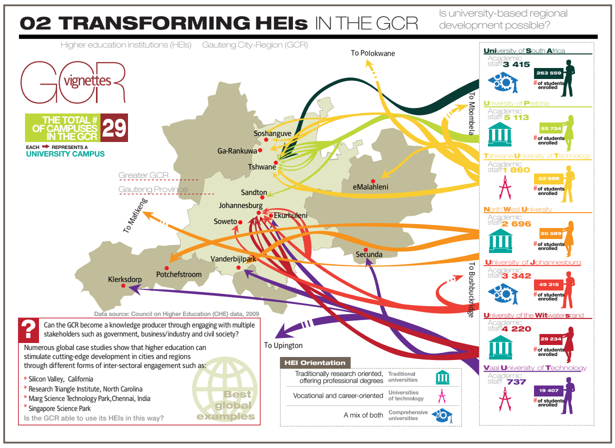

GCRO 2011 QoL Survey
Which town did the respondent move to Gauteng from

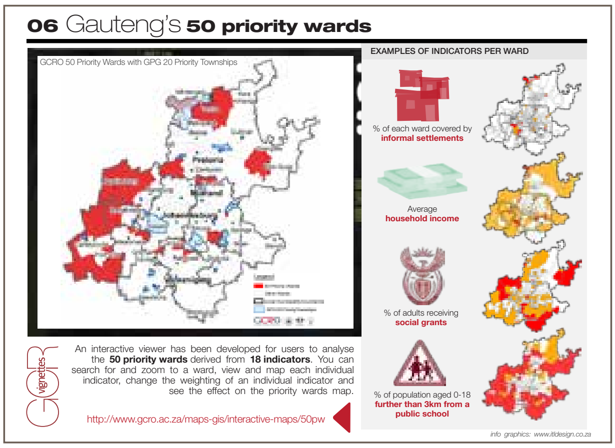

Languages most often spoken in households across Gauteng

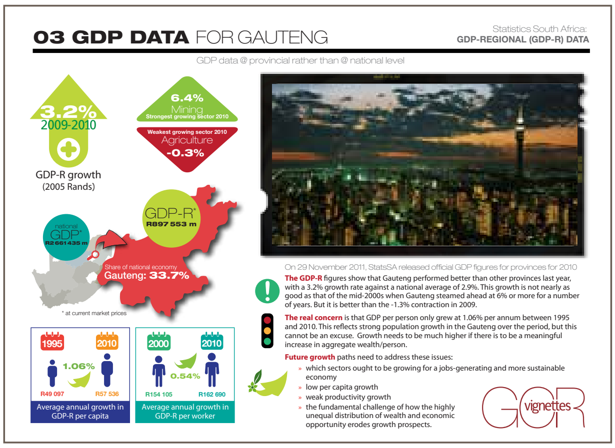

Building our Green Networks
Legend

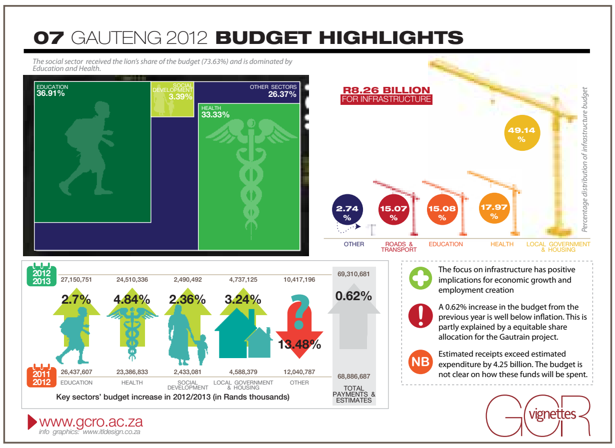

01 GREEN STATS
1
Gauteng is the largest waste producer in South Africa
3,6m
cubic metres of hazardous waste are produced every year
5,7m
tons of waste are produced by Gauteng every year
#1
Gauteng has the highest greenhouse gas emissions of any province
2
Transport is the 2nd largest consumer of energy in the province
35%
of all energy consumption is consumed by transport
The average amount of waste generated in Johannesburg is approximately 1.2 kg per person per day. People in informal settlements generate on average 0,16kg per day, whereas over 2 kg per day is common in affluent areas. In addition, Gauteng is the biggest producer of waste in South Africa.
GCR vignettes

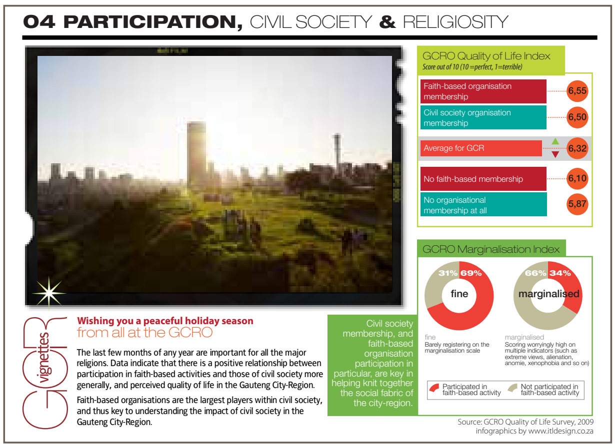

05 OECD TERRITORIAL REVIEW of the GCR
The Review says lot of interesting things about the GCR's spatial form, and its impact on social and economic opportunities
GCR DATA
COMPARISON
Surface area
18 178 km2
13th LARGEST in OECD database
Population density
623 persons/km2
OECD average 672 2008
Concentrated poverty rate*
39%
Compared to New York 25%
* % of poor living in high-poverty neighbourhoods
% of commuters using rail to get to work
9,5 %
Compared to Seoul 32,5%
Ratio of median house price to median household income†
23 times
Compared to Vancouver 9 times
†Cities where house prices are 5x annual household income are severely unaffordable.
On 30 November 2011, the Gauteng Provincial Government launched the OECD Territorial Review of the Gauteng City-Region. Under the leadership of the Gauteng Planning Commission the GCRO supported the Urban Development Unit in the OECD to conduct this 18 month research, peer-review and benchmarking exercise, in which the GCR was assessed in comparison with 90 other metropolitan-regions in the OECD's global MetroDatabase.
The largest GCRO project to date
Read more about this GCRO project, and find a link to an online version of the Territorial Review at:
http://www.gcro.ac.za/project/oecd-territorial-review
GCR vignettes
info graphics: www.itldesign.co.za

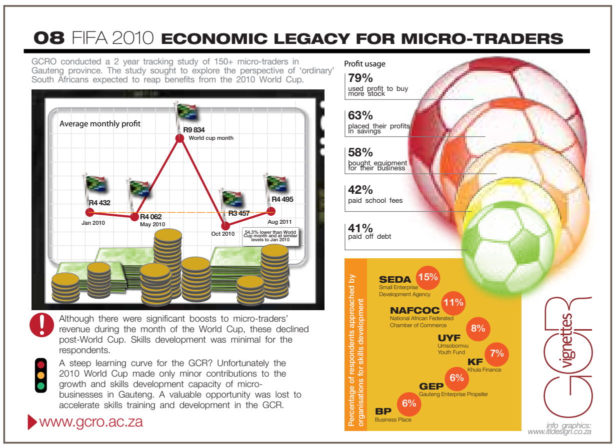

02 TRANSFORMING HEIs IN THE GCR
Is university-based regional development possible?
Higher education institutions (HEIs)
Gauteng City-Region (GCR)
GCR vignettes
THE TOTAL # OF CAMPUSES IN THE GCR 29
EACH REPRESENTS A UNIVERSITY CAMPUS
Greater GCR
Gauteng Province
To Polokwane
Soshanguve
Ga-Rankuwa
Tshwane
Sandton
Johannesburg
Soweto
Ekurhuleni
eMalahleni
Vanderbijlpark
Secunda
Potchefstroom
Klerksdorp
To Mafikeng
To Mbombela
To Bushbuckridge
To Upington
Data source: Council on Higher Education (CHE) data, 2009
University of South Africa
University of Pretoria
Tshwane University of Technology
North West University
University of Johannesburg
University of the Witwatersrand
Vaal University of Technology
3 415
5 113
1 860
2 696
3 342
4 220
737
Can the GCR become a knowledge producer through engaging with multiple stakeholders such as government, business/industry and civil society?
Numerous global case studies show that higher education can stimulate cutting-edge development in cities and regions through different forms of inter-sectoral engagement such as:
Silicon Valley, California
Research Triangle Institute, North Carolina
Marg Science Technology Park, Chennai, India
Singapore Science Park
Is the GCR able to use its HEIs in this way?
Best global examples
HEI Orientation
Traditionally research oriented, offering professional degrees
Traditional universities
Vocational and career-oriented
Universities of technology
A mix of both
Comprehensive universities

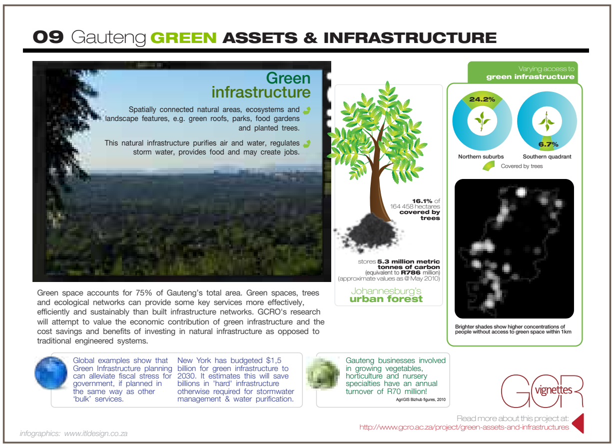

06 Gauteng's 50 priority wards
GCRO 50 Priority Wards with GPG 20 Priority Townships
EXAMPLES OF INDICATORS PER WARD
% of each ward covered by informal settlements
Average household income
% of adults receiving social grants
% of population aged 0-18 further than 3km from a public school
An interactive viewer has been developed for users to analyse the 50 priority wards derived from 18 indicators. You can search for and zoom to a ward, view and map each individual indicator, change the weighting of an individual indicator and see the effect on the priority wards map.
http://www.gcro.ac.za/maps-gis/interactive-maps/50pw
GCR vignettes
info graphics: www.itldesign.co.za

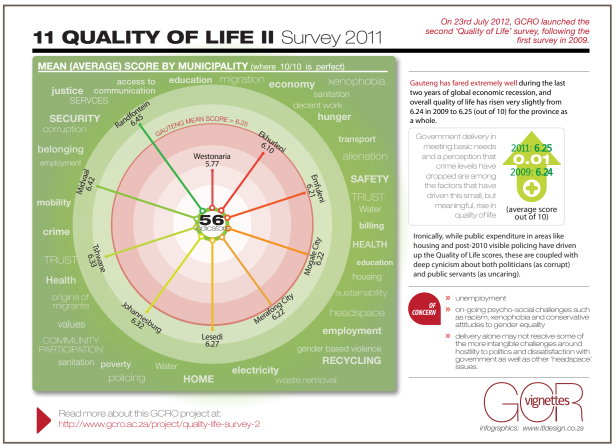

03 GDP DATA FOR GAUTENG
Statistics South Africa:
GDP-REGIONAL (GDP-R) DATA
GDP data @ provincial rather than @ national level
3.2%
2009-2010
GDP-R growth (2005 Rands)
6.4%
Mining
Strongest growing sector 2010
Weakest growing sector 2010
Agriculture
-0.3%
GDP-R*
R897 553 m
Share of national economy
Gauteng: 33.7%
* at current market prices
1995
2010
1.06%
Average annual growth in GDP-R per capita
2000
2010
0.54%
Average annual growth in GDP-R per worker
On 29 November 2011, StatsSA released official GDP figures for provinces for 2010.
The GDP-R figures show that Gauteng performed better than other provinces last year, with a 3.2% growth rate against a national average of 2.9%. This growth is not nearly as good as that of the mid-2000s when Gauteng steamed ahead at 6% or more for a number of years. But it is better than the -1.3% contraction in 2009.
The real concern is that GDP per person only grew at 1.06% per annum between 1995 and 2010. This reflects strong population growth in the Gauteng over the period, but this cannot be an excuse. Growth needs to be much higher if there is to be a meaningful increase in aggregate wealth/person.
Future growth paths need to address these issues:
which sectors ought to be growing for a jobs-generating and more sustainable economy
low per capita growth
weak productivity growth
the fundamental challenge of how the highly unequal distribution of wealth and economic opportunity erodes growth prospects.
GCR vignettes

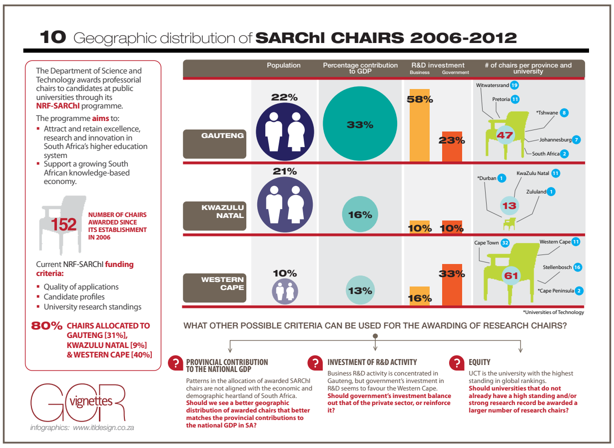

07 GAUTENG 2012 BUDGET HIGHLIGHTS
The social sector received the lion's share of the budget (73.63%) and is dominated by Education and Health.
EDUCATION 36.91%
SOCIAL DEVELOPMENT 3.39%
OTHER SECTORS 26.37%
HEALTH 33.33%
R8.26 BILLION FOR INFRASTRUCTURE
49.14 %
2.74 %
15.07 %
15.08 %
17.97 %
OTHER
ROADS & TRANSPORT
EDUCATION
HEALTH
LOCAL GOVERNMENT & HOUSING
Percentage distribution of infrastructure budget
2012 2013
27,150,751
24,510,336
2,490,492
4,737,125
10,417,196
69,310,681
2.7%
4.84%
2.36%
3.24%
13.48%
0.62%
2011 2012
26,437,607
23,386,833
2,433,081
4,588,379
12,040,787
68,886,687
EDUCATION
HEALTH
SOCIAL DEVELOPMENT
LOCAL GOVERNMENT & HOUSING
OTHER
TOTAL PAYMENTS & ESTIMATES
Key sectors' budget increase in 2012/2013 (in Rands thousands)
The focus on infrastructure has positive implications for economic growth and employment creation
A 0.62% increase in the budget from the previous year is well below inflation. This is partly explained by a equitable share allocation for the Gautrain project.
NB
Estimated receipts exceed estimated expenditure by 4.25 billion. The budget is not clear on how these funds will be spent.
www.gcro.ac.za
GCR vignettes

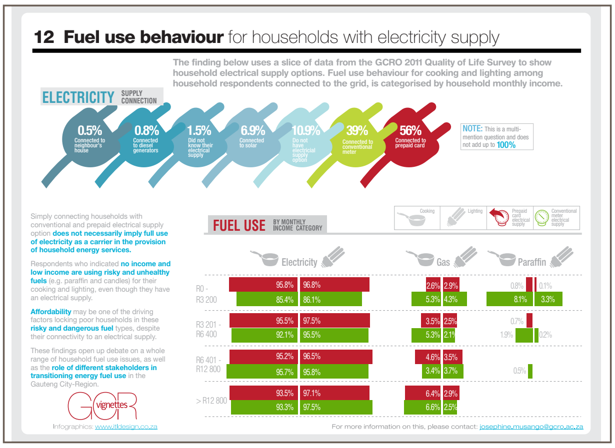

04 PARTICIPATION, CIVIL SOCIETY & RELIGIOSITY
GCRO Quality of Life Index
Score out of 10 (10 = perfect, 1 = terrible)
Faith-based organisation membership 6,55
Civil society organisation membership 6,50
Average for GCR 6,32
No faith-based membership 6,10
No organisational membership at all 5,87
GCRO Marginalisation Index
31% 69% fine
66% 34% marginalised
Participated in faith-based activity
Not participated in faith-based activity
Wishing you a peaceful holiday season
from all at the GCRO
The last few months of any year are important for all the major religions. Data indicate that there is a positive relationship between participation in faith-based activities and those of civil society more generally, and perceived quality of life in the Gauteng City-Region.
Faith-based organisations are the largest players within civil society, and thus key to understanding the impact of civil society in the Gauteng City-Region.
Civil society membership, and faith-based organisation participation in particular, are key in helping knit together the social fabric of the city-region.
Source: GCRO Quality of Life Survey, 2009
infographics by www.itldesign.co.za
GCR vignettes

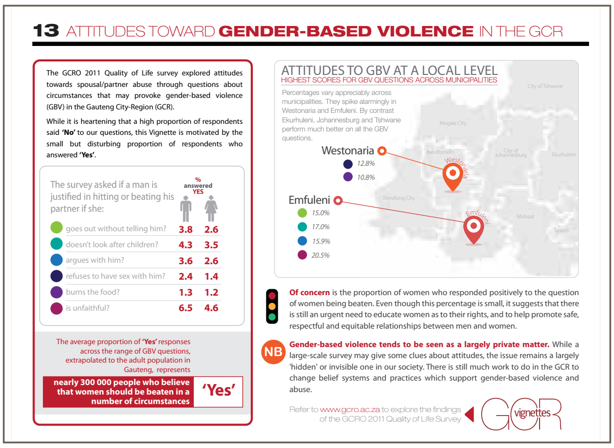

08 FIFA 2010 ECONOMIC LEGACY FOR MICRO-TRADERS
GCRO conducted a 2 year tracking study of 150+ micro-traders in Gauteng province. The study sought to explore the perspective of 'ordinary' South Africans expected to reap benefits from the 2010 World Cup.
Average monthly profit
R4 432
Jan 2010
R9 834
World cup month
R4 062
May 2010
R3 457
Oct 2010
R4 495
Aug 2011
Profit usage
79% used profit to buy more stock
63% placed their profits in savings
58% bought equipment for their business
42% paid school fees
41% paid off debt
Although there were significant boosts to micro-traders' revenue during the month of the World Cup, these declined post-World Cup. Skills development was minimal for the respondents.
A steep learning curve for the GCR? Unfortunately the 2010 World Cup made only minor contributions to the growth and skills development capacity of micro-businesses in Gauteng. A valuable opportunity was lost to accelerate skills training and development in the GCR.
www.gcro.ac.za
Percentage of respondents approached by organisations for skills development
SEDA 15%
NAFCOC 11%
UYF 8%
KF 7%
GEP 6%
BP 6%
GCR vignettes
info graphics: www.itldesign.co.za

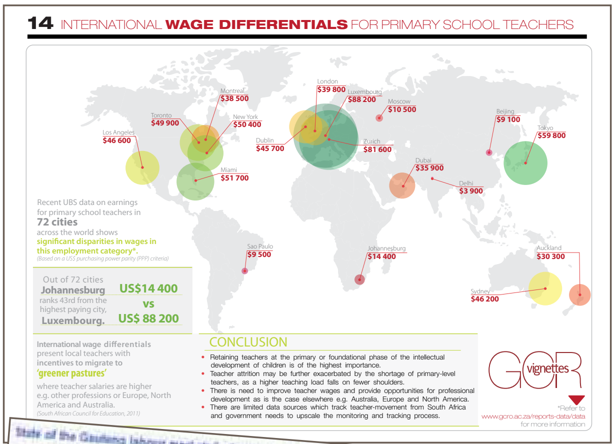


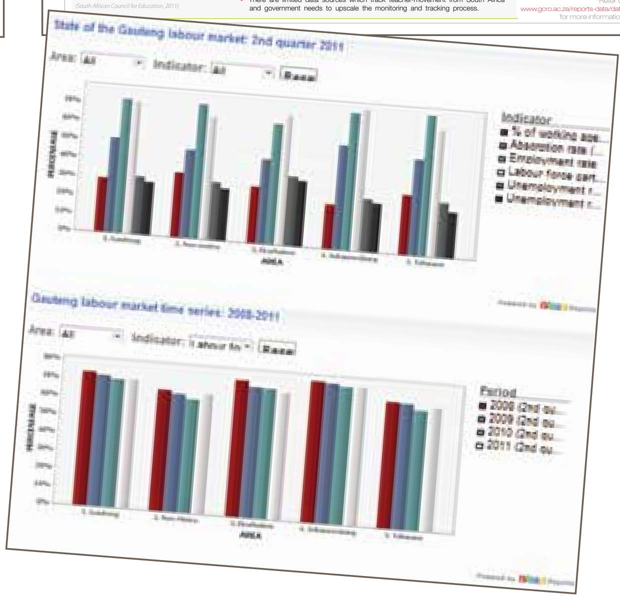












www.gcro.ac.za

GCRO
(Gauteng City-Region Observatory)

4th Floor University Corner
11 Jorissen St (Cnr Jorissen and Jan Smuts)
Braamfontein
Johannesburg
Gauteng
South Africa
Tel +27 11 717 7280
Fax +27 11 717 7281

info@gcro.ac.za www.gcro.ac.za