Soc Indic Res (2017) 131:887–930
DOI 10.1007/s11205-016-1294-5

Construction and Analysis of a Composite Quality of Life Index for a Region of South Africa

Talita Greyling[1] · **Fiona Tregenna**[1]

Accepted: 11 March 2016 / Published online: 18 March 2016

Abstract This study quantifies and analyses quality of life in the Gauteng City-Region of South Africa. First we adapt and extend a method used in research on market regulations to construct composite indices to the field of quality of life. In the adapted method we employ categorical principal components analysis suitable for the analysis of categorical data typically used in quality of life research. The newly constructed index is a comprehensive quality of life index, which includes objective and subjective as well as economic and non-economic indicators. This is the first composite index of its kind in South Africa. Second, this index is used to compare the quality of life of different demographic and socio-economic groups in the region. The quality of life scores of Africans, low income, female, older and urban informal dwellers are relatively low. Third, the explained variance of the dimensions of quality of life is compared across groups. The dimensions 'housing and infrastructure' and 'social relationships' explained the most variance for groups with lower and higher quality of life scores respectively. The results shed light on quality of life in this region, including the great unevenness of well-being. The study provides a basis for the measurement and analysis of quality of life in other regions and countries.

Keywords Quality of life · Well-being · Composite indices · Gauteng City-Region · Measuring instruments

JEL Classification C38 · I31 · O15 · O18 · R11

✉ Talita Greyling
talitag@uj.ac.za

Fiona Tregenna
ftregenna@uj.ac.za

[1] Department of Economics and Econometrics, University of Johannesburg, P O Box 524, Auckland Park, Johannesburg 2006, South Africa

1 Introduction

Quality of life has gained increasing prominence, both in the academic literature and in policy discourse. This derives in part from growing recognition of the inadequacy of economic growth as a measure of progress.

The improvement of the quality of life of all people in South Africa is high on the policy agenda at both national and regional levels (see for instance National Planning Commission 2012; Gauteng Planning Commission 2012). However, quality of life is more amorphous, multi-faceted and difficult to measure than economic growth. Recognising the importance of improving quality of life thus points to the importance of reliable and appropriate ways of measuring it. For both analytical and policy purposes, there is a need for a composite index of quality of life, which can measure the overall quality of life of people as well as tracking it over time and comparing it across different groups or areas. Such a measure could identify those demographic and socio-economic groups with relatively low quality of life, as well as identifying the specific dimensions of quality of life which should be prioritised in order to improve overall well-being.

This article analyses quality of life in the Gauteng City-Region (GCR) of South Africa. There are three key contributions to the literature. First, we construct a composite index of quality of life which is concise, objectively weighted and which includes objective and subjective as well as economic and non-economic measures. To construct the composite index we adapt and extend a method initially developed by Nicoletti et al., which is also included in the Handbook on the Constructing of Composite Indicators (OECD 2008).[1] The method of Nicoletti et al. (2000) has the advantage of explaining more variance in the dataset and thus better representing the data compared to alternative methods. The method was originally developed for the construction of composite indices in the field of market regulations, but is applied for the first time here to the field of quality of life. Second, this constructed composite index is used to calculate and compare the quality of life of different demographic and socio-economic groups in the GCR. The composite index allows for the decomposition of the index and analysis of the drivers of the differences in the levels of quality of life of the groups. Transparency on the drivers of the differences in quality of life is important for policy development, particularly in the light of the high inequality and uneven development within this region (Seekings 2001). Third, the relative explained variance of quality of life is compared for different demographic and socio-economic groups. This is accomplished by conducting separate categorical principal components analyses (CATPCAs) for each group. Optimal scaling is used to quantify the categorical data used in the analyses.

Gauteng, the region of focus of this study, contributed 7.7 % to the gross domestic product (GDP) of Africa and one-third of the national GDP of South Africa (OECD 2011), indicating the region's importance as an economic hub of the country and continent (Gauteng Planning Commission 2012). The region houses 22 % of the national population (11.2 million people) and encompasses seventeen municipalities that function as an integrated region (Gauteng Provincial Government 2012). It has some of the biggest cities in South Africa, including Johannesburg and the national capital of Tshwane (formerly Pretoria). The city-region faces deep socio-economic challenges, including high unemployment (26.9 %)[2] and low productivity growth (OECD 2011). Its rapid demographic and

[1] The method has also been applied by among others Groh et al. (2010), Rukšėnaitė (2011) and Szent-Iványi and Vigvári (2012), but not for analysing quality of life.

[2] The narrow unemployment rate of the GCR (OECD 2011).

economic development has also reinforced the spatial segregation inherited from apartheid. These disparities underscore the importance of analysing quality of life not just in the region as a whole, but for specific groups within it.

The remainder of this paper is structured as follows. Section 2 reviews the relevant literature on quality of life, and the methodology employed is discussed in Sect. 3. Section 4 describes the data used as well as explaining the selection of the indicators. In Sect. 5 we discuss the constructed composite index, and Sect. 6 compares the quality of life of different groups in the GCR using this index. Section 7 analyses which components of quality of life explain the most variance in the quality of life of the different demographic and socio-economic groups. Section 8 concludes.

2 Literature on the Measurement of Quality of Life

We begin by reviewing the international literature on composite quality of life indices and on the measurement of quality of life. Thereafter, we discuss existing measures and literature on quality of life in South Africa as well as findings comparing the quality of life of different socio-economic and demographic groups in the country.

2.1 Composite Indices of Quality of Life

Since the 1990 s there has been greater focus on constructing multidimensional composite indices of quality of life. Probably the best known and most widely used composite index is the Human Development Index (HDI) of the United Nations Development Program (UNDP), which is based on Sen's (1985) 'capabilities and functionings' theory of human development. It combines indicators of life expectancy, educational attainment and income into a composite index (UNDP 2010). The main criticisms of this measure are that it includes only three quality of life dimensions, the dimensions are exclusively objectively measured and the dimensions are equally weighted (Dowrick et al. 2003).

Other composite indices that are frequently used or referenced in the international literature or in policy include the Physical Quality of Life Index (Morris 1979), the Quality of Life Index (Dasgupta and Weale 1992), the Comprehensive Quality of Life Survey (Cummins et al. 1994), the Combined Quality of Life Indices (Diener 1995), the Index of Economic Well-being (Osberg and Sharpe 2000) and the Economist Intelligence Unit's Quality of Life Index (Economist Intelligence Unit 2005). Similar dimensions of quality of life tend to be used in most of these indices: housing, income, employment, community involvement, education, civic engagement, good governance, health, life satisfaction, safety, culture, work life balance, an environmental dimension and certain demographic indicators (Greyling 2013).

Interest in the measurement of quality of life heightened further with the publication of the Report on the Measurement of Economic Performance and Social Progress (Stiglitz et al. 2009). The report re-emphasised that GDP is not an adequate measure of economic progress and drew attention to the importance of developing a multidimensional measure of quality of life which includes both objective and subjective measures.

2.2 Quality of Life in South Africa

Research on the construction of composite quality of life indices in South Africa is limited but growing. Higgs (2007) proposed the 'Everyday Quality of Life Index' to measure well-

being at a national level and at a sub-national level (metropolitan cities and municipalities). His aim was to develop an index which encompass the whole human being's well-being. Rossouw and Naudé (2008) and Naudé et al. (2009) described the construction of non-economic quality of life measures and analysed which municipality in South Africa had the highest quality of life without considering economic factors. The three domains included in their indices were population basic needs, geography and human development.

Furthermore two development indices were published for South Africa, but both were last published in 2011. One was the Development Index of the South African Audience Research Foundation (SAARF 2011) which included indicators on personal data and media penetration, which they regarded as indicators of development. The other was the Development Index of the South African Institute of Race Relations (SAIRR 2011) which included only objective indicators on the economy, education, health, living conditions, gender, and crime.

Lastly at regional level a composite index of quality of life was constructed for the GCR by the Gauteng City-Region Observatory (GCRO). The index included ten dimensions of quality of life measured by 56 indicators that were both objectively and subjectively measured. Combining ten dimensions of quality of life into a single composite index is more than typically used internationally (for a review of the different dimensions and number of dimensions included in quality of life composite indices see Greyling 2013). The GCRO index also subdivided some of the dimensions as found in the literature, for example 'social relationships', into more than one dimension namely community and family relationships. The inclusion of a higher number of dimensions than used in standard measures also implied the inclusion of a higher number of indicators into the composite index. The index was constructed by calculating an average score for each dimension and adding the scores linearly (see GCRO 2011 for a detailed description of the construction of this index).

While the research on the construction of composite indices of quality of life provides important information on the topic, the suggested limited number of indices are mostly not comprehensive measures, in that they measure only objective or subjective quality of life or economic or non-economic quality of life. They are mostly equally weighted, which is generally not appropriate or optimal as it does not give an indication of the relative importance of the different dimensions. Moreover they often include a vast number of indicators which increases the complexity of the index and the cost of collecting the data. Furthermore, none of these indices' aim is purely or specifically to measure quality of life.

Furthermore, the existing indices mainly measure quality of life nationally. South Africa is highly diverse and unequal, in terms of differences and disparities between regions as well as between demographic and socio-economic groups. This makes it sounder, from an analytical and a policy point of view, to analyse the quality of life within a specific region of South Africa and to compare the quality of life of different socio-economic and demographic groups in the region. Using these comparisons, groups with lower levels of quality of life within a region can be identified and dimensions in which their quality of life is most lacking can be addressed. In this study these gaps in the literature are addressed.

Several other studies compare certain individual aspects of quality of life (such as income or education) amongst different racial groups in South Africa. As would be expected, African households are found to be worse off than other groups (see Klasen 2000; Higgs 2007; Posel and Casale 2011; GCRO 2011; Møller 2012; the Human Sciences Research Council 2013). These findings reflect South Africa's apartheid history and the ongoing legacy thereof. There are only a few studies comparing quality of life indicators for demographic and socio-economic groups other than race. This literature finds women,

people in rural areas and people over the age of 46 to have poorer levels of well-being than others (see Klasen 2000; Higgs 2007; Rossouw and Naudé 2008; GCRO 2011).

Finally, the literature regarding factors that play the biggest roles in influencing well-being in South Africa generally identify important factors such as housing, basic services, social relationships, education, health, employment and safety. Amongst these factors, housing and basic services, employment and education are found to contribute most to the well-being of people (see Bookwalter et al. 2006; Higgs 2007; Hinks and Gruen 2007; GCRO 2011; Posel and Casale 2011; Møller 2012).

3 Methodology

Section 3.1 discusses the construction of the composite index of quality of life. Thereafter, the methodologies used to compare quality of life for different population sub-groups and to determine the dimensions of quality of life which explain the most variance in the data for specific groups are explained in Sect. 3.2.

3.1 Methodology Used to Construct the Composite Index

Here, we discuss the different methods of weighting indices and the method of Nicoletti et al. (2000) as adapted to weight and construct composite indices of quality of life. We also describe the methods used to test the validity and the robustness of the constructed composite index.

3.1.1 Weighting Methods

Most studies, for the sake of simplicity, use equal weighting to construct composite indices (Hagerty and Land 2007), thus not attaching different importance to the various individual indicators. However, there are disadvantages of using equal weighting which are often overlooked (OECD 2008). If indicators are grouped into different dimensions, in this case of quality of life, and those dimensions are further aggregated into a composite index, then equal weighting of the indicators may imply unequal weighting of the dimension. This is because indicators in those dimensions which include more indicators will have greater weight than those with fewer indicators. Furthermore, if indicators that are highly correlated are combined into a composite index, an element of double counting is introduced (OECD 2008).

An alternative to equal weighting is to differentially weight indicators when constructing composite indices. There are a number of subjective and objective methods that can be used to weight composite indices (OECD 2008). A benefit of weighting indicators included in composite indices, in addition to addressing the previously mentioned limitations of equal weighting, is to reflect the importance or the preferences of communities for specific indicators, which can also shed light on priorities for policy. Subjective weighting methods depend on the judgement of experts or make use of participatory methods (OECD 2008). A concern in using subjective measures of weighting is that the weighting may not measure the importance of each individual indicator but rather the urgency or need for political intervention concerning the dimensions. It may also be that the subjectively selected weights do not have credibility among the communities or groups being analysed, thus limiting their participation and commitment.

Objective measures[3] to weight composite indices include, among other, multivariate statistical techniques, such as Factor Analysis (FA) and Principal Component Analysis (PCA).[4] According to Booysen (2002), PCA and FA are the multivariate statistical techniques most frequently used in the weighting of composite indices. PCA and FA group together individual indicators which are collinear to form a composite indicator that captures as much as possible of the information common to individual indicators (Tabachnick and Fidell 2007). Each factor or component reveals the set of indicators with which it has the strongest association, with limited correlation between the extracted components, thus each extracted component supplies additional relevant information (OECD 2008). Using PCA or FA, the aim is to account for the highest possible variation in the indicator set using the fewest possible components or factors.

With FA or PCA, the weights are derived from the factor loadings on the extracted components. The largest factor loadings are allocated to the indicators with the largest variation across the dataset, with smaller factor loadings allocated to those indicators with less variation across the dataset. This is a desirable property of FA and PCA in weighting composite indices constructed for the purpose of comparison, as done here, as the indicators that are similar across groups are of little interest and cannot explain the differences between groups. Further advantages of these methods are that no a priori assumptions or information on the weights of the different indicators included in the composite index are required (OECD 2008), the methods are simple though sophisticated and are suitable for the analysis of large datasets (Szent-Iványi and Vigvári 2012). In general, PCA is the preferred method to weight composite indices if an empirical summary of the dataset with a unique mathematical solution is needed (Tabachnick and Fidell 2007). The weights derived using PCA are fixed across all groups giving a good basis for comparison. In this study we selected PCA as the ideal method to weight the newly constructed composite index. We used Nicoletti et al.'s (2000) method, which employs and extends PCA, to construct a composite index, and which has the added advantage of creating composite indices that are more representative of the data compared to other methods based on PCA. We also adapted the basic PCA technique to categorical PCA (CATPCA) in order to extract components, as it is more suitable for the analysis of categorical data.

3.1.2 The Method of Nicoletti et al. (2000)

The standard method when applying PCA as a weighting technique is to use the factor loadings of the indicators on the first component (Klasen 2000). In some cases, the first component is sufficient to adequately represent the original indicators (Ram 1982). However, in other instances, as in this study, the explanatory value of the first component is insufficient to represent the data and alternative methods are needed. Such a method was developed by Nicoletti et al. (2000). This method considers the factor loadings of the first extracted component as well as the factor loadings of the subsequent extracted components to weight a composite index, thus preserving the maximum possible proportion of the total variation in the original dataset. This is the only method mentioned in the *OECD Handbook on Constructing Composite Indicators* that uses not only the first extracted component to derive composite indices, but also consecutive components.

Initially, the method of Nicoletti et al. (2000) was developed to construct composite indices in the field of market regulations, for which the data is continuous, therefore

[3] See OECD (2008) for a discussion of different weighting methods.

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

standard PCA was suitable for extracting the principal components. However, standard PCA has two important limitations relevant here: it assumes that the relationships between indicators are linear and that the data is scaled at the numeric level of either a ratio or an interval scale of measurement. These limitations make standard PCA underperform in the analysis of categorical data, which is the type of data used here and in much quality of life research. Nonlinear PCA methods have been developed to address these limitations of standard PCA (see Linting et al. 2007; Linting 2007 and for more details on the mathematics of nonlinear PCA see Gifi 1990; Meulman et al. 2004). Nonlinear PCA, also known as CATPCA, has the same purposes as standard PCA but it can be applied to indicators of mixed measurement levels which are not necessarily linearly related to one another (Linting et al. 2007).[5] In general, CATPCA performs better than standard PCA in the analysis of categorical data and the results typically explain more variance than using standard PCA.

Although CATPCA is a good alternative to standard PCA due to its ability to analyse categorical data and indicators with nonlinear relationships, it also has disadvantages. In applying CATPCA, different options need to be tried out, which affects the objectivity of statistical analysis as these options rest on subjective decisions by the researcher. Furthermore, due to the various options available in CATPCA the method is more complex than standard PCA and therefore less accessible to researchers.

To construct a composite index using the Nicoletti et al. (2000) method, individual indicators with the highest factor loadings on a specific extracted component are grouped into intermediate composite indices (ICI). The weighting of each of the indicators in the intermediate composite indices is derived by squaring the factor loadings of the indicators and scaling it to unity sum within each intermediate composite index. The squared factor loadings represent the proportion of the total variance of the indicator explained by the component. Once the intermediate composite indices have been constructed, they are aggregated by allocating a weight to each of them equal to the proportion of the explained variance of the component in the dataset. The weight of each consecutive intermediate composite index in the composite index falls as the explained variance of the component decreases.[6]

As well as the previously mentioned advantage of the method of Nicoletti et al. (2000), of having higher representative power, an additional advantage arising from the utilisation of more than the first extracted component is testing the capacity of the data to reflect the underlying theory (OECD 2008). Using PCA as the statistical method (or CATPCA as in this study), it extracts components that reveals the set of indicators with which it has the

[5] Using CATPCA, categorical indicators are transformed by assigning optimal scale values to the categories, thus transforming categorical indicators to numeric-valued indicators referred to as category quantifications. CATPCA, given the number of extracted components, finds optimal category quantifications, implying that the overall variance accounted for in the transformed indicators are maximised (Linting et al 2007). The category quantifications for an indicator together form that indicator's transformation. In the optimal scaling process, the initial information provided by the categorical data is retained in the category quantifications. However the degree to which the original data is retained depends on the optimal scaling level (nominal, ordinal or numeric) selected for each indicator (Linting et al. 2007). The transformed indicators have metric properties which allow the researcher to use all standard methods normally applied to continuous indicators.

[6] In mathematical terms the aggregation of the intermediate composite indices to derive at the composite index of quality of life (CIQoL) is as follows: $\text{CIQoL}_i = \left(\sum_{i=1}^{n} (ICI_i w_i)\right) \times 100$ where CIQoL is the composite index of quality of life, ICI_i = Intermediate Composite Index ($i = 1 \ldots n$), and w_i is the weight of the ith ICI determined by the explained variance of the intermediate composite index (extracted component) in the dataset.

strongest association. Thus, each extracted component provides additional relevant information (OECD 2008), and if these components mirror the theory it demonstrates the robustness of the method used to construct the composite index. Another advantage of the method derives from the construction of intermediate composite indices, as it is possible to trace back indicators and dimensions to reveal detailed information on the differences in quality of life between groups (Groh et al. 2010). This increases the analytical power of the method and, in a study such as this one, the capacity to draw out policy implications.

The last step in the construction of a composite index is to test the validity and the robustness of the index, which is done here in three ways. There are no existing directly comparable quality of life indices to which the newly constructed composite index can be benchmarked, as these indices differ significantly from the current index in time, geographical areas, composition and method of weighting of the indices (see Sect. 2.2). However certain conclusions drawn from comparing these previous measures of quality of life and results found in previous studies to the findings on quality of life of the current developed index (see Sects. 5–7). Secondly, we test the validity of the index by correlating it to a single traditional measure of well-being, namely life satisfaction. For ease of comparison we express the life satisfaction variable as a percentage[7] for each socio-economic and demographic group. The third test of the validity of the composite index is correlating it to two own constructed indices using the same data: one with equal weighting and another adding all the indicators to a sumscore (see Sect. 5). The degree and significance of the correlation between the indices give an indication of the validity of the newly constructed weighted composite index (see Groh et al. 2010).

The robustness of the newly constructed composite index was tested by running CATPCAs, using the finally selected set of data using different methods of treating the missingness in the data, with data imputation, pairwise deletion and listwise deletion (see Sect. 4 and Appendix 4 in Table 12). In addition, the selection of the optimal scaling method in applying CATPCA was varied between ordinal, nominal, and spline ordinal and spline nominal scaling methods to determine if the results were robust to choice of scaling (see Sect. 5). The choice of weights of the indicators was prescribed by the method, though we tested whether the extracted order of the dimensions of quality of life remained the same if we run a CATPCA on a similar set of indicators also selected from the preliminary list. In order to select the similar set of indicators, we did correlation analysis and replaced, where available, indicators selected in the final list with closely correlated indicators from the preliminary list. Where there were no closely correlated indicators, for example years of education, we retained the original indicator. In this manner we derived a second set of indicators similar to the first set, though not the optimal set as it explained less variance in the data (see Sect. 5 and Appendix 4 in Table 12 for the similar set of indicators). Although the similar set of indicators is not the optimal one, running a CATPCA on this set of indicators tests the reliability of the order of the extracted components. The order of the extracted components determined the weight of each consecutive extracted component, with the weighting decreasing from the first to the last extracted component.

[7] We derived the percentage by dividing the life satisfaction score of each respondent by the maximum life satisfaction score (5) and multiplying by 100.

3.2 Methods Used in Calculating Quality of Life Scores and Comparing the Order of the Extracted Components of Quality of Life for Different Groups

We next used the newly constructed composite index to calculate the quality of life score of each respondent in the dataset. To compare the quality of life scores of the different demographic and socio-economic groups (race, sex, age groups, area, and income groups) we computed the mean score of each group.

To test whether there are significant differences in the mean scores of the different groups (for example race) we used one-way between-groups analysis of variance (ANOVA). In cases where we compared the means of only two groups we used t-tests. Where we found significant differences between the mean scores of the groups, we calculated the effect sizes to establish the magnitudes of the differences. For this purpose we made use of the effect size statistic, *eta squared*.[8] To interpret the effect size statistic we used Cohen's (1988) rules of thumb stating that an effect size with an *eta squared* statistic of 0.01 is small, 0.06 is moderate and 0.14 is large. As the one-way between-groups ANOVA is an omnibus test statistic that reveals only whether there are significant differences in the mean scores of the analysed groups, but does not reveal between *which* groups, we also conducted post-hoc tests using the Tukey HSD test (Pallant 2007). The post-hoc tests reveal between which groups there are significant differences. We investigated possible reasons to account for the differences in the mean quality of life scores between the groups.

We also undertook regression analysis to determine whether and how background and socio-economic characteristics of the respondents affect their quality of life scores, and of particular interest here, how these relationships vary between the different demographic and socio-economic groups. To estimate the regression coefficients we used the Ordinary Least Square (OLS) estimation technique and ran diagnostic tests to test for normality of the distribution of the residuals, multicollinearity, homoscedasticity and the independence of the residuals. The findings from this part of the analysis are presented at the end of Sect. 6.

Finally we used CATPCA to extract the principal components of the different demographic and socio-economic groups in the region to highlight the dimensions of quality of life which explained the most variance in the data of the different groups. According to the methodology of Nicoletti et al. (2000) the dimension which explains the most variance in the data carries a higher weight in the composite index. These results shed some light on the order of the priorities of the different groups. These results are discussed in Sect. 7.

4 Data and the Selection of the Indicators

This study utilises a dataset collected by the GCRO (2009) specifically to measure quality of life. The survey covered many dimensions of quality of life including both objective and subjective indicators. The questions asked in the survey render it ideal for this study. The

[8] To calculate *eta squared* when using ANOVA we used the formula $Eta\ squared = \frac{sum\ of\ squares\ between\ groups}{Total\ sum\ of\ squares}$, and when using t tests we used the formula $Eta\ squared = \frac{t^2}{t^{2+(N1+N2-2)}}$ (for more information on the calculation of effect sizes see Pallant 2007).

only notable omission is of indicators measuring the environmental dimension of quality of life.

The GCRO used a multi-stage cluster sampling method as no sampling frame of the entire population exists (Development Research Africa 2009). The sample was drawn in three stages. In the first stage, all 602 wards in the GCR were selected. Dwelling units within the wards were selected in the second stage, and respondents within dwelling units were selected in the third stage. During the second stage, the field teams undertook a certain number of interviews per ward, based on the population of the ward. Random starting points were identified from ward maps. The number of starting points per ward ranged between one and six, pre-specified depending on the number of interviews to be undertaken in a ward. From each starting point, the 20th household to the left was identified for the first interview, thereafter the next 20th household to the left for selected, and so on, until the required number of interviews for the ward was obtained. In cases of household refusal, the household was substituted with the one on the left. If there was still no interview completed then a further substitution, to the right of the originally selected household, was made. In instances of non-contact, where there was nobody at home after at least two visits at different times on the same day, the same substitution method was followed. During the third stage, one randomly selected household respondent was selected per household. All household members aged 18 years of age or older were eligible for selection (in child-headed households, the oldest child was interviewed). In a household with more than one eligible person the Kish[9] grid was used to randomly select one individual per household after listing all eligible individuals in the household. Substitutions could not be made between individuals within a household in cases of refusals or non-contact after at least two visits (in these instances, the household substitution method described in stage 2 above was followed). All interviews were face-to-face, with no proxy interviews. The pre-specified number of interviews were obtained per ward, using the random selection procedure among households and among individuals within households.[10] Weights were allocated per municipality in accordance with the census data (Statistics South Africa 2002) to ensure the data is representative. Missing data in all the variables were checked against the questionnaire, and variables were cross-checked for inconsistencies. Wherever possible, respondents were contacted to verify responses or to see if there was an explanation for inconsistencies.

A total of 6636 respondents in the 602 wards in all 17 municipalities of the GCR were interviewed (GCRO 2011).[11] Table 1 summarises demographic and socio-economic characteristics. The vast majority of the respondents in the sample resided in urban areas, indicating the degree of urbanisation in the GCR. The low level of household income for most respondents is apparent.[12] The share of the income of the lowest decile of income

[9] The Kish grid is commonly used in selecting cases at random when more than one case is eligible for inclusion when the interviewer calls at a sampled address or household.

[10] No information is available in the dataset as to whether respondents were in the originally selected households or in substituted households.

[11] Due to a field sampling error, some municipalities were over-sampled and a few were under-sampled. Therefore the data was weighted in accordance to the Census 2001 (Stats SA 2002) population data.

[12] The income figure includes salaries, grants, pensions and any other source of income. Household income was used in the analysis as neither individual income nor a continuous income indicator were available in the survey. An individual income measure was calculated, converting bracket midpoints to a per capita measure. This was found to be highly correlated with the household income measure. The estimated per capita measure was also used in the analysis as a robustness check, and the results were highly consistent with those from the household measure, which was preferred. R1 is approximately US$0.9 (March 2014).

Table 1 Descriptive statistics of the socio-economic and demographic groups

	Frequency (*N*)	% of Sample
Geographical area		
Urban formal (built-up town or city area)	4081	61.5
Urban informal (including peri-urban)	2263	34.1
Tribal- and- farming	292	4.4
Income groups[a]		
R0–R800	976	14.7
R801–R1600	1481	22.3
R1601–R3200	1594	24.0
R3201–R12,800	1740	26.2
R12,801–R51,200	711	10.7
R51,201+	134	2.0
Race		
African	5482	82.6
Asian/Indian	63	0.9
Coloured	203	3.1
White	888	13.4
Sex		
Male	2694	40.6
Female	3942	59.4
Age		
18–20	464	7.0
21–30	1977	29.8
31–40	1638	24.7
41–50	1154	17.4
51–65	941	14.2
66–75	337	5.1
75+	125	1.9

Note Asians were dropped from the analyses in Sects. 6 and 7, due to their small numbers in the sample, in order to ensure statistical robustness of the analysis

Source Authors' calculations based on Quality of Life Survey (GCRO 2009)

[a] Average household income per month

earners in the GCR was only 0.2 % compared to the top decile earners' share of 67.7 % (GCRO 2011). This reflects the considerable skewness of the income distribution in the GCR.

The selection of the indicators to be included in the composite index of quality of life was guided by the recommendations of the Stiglitz Report (2009). These dimensions included: health, education, everyday activities including work and housing, participation in the political process, social relationships, safety from bodily harm and the environment in which people live. This can be considered a condensed list from the list of fourteen dimensions mentioned earlier from the wider literature. To identify the indicators used to measure these dimensions we reviewed existing international and South African indices (see Sect. 2.2). Thereafter we studied the quality of life questionnaire of the GCRO (2009) and identified all possible indicators which could be used to measure the dimensions. These

indicators referred to as the 'preliminary list' are shown in Appendix 2. However as mentioned previously there were no indicators included in the survey to measure the environmental dimension, therefore this dimension could not be included in the composite index.

We normalised the data by recoding all the ordinal indicators to have the same direction of coding, such that one indicated the most 'negative'. We also transformed continuous data such as education into ordinal data. Nominal indicators were transformed into dichotomous indicators with a value of either one or two (see Appendix 2 for the coding of the indicators).

After the recoding of all the indicators in the preliminary list, we computed the percentage of missing values for each of the indicators. Very few indicators had missing values, with missing values constituting less than 1% of most indicators. When analysing the missingness in the indicators it was found to be random across municipalities. In the analyses described in Sect. 5 we imputed the missing values using the corresponding modes of the indicators. To test the robustness of treating the missingness through imputation we also undertook listwise deletion and pairwise deletion in the rotated matrices used in Sect. 5 (see Appendix 4 in Table 12). We found the results of the analyses using the different methods of treating the missingness of the data to be very similar, and therefore used the imputed dataset in the final analysis given the higher number of observations.

In building the composite index of quality of life, our aims in the selection of the final set of indicators were that the set should be parsimonious and explain the most possible variance in the data. We analysed the preliminary selected set of indicators (see "Appendix 2" for the list) to determine which of these indicators were highly correlated. Where specific indicators were found to be highly correlated, for example 'feel safe in area where you live in the day' and 'feel safe in area where you live after dark', we used the indicators alternatively in sets of indicators and ran successive rounds of CATPCA to determine which of the indicators was best explained by the extracted component, thus having the highest communality.[13] The derived communalities of the indicators were then used to guide the selection process. Furthermore we also varied the scaling levels (spline ordinal, spline nominal, ordinal and nominal) and the number of extracted components (between one and seven[14]) as a robustness check. Finally we selected a set of 15 indicators which combined explained the most variance in the data and was therefore optimally representative of the data. The final set of selected indicators and their descriptive statistics are shown in Table 2.

5 Composite Quality of Life Index

We followed the methodology explained in Sect. 3.1 to construct the composite index of quality of life. To derive the weights of the composite index, following the adapted method of Nicoletti et al. (2000), we ran a CATPCA on the selected indicators. The following steps and decisions were taken in preparation for this.

Included in the final set of 15 indicators were 11 ordinal indicators and four nominal indicators. We selected the ordinal scale of quantification for all 11 ordinal indicators and

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

[14] Seven was the total number of dimensions of quality of life used to guide the selection of the indicators (see the list of dimensions in Sect. 4).

Table 2 Indicators included in the Composite Index of Quality of Life

Indicator	Type of data	Min	Max	Mode
Type of dwelling	Dichotomous	1	2	2
Electricity for lighting	Dichotomous	1	2	2
Water on premises	Dichotomous	1	2	2
Satisfied with time available for own things	Ordinal	1	5	4
Satisfied with time available for friends	Ordinal	1	6	5
Satisfied with time available for family	Ordinal	1	5	4
Income	Ordinal	1	6	4
Satisfied with amount of money	Ordinal	1	5	2
Employment status	Dichotomous	1	2	2
Years of education	Ordinal	1	6	5
Satisfaction with health	Ordinal	1	4	3
Health prevent you from taking part in social activities	Ordinal	1	4	4
Feel safe during the day	Ordinal	1	5	4
Country is going in right direction	Ordinal	1	5	4
Elections was free and fair	Ordinal	1	5	4

Source Authors' calculations using GCRO data (GCRO 2009)

the nominal scale of quantification for all four nominal indicators. We preferred the ordinal and nominal scales of quantification to the spline quantifications. The reason for this preference is that it has been shown that if the number of categories of the indicators is relatively small, as in this study, ordinal and nominal quantifications lead to higher levels of explained variance than do spline quantifications as spline quantifications are more restrictive (Linting et al. 2007). However to evaluate the robustness of the scaling choices both options were tested. The results using spline quantifications differed only slightly from the solution using ordinal and nominal quantifications. The explained variance using spline quantifications was 65.9 % and using ordinal and nominal quantifications was 66.1 %.

To decide on the number of components to extract, we used the scree plot as a guideline (see Fig. 1). The scree plot indicated the extraction of six components judging from the 'elbow' of the scree plot which was located at the sixth component. To simplify the structure of the extracted solution we rotated the results. As rotation is not yet a standard provision in CATPCA we used varimax rotation as the data was orthogonal (the correlation of the components was less than 3 %—see "Appendix 3" for the component correlation matrix) within standard PCA to rotate the transformed indicators. After rotating the data we found the six factor component extraction was not theoretically meaningful and added little value (see Fabrigar et al. 1999). Therefore guided by the Kaiser rule,[15] stating that the number of components with an eigenvalue of 1.0 or more should be extracted, we extracted five components. The five extracted components made theoretical sense and mirrored the initial dimensions used as a guideline to select the indicator variables (see Sect. 4),

[15] The Kaiser rule, known as the Kaiser's criterion or the eigenvalue rule is related to the eigenvalue of each principal component. The eigenvalue (variance) for each principal component indicates the percentage of variation explained in the data. Using this rule, components with an eigenvalue of 1.0 or more are extracted.

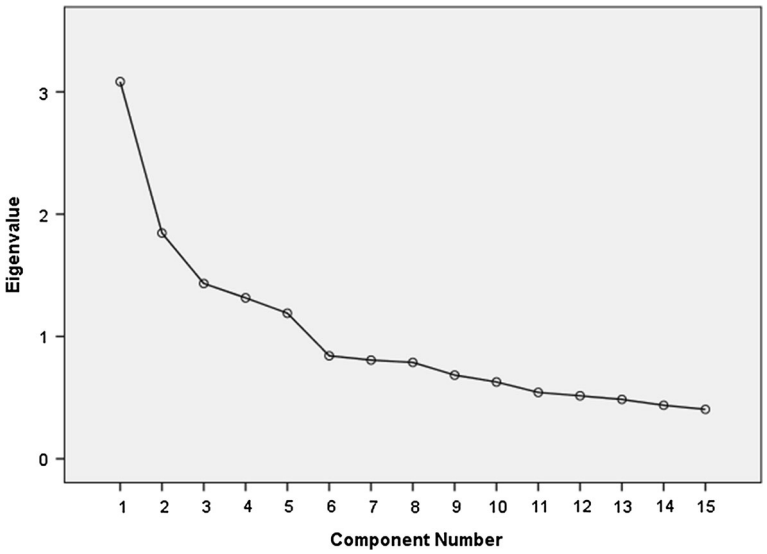

Fig. 1 Scree Plot indicating the quality of life components. *Source* Authors' calculations using GCRO data (GCRO 2009)

although certain of the dimensions loaded on the same component where other dimensions split over two separate dimensions and loaded on two components. The extracted components explained 60 % (see Table 3) of the variation in the dataset, which is an acceptable value of explained variance to be used in further analyses (for comparative studies see Vyas and Kumaranayake 2006; Rossouw and Naudé 2008; Naudé et al. 2009; and Rossouw and Pacheco 2012).

Table 3 shows the rotated component matrix of the five extracted components with varimax rotation. From the factor loadings of the indicators on the extracted components, the set of indicators with the strongest association with the component was revealed. The results reflected the dimensions of quality of life as found in the literature and described in the Stiglitz Report (2009). However, certain of the dimensions loaded on the same component, such as the 'participation in the political process' and the 'safety from bodily harm' dimensions, and others divided into two separate dimensions and loaded on two components, such as the 'everyday activity' dimension. We named each extracted component appropriately to describe the set of indicators with the highest factor loadings on the component (these names differ from those used in the Stiglitz Report to be more descriptive of the extracted components here). The names, in order of the extracted components, are as follows: housing and infrastructure, social relationships, socio- economic status, health, and governance and safety. This result is broadly consistent with the findings of Higgs (2007) and the GCRO (2011).

To further test the robustness of the order of the extraction of the components, CATPCA was performed on a similar set of indicators, also selected from the preliminary list, as was explained in the methodological section (see Sect. 3.1). The rotated component matrix, using a similar set of indicators, revealed the same order of extraction of the components (see Appendix 4 in Table 12) as was revealed using the final selected set of indicators,

Table 3 Factor loadings used in the calculations to weight the composite index based on the rotated component matrix

	Extracted components					Squared factor loadings, scaled to unity sum				
	1	2	3	4	5	1	2	3	4	5
Type of dwelling	**.75**	−.02	.10	.02	.13	**.29**	.00	.00	.00	.00
Electricity for lighting	**.83**	.00	.02	.11	−.05	**.35**	.00	.00	.00	.00
Water on premises	**.83**	.05	.08	.08	.03	**.36**	.00	.00	.00	.00
Satisfied with time available for family	−.03	**.81**	.8	−.04	−.02	.00	**.38**	.00	.00	.00
Satisfied with time available for own things	.06	**.78**	.06	−.03	.05	.00	**.35**	.00	.00	.00
Satisfied with time available for friends	−.01	**.69**	.23	.08	.05	.00	**.27**	.00	.00	.00
Income	.11	.33	**.69**	.24	.08	.00	.00	**.32**	.00	.00
Satisfied with amount of money	.09	.09	**.71**	.10	.01	.00	.00	**.27**	.00	.00
Employment status	.02	.02	**.76**	−.24	.08	.00	.00	**.27**	.00	.00
Years of education	.06	.29	**.49**	.41	−.08	.00	.00	**.14**	.00	.00
Satisfaction with health	.03	−.02	.07	**.83**	.05	.00	.00	.00	**.52**	.00
Health prevent you from taking part in social activities	.15	−.05	−.06	**.79**	.08	.00	.00	.00	**.48**	.00
Feel safe during the day	.11	.04	.12	.15	**.48**	.00	.00	.00	.00	**.16**
Country is going in right direction	.02	.04	−.04	.00	**.78**	.00	.00	.00	.00	**.42**
Elections was free and fair	−.04	−.02	−.08	−.07	**.78**	.00	.00	.00	.00	**.42**
Total % of explained variance of data	19.43	13.13	10.01	9.81	7.48					
(Eigenvalue)	(2.91)	(1.99)	(1.61)	(1.44)	(1.14)					
[a]Total % of explained variance of extracted components	.32	.22	.17	.16	.13					

Note (1) Varimax rotation was used. (2) Factor loadings in bold indicate the highest factor loadings on a specific component

Source GCRO dataset (GCRO 2009)

[a] Extracted components explain 60 % (19.43 + 13.13 + 10.01 + 9.81 + 7.48) of the variance of the data

reflecting the robustness of the order of the extracted components. This supports the results on the dimensions of quality of life which explains the most variance in the dataset.

Having checked the robustness of the order of the extracted components, we proceeded to construct the intermediate composite indices representing the different dimensions of quality of life. The indicators with the highest factor loadings on a component (in bold in Table 3) were grouped together and weighted according to the squared factor loadings scaled to unity sum (as explained in Sect. 3).

The indicators with the highest factor loadings on the first extracted component were 'water on premises' (0.83), 'electricity used for lighting' (0.83) and 'type of dwelling' (0.75). After the factor loadings of each of the indicators were squared and scaled to unity sum (see right hand side of Table 3), the weights of the indicators in the intermediate composite index were: 'water on premises' (36 %), 'electricity used for lighting' (35 %)

and 'type of dwelling' (29 %). These indicators were aggregated to derive an intermediate composite index designated 'housing and infrastructure'. The 'housing and infrastructure' intermediate composite index explains 32 %[16] of the extracted components variance (see Table 3). Therefore it has the highest weight in the composite index of quality of life. The indicators included in the 'housing and infrastructure' intermediate composite index vary greatly among the respondents and have an important role in accounting for differences in the levels of quality of life.

Economically, this dimension represents one of the basic needs of people. According to Maslow's hierarchy of needs, it is one of the first needs that people need to fulfil. The finding concerning the importance of this dimension is consistent with previous research (see for instance Bookwalter et al. 2006; Hinks and Gruen 2007; GCRO 2011; Møller 2012). These studies show 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 well-being of the people living in South Africa.

The indicators 'satisfied with time available for family' (0.81), 'satisfied with time available for own things' (0.78) and 'satisfied with time available for friends' (0.69) had the highest factor loadings in the second component. In the same manner as described for the construction of the 'housing and infrastructure' intermediate composite index, the intermediate composite index 'social relationships' was constructed. Social relationships are also a need described by Maslow (1943), who argued that people need to be loved and cherished and to feel part of a community. This dimension explained the second most variance in the five extracted components and carried a weight of 22 % (13.3/60—see Table 3) in the composite quality of life index.

In the third component the indicators with the highest factor loadings were 'employment status' (0.76), 'income' (0.69), 'education' (0.49) and 'satisfaction with amount of money' (0.71) which carried weights of 27, 32, 14 and 27% respectively in the intermediate composite index. These indicators are related to economic issues, with education playing an important role in employment and income (Berenger 2007). These indicators were weighted and aggregated to form the 'socio-economic status' intermediate composite index. This dimension explained the third most variance in the extracted components, and carried a weight of 17 % (10.01/60—see Table 3) in the composite quality of life index.

'How often does health prevent you from taking part in social activities' (0.79) and 'satisfaction with health' (0.83) were the two indicators with the highest factor loadings on the fourth component. We weighted and aggregated the two health indicators in the same manner as previously explained, with weights of 52 and 48 % to form the intermediate composite index 'health'. This dimension is related to the basic functioning of people, as health is essential to fulfil the end goals of human life and is integrally related to virtually all other aspects of life. This dimension explained 16 % (9.81/60) of the variance in the extracted components and had a corresponding weight in the composite index of quality of life.

In the fifth component designated 'governance and safety', the indicators with the highest factor loadings were 'feel safe during the day' (0.48), 'country is going in the right direction' (0.78) and 'elections was free and fair' (0.78); these indicators reflect functionings of a higher order. The weight allocated to each of the indicators included in the intermediate composite index was 16, 42 and 42 % respectively. This dimension explained

[16] 32 % was derived by dividing 19.43 % of the variance explained by the first extracted component by 60 %, which is the sum of the variances of each of the five extracted components (19.43 + 13.13 + 10.01 + 9.81 + 7.48) (see Table 3).

13 % (7.48/60) of the variance in the extracted components and carried this weight in the composite index of quality of life. In the literature, it has been shown that satisfaction with the government of a country is highly correlated with well-being (Dasgupta and Weale 1992). Finally, we constructed the composite index of quality of life by weighting and aggregating each of the intermediate composite indices with the explained variance in the data.

After the construction of the composite index, the validity of the index was tested through correlation analysis (see Sect. 3.1). We constructed two quality of life indices using the same indicator variables, firstly with equal weighting and secondly by adding the indicators to derive at a sumscore index and we constructed a life satisfaction index. We then tested the correlation between the newly constructed composite index of quality of life and the three separate indices: the own constructed equally weighted composite quality of life index; the sumscore index; and the life satisfaction index (subjective well-being) derived from the dataset.

The newly constructed composite quality of life index was positively and moderately correlated with the life satisfaction index ($r = .49$, $p < .01$), and positively and strongly correlated with the own constructed equally weighted composite quality of life index ($r = .86$, $p < .01$) and the sumscore index ($r = .89$, $p < .01$). In Fig. 3 in Appendix 4, a comparison of the four indices is shown across the various demographic and socio-economic groups. The ordering of scores of all four indices is similar across the groups.

These validity checks, along with the robustness tests described earlier, suggest that the index is a valid and reliable one for measuring quality of life in South Africa. Furthermore, as an additional method to evaluate the validity of the new composite index we compared the results of the current study and previous studies when the indices are applied to the comparison of quality of life of different groups (see Sect. 6 and 7 for a discussion of the comparative findings where applicable).

6 Comparison of the Quality of Life Scores of Different Groups

The newly constructed composite quality of life index can be used to compare the quality of life scores of different demographic and socio-economic groups within the GCR. This disaggregated and comparative analysis is especially important in the light of the considerable diversity and heterogeneity within the GCR. Figure 2 shows the mean quality of life scores for the different groups. Quality of life is compared here by geographical area (comparing urban formal, urban informal and tribal-and-farming areas), income category, race, sex and age. The differences between groups are discussed further below.

To determine if these differences in the means of the quality of life scores between the groups are significant, we conducted *t* tests in cases where only two groups were compared and one-way between-groups ANOVAs where more than two groups were compared (see Table 4). After conducting one-way between-groups ANOVA tests we measured the actual differences in the means by calculating *eta squared* (an effect size statistic) and conducting post-hoc tests using the Tukey HSD test (Pallant 2007) (as set out in Sect. 3.2).

We now proceed to compare the quality of life scores and patterns between the various groups, including attempting to shed light on what lies behind the differences in scores. In each case, we begin by comparing the quality of life scores across the relevant groups, testing these differences as to whether they are significant using ANOVA tests, analysing the differences in the indicator variables between groups that can account for the

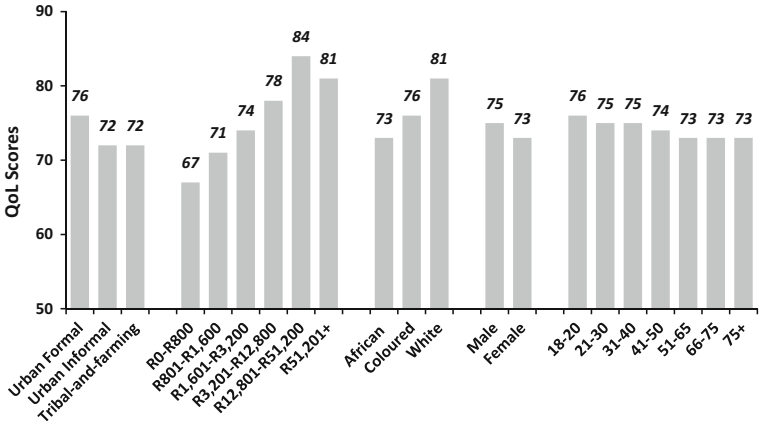

Fig. 2 Quality of life scores of different demographic and socio-economic groups. *Source*: Authors' calculations using GCRO data (GCRO 2009). *Note*: (1) Scores are shown as a percentage, with 100 % the maximum possible score. To attain a 100% score, each respondent within a group would have to score the highest value for each indicator within each dimension of quality of life. (2) Asians are excluded from the analysis due to the small number of the former, in order to ensure statistical robustness of the analysis. (3) QoL Scores = Quality of Life Scores

differences in scores across groups (taking into account the differences across groups in indicator variables included in the index, as well as the weighting of those variables in the index).

Comparing racial groups, Whites were found to have a higher mean quality of life score than Africans and Coloureds (see Fig. 2).[17] To test if these differences are significant we used a one-way between-groups ANOVA (see Table 4). We found a significant difference at the one percent level ($p < .01$) in the quality of life scores of the different race groups ($F(3) = 154.69, p = .000$). The difference between the mean scores using *eta squared* was 0.07, which using Cohen's (1988) rules of thumb is moderate. Post hoc comparisons using the Tukey HSD test indicated that the mean quality of life scores of all three groups were statistically different from one another. The difference is especially stark between Africans and Whites.

Table 5 shows the results for each variable included in the index for each racial group, as well as the weight of each variable in the quality of life index.[18] This allows for interpretation as to which of the variables may explain the overall differences in quality of life scores between the racial groups. Factors in the quality of life index for which there are especially large differences between racial groups (notably between Africans and Whites) as well as having high weighting in the index include 'water on premises', 'type of dwelling', and the main economic factors taken collectively ('income', 'satisfied with amount of money' and 'employment status'). 97 % of Whites reside in formal dwelling compared to 70 % of Africans, and 98 % of Whites have water on premises compared to 63% of Africans. The three indicators in the 'housing and infrastructure' dimension

[17] The number of Asians included in the sample were too small for robust statistical analysis, therefore Asians were excluded for the analyses in Sects. 6, 7.

[18] For reasons of space, Table 5 only shows racial groups. For the other group categories, similar tables are not shown, but relevant results are noted in the text; the additional tables are available from the authors on request.

Table 4 Results of Between-group One-way ANOVAs and *t* tests

	Sum of squares	*df*[a]	Mean square	F-statistic	P-statistic	Eta squared
Geographical area						.06
Between groups	29,052	2	14,526	49	.000	
Within groups	435,027	5590	78			
Total	464,079	5692				
Income groups						.03
Between groups	125,199	5	25,040	411	.000	
Within groups	341,956	5610	61			
Total	457,155	5615				
Race						.07
Between groups	35,679	3	11,893	149	.000	
Within groups	431,476	5612	77			
Total	467,155	5615				
Age						.01
Between groups	6070	6	1012	12,607	.000	
Within groups	455,569	5557	82			
Total	461,639	5563				

	t-statistic	*df*	Significance (2-tailed)	Mean difference	Std. error difference	Eta squared
t test for equality of means						
Sex	5.80	5614	.000	3	.25	.0005

Source GCRO data (GCRO 2009)

[a] *df* = degrees of freedom

collectively contribute almost one third of the overall weight of the quality of life index, and Whites have higher scores than Africans and Coloureds in all three variables in this dimension; this dimension contributes to the fact that Whites have an overall higher quality of life score than do Africans, as shown in Fig. 2. There are also very large differences between racial groups in the economic variables, although the collective weight of this dimension (15 % of the composite index) is less than the 'housing and infrastructure' dimension. The average income of African households is approximately R2200 per month compared to R9500 for Whites (GCRO 2009). Only 39 % of the African respondents reported being employed compared to 59 % of the White respondents.[19]

There are large difference between racial groups in education levels, but the weight of this dimension is comparatively low. Interestingly, the 'safety and governance' indicators of Africans were on average higher than that of Whites, especially pertaining to the governance indicators 'country is going in the right direction' and 'elections was free and fair', implying that Africans generally have greater confidence in the political situation than do Whites.

We now turn to the other group categories: geographical area; income groups; men and women; and age groups, and consider the differences in quality of life shown in Fig. 2. In terms of geographical area, we found that the differences in the means of the quality of life

[19] Employed is defined as a person who has worked in the last seven days before the interview.

Table 5 Frequencies on comparison of indicators of different races

Indicator	African	Coloured	White
Type of dwelling (0.09)*			
Informal	25.7	10.8	2.6
Formal	70.3	89.2	97.4
Electricity for lighting (0.11)			
No electricity	11.1	3.0	0.1
Have electricity	88.9	97.0	99.9
Water on premises (0.12)			
No water on premises	37.5	15.3	2.5
Have water on premises	62.5	84.7	97.5
Satisfied with time available for own things (0.08)			
Very dissatisfied	4.9	4.4	2.1
Dissatisfied	17.7	16.7	11.2
Neutral	18.1	20.2	12.4
Satisfied	48.1	43.3	51.5
Very satisfied	13.1	15.3	22.8
Satisfied with time available for friends (0.08)			
No Friends	2.9	1.5	1.6
Very dissatisfied	1.0	0.0	0.7
Dissatisfied	5.5	3.9	2.4
Neutral	21.0	25.1	13.5
Satisfied	53.3	51.2	48.0
Very satisfied	16.3	18.2	33.1
Satisfied with time available for family (0.06)			
Very dissatisfied	1.4	0.5	0.6
Dissatisfied	7.9	4.9	3.3
Neutral	11.4	9.4	5.5
Satisfied	53.9	55.2	44.0
Very satisfied	25.4	30.0	46.7
Income (0.05)			
R0–R800	17.0	5.5	1.3
R801–R1600	25.4	15.3	3.2
R1601–R3200	26.6	16.9	7.0
R3201–R12,800	24.8	33.3	34.3
R12,801–R51,200	5.3	23.5	45.8
R51,201+	0.9	5.5	8.5
Satisfied with amount of money (0.05)			
Very dissatisfied	27.3	17.2	9.5
Dissatisfied	40.5	39.9	19.5
Neutral	17.6	23.2	24.9
Satisfied	12.4	14.8	34.2
Very satisfied	2.3	4.9	11.9
Employment Status (0.05)			
Unemployed	61.4	47.0	41.0

Table 5 continued

Indicator	African	Coloured	White
Employed	38.6	53.0	59.0
Years of education (0.02)			
0 years	4.5	1.1	0.3
1–3 years	3.7	0.5	0.1
4–6 years	7.8	4.3	0.4
7–8 years	20.4	12.3	4.5
9–11 years	27.6	26.2	18.3
12 years	28.9	42.2	41.9
Tertiary training	7.0	13.4	34.6
Satisfaction with health (0.08)			
Very dissatisfied	2.0	2.0	0.2
Dissatisfied	12.2	8.4	6.6
Satisfied	51.4	55.7	49.2
Very satisfied	34.5	34.0	43.9
Health prevent you from taking part in social activities (0.08)			
Always	4.5	4.4	2.9
Some of the time	23.4	26.6	20.3
Hardly ever	22.1	19.2	21.7
Never	50.0	49.8	56.7
Feel safe during the day (0.02)			
Very unsafe	5.8	4.4	4.8
Bit unsafe	9.3	7.4	8.8
Neutral	7.2	6.4	9.6
Fairly Safe	41.0	48.3	44.7
Very safe	36.6	33.5	32.1
Country is going in the right direction (0.05)			
Strongly disagree	5.9	9.4	16.7
Disagree	12.7	13.8	21.6
Neither agree or disagree	19.1	22.2	21.8
Agree	50.7	44.6	29.8
Strongly agree	11.6	9.9	10.0
Elections was free and fair (0.05)			
Strongly disagree	2.9	4.9	7.9
Disagree	4.2	4.4	11.3
Neither agree or disagree	9.2	11.3	18.9
Agree	57.8	56.2	45.3
Strongly agree	25.9	23.2	16.6

Source Authors' own analysis using GCRO data (GCRO 2009)

* The weight of the variable in the Quality of Life Index

scores of tribal-and-farming residents, urban informal residents and urban formal residents were significant at the one percent level ($p < .01$) with an *F-statistic* of 186.7 ($df = 2$) and p value = .000. Measuring the actual differences between the mean quality of life scores of the different geographical groups it was found that the differences were moderate

considering the effect size statistic (*eta squared* = .06). After administering post-hoc comparisons using the Tukey HSD test it was found that there was a significant difference between the mean quality of life score of urban formal residents and tribal-and-farming residents and between urban formal residents and urban informal residents, yet no evidence was found of a significant difference in the mean quality of life scores of tribal-and-farming residents and urban informal residents. This implies that the quality of life experienced on average by tribal-and-farming residents and urban informal residents is similar, but lower than the quality of life of urban formal residents.

By way of interpretation, the differences in quality of life scores between tribal-and-farming areas and urban informal areas on the one hand, and urban formal areas on the other hand, are accounted for primarily by the economic factors in the index. There are large differences between the geographical areas in all three economic indicators ('income', 'satisfied with amount of money' and 'employment status'). The average monthly income of respondents in tribal-and-farming areas is R1,200 and in informal urban areas is R2400, compared to R4800 for respondents in formal areas. In terms of employment status, only 47 and 55 % of respondents in tribal-and-farming areas respectively are employed, compared to 70 % of the urban formal group. There are also significant differences in education levels (although the weight of this variable in the composite index is comparatively low) and in the housing and infrastructure variables, especially in access to water on premises (14 % for tribal-and-farming areas, 61 % for urban informal areas and 78 % for urban formal areas). The big differences in the housing and infrastructure dimension of quality of life were also reflected in the CATPCA (see Sect. 5) showing that this dimension explained the most variance in the data.

Comparing the mean quality of life scores of different income groups, using a one-way between-groups ANOVA, we found a significant difference at the one percent level ($p < .01$) between the six income groups ($F(5) = 410.80$, $p = .000$) (see Table 4). However, the actual differences in the mean quality of life scores of the income groups were relatively small with the effect size measured by *eta squared* equal to .03. After conducting post-hoc tests using the Tukey HSD test we found the mean quality of life scores of all the income groups to be significantly different from one another except for the highest last two income groups (including respondents earning between R12,801 and R51,200 and respondents earning more than R51,201), which did not differ significantly.

Considering the mean values of the quality of life scores, the findings show an 'inverted-U' relationship between income and quality of life. This result corresponds with the findings of Easterlin (1974) concerning subjective well-being: as income increases, subjective well-being increases up to a point, after which it declines. It is noteworthy that our findings on quality of life, a multidimensional measure, accord with Easterlin's findings on subjective well-being, a single-dimensional measure. Descriptive analysis of the data suggests that the dip in quality of life for the top bracket of income earners (R51,201) is associated with the lower levels of 'satisfaction with time spent with family and friends', and 'satisfaction with leisure time'.

The positive relationship between income and quality of life up to a point is arguably related to the important role of income in financing the meeting of basic needs and in enhancing capabilities. Low income is also associated with unemployment (or underemployment), as well as with low levels of education, both of which have a negative association with quality of life. Only 4 % of the lowest income groups had tertiary training compared to almost 62 % among the income group earning more than R51,201 per month. Comparing dwelling types and infrastructure (a dimension contributing almost a third of the overall weight of the composite index), 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. Probing deeper, the quality of life survey (GCRO 2009) in an additional question provided certain explanations for people's life satisfaction of which they had to choose the most appropriate option. The options most often selected by the lowest income groups were a lack of income, high costs of living and a shortage of employment opportunities (GCRO 2009). In comparison 75 % of the wealthier respondents reported being either satisfied or very satisfied with life (GCRO 2009).

Turning to the differences in quality of life scores between women and men, using an independent–samples t-test (see Table 4) we found that there was a significant difference in the quality of life scores of men (Mean (M) = 71, Standard Deviation (SD) = 9.22) and women (M = 68, SD = 9.00) with the t-stat statistic (df = 5614) equal to 5.80 (p = .000, two-tailed). The magnitude in the differences in the means (mean difference = 3) evaluated using Cohen's (1988) rules of thumb was very small (*eta squared* = .0005). From our analysis of the GCRO data, the differences in the mean quality of life scores can primarily be explained by the 'employment' and 'income' indicators. Men are 20 % more likely to be employed than women. Part of the explanation for this could be the lower levels of female education, as well as reasons listed by female respondents for not finding employment such as that they had to look after children or family members and that they live too far away from employment opportunities (GCRO 2009) as well as gendered segmentation and possible discrimination in the labour market. The average income earned by women is slightly less than that of men (GCRO 2009). There were not significant differences between men and women in the 'housing and infrastructure' indicators, unsurprisingly since men and women typically reside together in the same dwellings. That this dimension has the highest weighing in the composite index, and the relatively small differences between women and men in this dimension, partly accounts for the small (albeit significant) difference in men's and women's quality of life scores.

Finally, the mean quality of life score is fairly consistent across the different age groups, with a slight decline in quality of life for older respondents (see Fig. 2). To test if the differences between the age groups were significant, we conducted a one-way, between-groups ANOVA (see Table 4). We found a significant difference in the quality of life scores at a 1% level ($p < .01$) between the seven age groups ($F(6) = 12.340$, $p = .000$). Notwithstanding the finding of significant difference between the mean quality of life scores of the different age groups, the actual difference in the mean scores were quite small. The effect size, calculated using *eta squared*, was 0.01. Post-hoc comparisons using the Tukey HSD test showed the mean quality of life score of the age group between 18 and 20 years to be significantly different from all the other age groups. Furthermore, we found significant differences between the age groups 21–50 years and those older than 51 years.

This finding of similar quality of life across age groups differs from the common finding in the literature of an U-shaped relationship between age and well-being, with the young and the old having higher levels of well-being than middle-aged persons (Diener et al. 1999). The finding that older persons (over the age of 51) in the Gauteng City-Region have slightly lower levels of quality of life might reflect the deteriorating health and difficult economic situations of many of the older respondents (GCRO 2009). Rossouw and Naudé (2008) found that the well-being of people over the age of 46 decreases, similar to our finding. The finding of lower levels of quality of life of older people was also shown in the results of a study by Hansen and Slagsvold (2012) for Norway. Although the overall quality of life scores indicated by the

composite index are fairly consistent across age groups, there are differences in the various dimensions of the index. For example, the income for the age group 31–50 is higher than that of the young and the old, whereas the health of the younger age groups is superior to that of the older age groups. 21 % of the young were dissatisfied with the time they had to themselves compared to 13 % of the oldest two groups. 61 % of the youngest group felt that the country is going in the right direction compared to 54 % of the oldest two groups. It appears that the difference between the different dimensions of quality of life across the age groups balance out in the composite quality of life index.

To summarise, it was shown that there are significant differences in the mean quality of life scores of the different demographic and socio-economic groups. The groups with the lowest quality of life scores within each category were respondents living in urban informal and tribal-and-farming areas, respondents in households earning less than R800 per month, Africans, females and people over the age of 51.

Finally, we undertook regression analyses to determine if certain background- and socio-economic characteristics of the respondents are likely predictors of their levels of quality of life, and of particular relevance here, to investigate how these relationships vary across groups.

We selected relevant background and socio-economic characteristic variables available in the data, but which were not included as indicators in the composite index. In the selection of the variables we were also guided by previous research including among other Cummins (1996) and Posel and Casale (2011). The variables included as explanatory variables in the regressions are: marital status, sex, race, age categories, geographical area, crime, ownership of a television (as a proxy for asset ownership), perceived social status, satisfaction with dwelling, and life satisfaction (see Appendix 1 in Table 8 for the coding of the variables).

The regression was run initially as a baseline specification for the full sample, and subsequently separately for the various demographic and socio-economic groups.[20] The baseline specification is as follows:

$$CIQoL_i = a_i + b_m X_{mi} + u_i$$

where CIQoL$_i$ represents the Quality of Life score of each respondent i(i=1 … n), a_i is the constant, b_m provides the coefficients that indicate the relative importance of different contributors to quality of life, X_{mi} is a vector of m explanatory background and socio-economic characteristics variables of the respondents i, and u_i is an error term including all the factors that were not captured in the quality of life equation.

Table 13 (Appendix 5) presents the results for the overall sample in the baseline regression and for the race groups and geographic areas; Table 14 (Appendix 5) for the income categories and sexes; and Table 15 (Appendix 5) for the various age groups. Where results are compared across groups classified according to a variable that is an explanatory variable in the baseline specification, that variable is omitted from the regressions run for those groups (for example the omission of race as an explanatory variable in the three regressions run separately for racial groups, presented in Table 15).

The regressions were estimated using OLS, with the newly constructed quality of life index as the dependent variable and the previously mentioned background and socio-economic characteristic variables as explanatory variables. We conducted diagnostic tests and found no evidence of multicollinearity,[21] the residuals were normally distributed and

[20] See Table 1.

[21] The tolerance value of all the variables were >0.1 and the Variance Inflation Factor (VIF) of all the variables were <2. On inspection of the correlation matrix no signs of multicollinearity were found.

from the standardised residuals scatterplot it is apparent that the residuals are homoscedastic and linear as well as being independent (Durbin Watson Statistic = 1.8). F-statistics indicate satisfactory goodness of fit of the regressions.

From the baseline regression, run across the entire sample, the $R^2 = .34$, thus 34 % of the variance in the quality of life composite index was explained by the independent variables. This R^2 seems relatively low, however it compares well to other studies in the literature which also find a R^2 of approximately 30 % (Posel and Casale 2011). The following variables are significant at either the 1 or 5 % level in explaining quality of life: marital status (being married having a positive effect); sex (being male having a positive effect); being white (having a positive effect); being aged above 40 (falling into the categories 41–50 or over 50 both having a negative effect); living in an urban formal area (having a positive effect); ownership of a television (as a proxy for asset ownership, having a positive effect); and perceived social status, satisfaction with dwelling, perceptions of improvement in the local crime situation, and life satisfaction (all having a positive effect) (see Table 13 in Appendix 5).

The regressions run separately for the race groups have a R^2 of between .24 and .30 implying that the independent variables in the regressions run separately for each race group explained approximately a third or less of the variance in the composite index of quality of life, similar to the baseline model. The results in Table 13 reflect how the various characteristics differentially affect quality of life for the different race groups. Perceived social status, satisfaction with dwelling and life satisfaction have a positive and significant effect on quality of life across all race groups. This is consistent with the findings from the CATPCA analysis showing housing and economic factors are important predictors of quality of life. Being older than 51 is negatively related to the quality of life of all the race groups. Some variables are significant in explaining the quality of life of Africans only, such as having a TV. The overwhelming majority of Whites can afford a TV, so for households who do not have a TV this is typically a conscious lifestyle choice, whereas for Africans ownership of a TV is generally likely to be associated with affordability and having this type of asset can contribute to higher levels of quality of life.

From the regressions run separately for the three geographical areas (see Table 13), the R^2 is approximately 0.3 in each of the cases, implying that similar to the regressions based on different race groups the independent variables in the regressions based on geographical areas explain approximately a third of the variance in the dependent variable. Regarding the different geographical areas, some important commonalities but also interesting differences can be observed in terms of the determinants of quality of life. Being aged above 50, ownership of a TV, perceived social status, and life satisfaction are significant for all three areas. However, being married and being male are significant predictors of quality of life only in urban formal areas; this may be indicative of the more heterogeneous ways in which sex and marital status affect quality of life in other areas.

Comparing the results across income groups (see Table 14 in Appendix 5), the R^2 is significantly higher for people earning an income of more than R51,000 than for people earning less, thus the independent variables for the higher income groups explain more of the variance in the dependent variable compared to the other income groups. Living in an urban formal area and being satisfied with life are significant and positive in predicting quality of life for all income categories. Being older than 51 years of age is significant and negatively related to quality of life of all income groups. An interesting finding is that the economically-related characteristics of TV ownership and social status are only significant for the lower-middle income categories. This is consistent with the findings from the CATPCA analysis, suggesting that as the basic economic needs are largely addressed for upper-income groups, other factors become more important in determining quality of life.

The overall pattern of the econometric results is similar for female and male respondents (refer to Table 15 in Appendix 5) the R^2 is also in line with the base line model. Indeed, the overall quality of life scores were similar for women and men. One interesting difference is that being married has a positive and significant effect on men's quality of life, whereas the coefficient is positive but not significant for women. This provides an interesting perspective on the relative benefits of marriage for enhancing men's and women's quality of life. A second interesting difference between men and women is that the negative effects of age on quality of life appear to begin earlier for women than for men, with the dummy for age between 41 and 50 being significant for women but not for men, albeit negative in both cases.

Finally, in terms of age groups (see Table 15) the R^2 corresponds with the base line model and the only variable that is strongly significant across all age groups is life satisfaction. The coefficient for marital status is highest, and is also strongly significant, for the 75+ age group, pointing to the central importance of companionship for the quality of life of the elderly. For most age groups, staying in an urban formal area, having a higher level of perceived social status, and having an improved crime situation, are significant (at the 1 or 5 % levels) related to quality of life (see Table 15). Being White predicts higher levels of quality of life than being of any other race group across all the age groups, except for the age group 18–20. Satisfaction with dwelling is not significant in predicting respondents' quality of life between the ages 31–50, but significant for younger and older respondents.

The econometric analysis adds to the CATPCA and other empirical analysis presented here, and the findings are broadly consistent. The regressions for separate groups are illuminating as to how the characteristics modelled as explanatory variables differentially affect the quality of life of different groups. This links to the analysis in the next section.

7 Comparison of the Order of the Extracted Components (Dimensions of Quality of Life) of Different Groups

The previous section compared quality of life for different groups according to the constructed composite quality of life index. Here, we run CATPCAs separately for the various demographic and socio-economic groups within the GCR using the same indicators and number of extracted components (five), as was used in earlier analyses (see Sect. 5). This is to determine which dimensions of quality of life explain the most variance in the data for each group. The extracted components for each group are shown in Table 6. The ranking of the components for each group shows which component explains the most variance in the data for that group.

For most groups, the most variance in the data was explained by the 'housing and infrastructure' component. This implies that the indicators 'type of dwelling', 'water on premises' and 'electricity for lighting' (which make up this component) vary considerably within these groups.

For the groups Coloureds, Whites, respondents with a monthly income between R3,201 and R51,200 and people older than 31 and younger than 50 and urban formal dwellers, the dimension 'social relationships' explained the most variance in the data. This implies that the indicators '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. As 'social relationships' explained most of the variance in the data, the other dimensions such as 'housing and infrastructure' were more uniform among these respondents, with most the respondents living in formal housing and having electricity and water. It is possible that the basic needs of many members of these groups have already largely been met.

Table 6 The order of the first extracted components

Group	Components				
	1	2	3	4	5
Geographical area					
Urban formal	Social relationships (18.98)	Socio-economic status (12.89)	Health (9.91)	Safety and governance (7.98)	Housing and infrastructure (7.42)
Urban informal	Housing and infrastructure (17.96)	Social relationships (18.90)	Socio-economic status (9.50)	Health (8.12)	Safety and governance (7.70)
Tribal and farming	Housing and infrastructure (17.56)	Socio-economic status (15.45)	Social relationships (10.66)	Health (9.69)	Safety and governance (7.38)
Race					
Africans	Housing and infrastructure (23.26)	Social relationships (15.00)	Socio-economic status (8.47)	Health (9.63)	Safety and governance (6.97)
Coloureds	Social relationships (19.91)	Socio-economic status (10.66)	Housing and infrastructure (8.48)	Health (7.97)	Safety and governance (7.96)
Whites	Social relationships (16.96)	Socio-economic status (11.54)	Safety and governance (9.6)	Health (8.00)	Housing and infrastructure (6.97)
Income					
R0–R800	Housing and infrastructure (17.65)	Social relationships (13.15)	Health (8.75)	Socio-economic status (7.98)	Safety and governance (6.68)
R801–R1600	Housing and infrastructure (16.07)	Social relationships (13.00)	Health (11.15)	Socio-economic status (9.50)	Safety and governance (8.39)
R1601–R3200	Housing and infrastructure (16.53)	Social relationships (12.45)	Health (11.02)	Socio-economic status (9.44)	Safety and governance (7.73)
R3201–R12,800	Social relationships (16.56)	Housing and infrastructure (13.10)	Socio-economic status (9.65)	Health (9.65)	Housing and infrastructure (12.23)
R12,801–R51,200	Social relationships (19.65)	Safety and governance (9.13)	Socio-economic status (8.36)	Health (9.65)	Housing and infrastructure (12.23)
R51,201+	Socio-economic status (8.36)	Safety and governance (9.13)	Social relationships (19.65)	Health (9.65)	Housing and infrastructure (12.23)

Table 6 continued

Group	Components				
	1	2	3	4	5
Age					
18–20	Housing and infrastructure (18.08)	Social relationships (14.82)	Socio-economic status (8.42)	Health (8.00)	Safety and governance (7.68)
21–30	Housing and infrastructure (19.06)	Socio-economic status (13.82)	Social relationships (8.34)	Health (8.01)	Safety and governance (7.82)
31–40	Socio-economic status (21.68)	Social relationships (14.64)	Housing and infrastructure (8.70)	Safety and governance (8.10)	Health (7.20)
41–50	Socio-economic status (21.08)	Social relationships (13.87)	Housing and infrastructure (9.46)	Safety and governance (8.73)	Health (7.60)
51–65	Social relationships (20.76)	Socio-economic status (13.06)	Housing and infrastructure (10.59)	Health (8.73)	Safety and governance (9.60)
66–75	Housing and infrastructure (19.05)	Social Relationships (13.67)	Health (9.66)	Socio-economic status (8.75)	Safety and governance (6.89)
75+	Housing and infrastructure (20.62)	Social Relationships (12.01)	Health (10.8)	Socio-economic status (9.2)	Safety and governance (7.1)
Sex					
Male	Housing and infrastructure (20.67)	Socio-economic status (16.04)	Social relationships (8.81)	Safety and governance (7.97)	Health (7.67)
Female	Housing and infrastructure (19.00)	Socio-economic status (14.08)	Social relationships (9.70)	Health (8.60)	Safety and governance (7.40)

Note (1) Values in parentheses indicate the percentage of total explained variance of each component, (2) Asians are excluded from the analysis due to small numbers, to ensure statistical robustness of the analysis

Source Authors' own analysis using GCRO data (GCRO 2009)

The component 'socio-economic status' ranked in the top three positions for most groups. While not explaining the most variance in the data for each group, this component explained a considerable proportion of the variance for most groups. This suggests that this dimension has an important effect on the quality of life across the population of the GCR, and varies considerably among all respondents. Economic reasons were also noted as among the main determinants of the subjective well-being of those persons dissatisfied with life (GCRO 2009).

Based on the calculated quality of life scores of the different demographic and socio-economic groups (presented in Sect. 6), and the findings here as to which components explained the largest proportion of the variance in the data, we found that the groups in which 'social relationships' explained the most variance in the data were generally the groups with the highest quality of life scores. Thus we could argue that to improve the well-being of people that had higher quality of life scores the 'social relationship' component should be addressed (although this component is largely beyond the reach of public policy interventions).

The groups experiencing lower levels of quality of life corresponded with those groups in which 'housing and infrastructure' explained the most variance in the data. Thus a case could be made that to improve the well-being of people with lower levels of quality of life—who should arguably be the primary concern of policymakers—the component 'housing and infrastructure' should be addressed. This component included the indicators 'housing type', 'electricity for lighting' and 'water on premises', though it could be assumed that any indicator closely related to these indicators such as 'sanitation' should also be addressed. Furthermore, as 'socio-economic status' explained a considerable portion of the variance in the data 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 well-being of all people in the GCR (see Stroup and Stephen 1992).[22]

8 Conclusion

South Africa, including the GCR, has a highly diverse population. In addition, inequality is extremely high by international standards. Improving people's quality of life of the people and addressing the wide disparities in the society have increasingly been recognised as important public policy objectives in South Africa, as is the case internationally. Making progress in this regard, and being able to quantify and evaluate such progress, requires appropriate measurement of quality of life. This is not straightforward, given the multi-dimensional nature of the concept of quality of life. Therefore a composite quality of life index is required, which can assess each dimension of quality of life, and weight and aggregate these dimensions in a meaningful and scientific way.

This article addresses the need for a composite index to measure and analyse quality of life in the GCR. The article contributes to the literature on the measurement of quality of life by constructing a composite index, which uses rigorous and appropriate methodologies, is concise, objectively weighted and includes the dimensions of quality of life as defined in the Stiglitz Report (2009). The application of the index to compare and analyse quality of life across various groups in the GCR, contributes to a deeper understanding of the quality of life of the population of the GCR. Furthermore, the application of the index allows for the identification of specific socio-economic and demographic groups with

[22] Results from the SAARF development index (2011) also noted the importance of the 'socio-economic status' dimension in improving South Africans' well-being and highlighted the need for higher levels of employment.

lower quality of life as well as the specific dimensions which most contribute to the lower levels of quality of life. These findings inform policy so that the well-being of groups with lower levels of quality of life can be improved and a more equitable society be created. This is the first measure of this type constructed for the GCR or used in South Africa and makes a meaningful contribution to understanding the complex society of South Africa.

To construct the index we extended the method developed by Nicoletti et al. (2000) to construct composite indices and applied it to the measurement of quality of life. The methodology of Nicoletti et al. (2000) was adapted and extended using CATPCA, which is appropriate for the analysis of categorical data, to objectively weight the index according to the explained variance of the data. The method incorporates not only the first principal component to weight the index, but also additional components to achieve a better representation of the data. Guided by the Stiglitz Report (2009) and existing composite quality of life indices as well as data availability, we identified the dimensions and the indicators most often included in composite indices to measure quality of life. Using CATPCA we selected fifteen indicators from the preliminary identified indicators. The final selection of indicators included both objective and subjective as well as economic and non-economic indicators and explained, relative to other sets of indicators, the most variance in the data and thus gave the best representation of the data. Using the selected indicators in a CATPCA, five components were extracted, each representing a different dimension of quality of life. These components accorded with the dimensions of quality of life found in the literature and were designated 'housing and infrastructure', 'social relationships', 'socio-economic status', 'health' and 'safety and governance'. Corresponding to the extracted components representing the different dimensions of quality of life we constructed five intermediate composite indices. These intermediate composite indices were aggregated by weighting each according to the percentage of variance explained by the component. The index was found to be robust to various sensitivity and validity checks.

Comparing the quality of life of the sub-groups in the GCR we found the quality of life scores of the following demographic and socio-economic groups within the respective categories to be lower than the other: tribal-and-farming residents, urban informal dwellers, low income earners, females, Africans and older respondents. Running regression analyses, we found various background and socio-economic characteristics significant in predicting the quality of life of the different groups. Being married, male, white, younger, staying in an urban area, being satisfied with your dwelling and life and having a higher level of perceived social status predicts higher levels of quality of life.

Next, comparing the order of the extracted components for each demographic and socio-economic group to determine which dimension contributed most in explaining the variance in quality of life of the specific group. We ran a CATPCA separately for each separate group. The first extracted component and the order of subsequent components varies across the groups. For African respondents, tribal-and-farming communities, urban informal residents, the majority of age groups, both sexes and low income earners, the component 'housing and infrastructure', which represents a basic need, explains the most variance in the data and therefore has the highest weight in their composite indices. For White and Coloured respondents as well as for high income earners and respondents in urban formal areas, the dimension designated 'social relationships', which is a higher order need, explains the most variance in the data. Basic needs, such as shelter and access to water, are already largely fulfilled for these groups. Furthermore, the dimension 'socio-economic status' represents a considerable proportion of the explained variance of all the groups.

Considering together the results from the different parts of the empirical analysis yields interesting new insights into understanding the South African society in the GCR, not revealed in previous literature. In the groups with relatively low overall quality of life scores, the dimension 'housing and infrastructure' explained the most variance in the data.

Therefore it might follow that to improve the quality of life of these groups the indicators included in this dimension – 'type of dwelling', 'electricity for lighting' and 'water on premises' – as well as the indicators closely associated with these such as 'sanitation' and 'electricity for cooking' should be prioritised. This is especially important in the case of residents in tribal-and-farming communities, in which the access to water on premises and electricity for lighting is much lower than in other geographical areas.

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. This implies that the indicators 'time with family and with friends' as well as 'leisure time'—specifically the perceived lack of adequate social time—are important to the quality of life of these groups. Within all the groups the dimension 'socio-economic status' explained a considerable proportion of the variance in the data. This dimension was identified as contributing to lower levels of quality of life. Therefore we can argue that improved 'socio-economic status' can directly improve the well-being of all groups of people in the GCR.

In addition to shedding light on quality of life issues in South Africa, this study sets out an approach to the measurement and analysis of *quality* of life that can be helpful in similar studies in other countries or regions thereof.

Acknowledgments We thank the Gauteng City-Region Observatory (GCRO) for access to the GCRO Quality of Life Survey data set and financial support. We also thank Economic Research Southern Africa (ERSA) for their financial support.

Appendix 1: Recoding of Nominal Indicators

The recoding of the nominal indicators was based on the guidelines provided in the Reconstruction and Development Programme (1996). This policy was developed by the first democratic government to address South Africa's development challenges. It set out guideline standards for service delivery, which are used here (See Tables 7, 8).

Table 7 Coding of indicators

Indicator	Type of indicator	Coding	Description	Explanation for method of coding
Type of dwelling	Nominal	1	Informal dwelling	All formal structures were coded as 2 and informal structures were coded as 1
		1	Traditional dwelling	
		2	House or formal structure	
		2	Flat	
		2	Town/cluster/ semi-detached	
		2	Unit in retirement village	
		2	Room/flatlet in main dwelling	
		2		
		2	Hostel	

Table 7 continued

Indicator	Type of indicator	Coding	Description	Explanation for method of coding
Water on premises	Nominal	1	Water not on premises	If water was accessible on the yard it was coded as 2, if not, it was coded as 1
		2	Water in dwelling	
		2	Yard tap	
Electricity used for lighting	Nominal	1	Gas/LPG	More advanced energy sources were coded as 2 and the other as 1
		1	Paraffin	
		1	Wood	
		1	Candles	
		2	Solar energy	
		2	Electricity	
Employment status	Nominal	1	Unemployed	If a person did any type of paid work in the past 7 days it was coded as 2 otherwise as 1
		2	Employed	
Education	Ordinal	1	0 years	Years of education
		2	1–3 years	
		3	4–6 years	
		4	7–8 years	
		5	9–11 years	
		6	12 years	
		7	Tertiary training	

Source GCRO dataset (2009)

Table 8 Coding of variables included in regression analyses

Variables	Coding
Marital status	
Married/have a partner	1
Single	0
Sex	
Male	1
Female	0
Race	
African	Reference group
White, other	1,0
Coloured, other	1,0
* Age	
18–40	Reference group
41–50, other	1,0
51+, other	1,0
Geographical area	
Formal	1

Table 8 continued

Variables	Coding
Urban informal and tribal-and-farming	0
TV	
Have a TV	1
No TV	0
Perceived social status	
Poor	1
Working class	2
Middle class	3
Upper class	4
Satisfaction with dwelling	
Very dissatisfied	1
Dissatisfied	2
Neutral	3
Satisfied	4
Very satisfied	5
Crime	
Crime in area—worse	1
Crime in area—same	2
Crime in Area—improved	3
Life Satisfaction	
Very dissatisfied	1
Dissatisfied	2
Neutral	3
Satisfied	4
Very satisfied	5

Source GCRO dataset (2009)

* The age variable was coded to reflect the literature stating that the quality of life of people decreases between approximately the ages of 41 to 50

Appendix 2

See Table 9.

Table 9 Preliminary list of indicators from which final selection were made

Indicator	Type	Min	Max	Mean	Mode
Health					
Satisfaction with health	Ordinal	1	4	3.2	3
How often does health prevent you from doing your daily work	Ordinal	1	4	3.2	4
How often does health prevent you from taking part in social activities	Ordinal	1	4	3.2	4
Education					
Years of education	Ordinal	1	6	4.1	5

Table 9 continued

Indicator	Type	Min	Max	Mean	Mode
Everyday Activities (including work and housing)					
Employment status	Dichotomous	1	2	1.6	2
Income	Ordinal	1	6	3.0	4
Satisfaction with standard of living	Ordinal	1	5	3.2	4
Satisfied with amount of money	Ordinal	1	5	2.4	2
Perceived social status	Ordinal	1	5	1.8	2
Life satisfaction	Ordinal	1	5	3.2	4
Satisfied with time available for own things	Ordinal	1	5	3.5	4
Type of dwelling	Dichotomous	1	2	1.8	2
Water on premises	Dichotomous	1	2	1.7	2
Electricity used for lighting	Dichotomous	1	2	1.9	2
Type of sanitation	Dichotomous	1	2	1.5	2
Satisfaction with dwelling	Ordinal	1	5	3.2	4
Participation in the political process					
Satisfied with local government	Ordinal	1	5	2.9	3
Politics is not a waste of time	Ordinal	1	5	3.7	4
Judiciary is free	Ordinal	1	5	3.1	4
Country is going in the right direction	Ordinal	1	5	3.4	4
Election was free and fair	Ordinal	1	5	3.9	4
Social relationships					
Satisfied with time available for family	Ordinal	1	5	4.0	4
Satisfied with time available for friends	Ordinal	1	6	4.8	5
Satisfied with marriage	Ordinal	1	5	4.3	5
Safety from bodily harm					
Feel safe in area where you live in the day	Ordinal	1	5	3.93	4
Feel safe in area where you live after dark	Ordinal	1	5	2.28	3
Feel safe at home	Ordinal	1	5	4.38	5

Source Authors' calculations of the GCRO dataset (2009)

Appendix 3

See Table 10.

Table 10 Component correlation matrix

Component	1	2	3	4	5
1	1.00	.05	.00	.19	.28
2	.05	1.00	.09	.20	.15
3	.00	.09	1.00	.03	.00
4	.19	.20	.03	1.00	.14
5	.28	.15	.00	.14	1.00

Source Authors' calculations of the GCRO dataset (2009)

Appendix 4

See Tables 11, 12; Fig. 3.

Table 11 Robustness check using different methods of treating missingness

Different methods of treating missingness	KMO statistic	Bartlett's test (Approx. Chi square)	% of variance explained	Order of extraction of dimensions	% of variance explained by each component
Excluding cases pairwise	.714	17,8512***	59.89	Housing and infrastructure	19.27
				Social Relationships	12.27
				SES	10.28
				Health	8.82
				Safety and Governance	7.56
Excluding cases listwise	.714	17,824***	59.87	Housing and infrastructure	19.43
				Social Relationships	13.13
				SES	10.01
				Health	9.81
				Safety and Governance	7.48
Data imputation (mode)	.714	17,849***	59.89	Housing and infrastructure	19.44
				Social Relationships	13.14
				SES	10.02
				Health	9.81
				Safety and Governance	7.48

Source Authors' calculations using the GCRO dataset (2009)

Table 12 Order of extracted components based on the rotated component matrix of a dataset similar to the final set of data

	Extracted components				
	1	2	3	4	5
Type of dwelling	**.73**	−.04	.07	.05	.00
Electricity for lighting	**.80**	.02	.05	.00	−.03
Sanitation	**.77**	.00	.09	.05	.08
Satisfied with marriage	−.07	**.71**	.13	−.09	−.11
Satisfied with time available for own things	.00	**.67**	.05	−.09	.31
Satisfied with time available for friends	−.05	**.83**	.00	.18	.02
Income	.29	.21	**.73**	.04	.01

Table 12 continued

	Extracted components				
	1	2	3	4	5
Satisfied with standard of living	.30	.30	**.52**	.09	.19
Employment status	−.10	.04	**.81**	−.02	.10
Years of education	.19	.09	**.55**	.37	−.08
Satisfaction with health	.00	−.07	.17	**.82**	.04
Health prevent you from doing your daily work	.15	−.10	−.02	**.83**	.03
Feel safe at work	.00	.12	.15	.15	**.61**
Politics is not a waste of time	.04	.13	−.03	.17	**.58**
Satisfaction with local government	.15	.10	.02	.07	**.69**
Explained variance (Eigenvalue)	2.88	1.79	1.45	1.20	1.08
Total percentage of explained variance	.34	.21	.17	.15	.13

Note Varimax rotation was used. Factor loadings in bold indicate the highest factor loadings on a specific component

Source GCRO dataset (GCRO 2009)

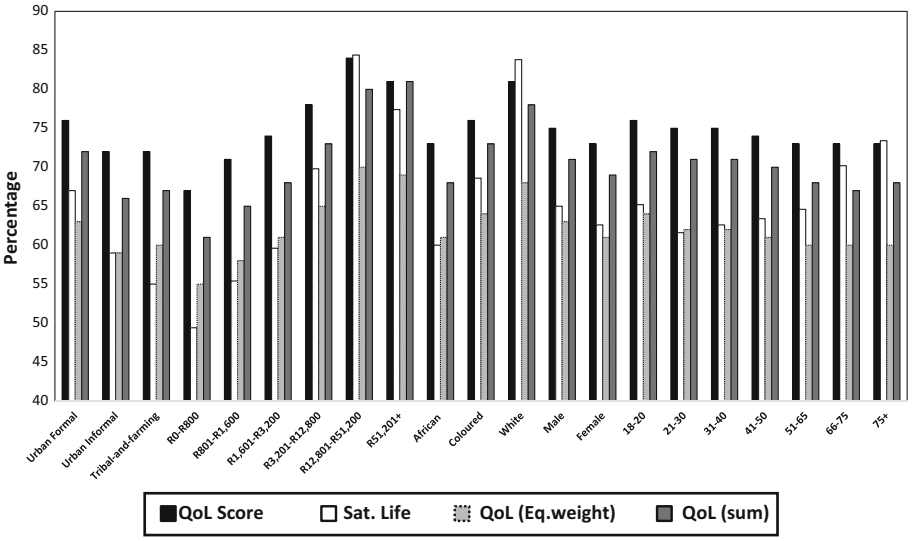

Fig. 3 Comparison of the newly created Quality of Life Score, the Life satisfaction Index and the Quality of Life Score (Equal weighting) and the Quality of Life Sumscore. *Source*: Authors' calculations using GCRO data (GCRO 2009). *Note*: (1) Scores are shown as a percentage, with 100 % the maximum possible score. (2) Asians are excluded from the analysis due to their small number in the dataset, in order to ensure statistical robustness of the analysis

Appendix 5: Regression Results

See Tables 13, 14 and 15.

Table 13 Regression results pertaining to the baseline regression, race and geographical area

	Baseline	Race			Geographical Area		
		African	Coloured	White	Urban formal	Urban informal	Tribal-and-farming
(Constant)	58.10*** (.45)	58.36*** (.48)	52.68*** (3.96)	57.33*** (2.79)	60.72*** (.68)	57.00*** (.70)	61.04*** (2.10)
Marital status (married = ref)	.50** (.24)	.44* (.26)	.82 (1.61)	.61 (.77)	.78** (.31)	.22 (.41)	−.01 (1.23)
Sex (male = ref)	.64** (.20)	.61*** (.22)	−.74 (1.21)	1.40** (.56)	.74*** (.26)	.25 (.35)	1.76* (.97)
White (African = ref)	2.10*** (.36)	–	–	–	2.10*** (.37)	5.81*** (1.80)	.55 (2.49)
Coloured (African = ref)	.23 (.59)	–	–	–	1.17* (.68)	−1.88 (1.20)	−2.06 (3.83)
Age 41–50 (18–40 = ref)	−.27*** (.06)	−.29*** (.07)	−.18 (.35)	−.23 (.18)	−.29*** (.08)	−.22* (.11)	−.27 (.33)
Age 51+ (18–40 = ref)	−3.34*** (.25)	−3.32*** (.28)	−6.32*** (1.62)	−3.46*** (.63)	−3.28*** (.32)	−3.08*** (.46)	−5.51*** (1.43)
Urban formal (urban formal = re f)	1.94*** (.22)	1.88*** (.22)	4.48*** (1.38)	.06 (1.27)	–	–	–
TV (Have a TV = ref)	2.36*** (.29)	2.50*** (.30)	3.29 (3.17)	.43 (1.72)	2.07*** (.46)	2.35*** (.43)	3.73*** (1.09)
Social status (0 = poor, 4 = upper class)	1.78*** (.14)	1.83*** (.15)	1.60** (.80)	1.11*** (.41)	1.64*** (.17)	1.95*** (.26)	2.11*** (.73)
Satisfaction with dwelling (1 = very unsatisfied, 5 = very satisfied)	.41*** (.09)	.31*** (.09)	1.50** (.74)	1.43*** (.38)	.54*** (.11)	.13 (.17)	−.15 1 (.41)
Crime (crime worse = 1, crime improved = 3)	.79*** (.13)	.92*** (.14)	.40 (.78)	.07 (.38)	.47*** (.16)	1.35*** (.22)	.55 (.597)
Life satisfaction (1 = very unsatisfied, 5 = very satisfied)	2.31*** (.098)	2.20*** (.10)	2.58*** (.64)	3.44*** (.34)	2.23*** (.12)	2.61*** (.18)	1.78*** (.44)

Table 13 continued

	Baseline	Race			Geographical Area		
		African	Coloured	White	Urban formal	Urban informal	Tribal-and-farming
Adjusted R^2	.34	.29	.41	.27	.30	.30	.24

Note The dependent variable in the regression is the newly constructed quality of life index. ***significant at 1 %, **significant at 5 %, *significant at 10 %. Standard errors are indicated in parentheses

Source Authors' own analysis using GCRO data (GCRO 2009)

Table 14 Regression results pertaining to income and sex

	Income categories						Sex	
	0–800	801–1600	1601–3200	3201–12,800	12,801–51,200	51,201–+	Male	Female
Variables								
(Constant)	56.40*** (1.12)	61.82*** (.86)	64.08*** (.93)	66.19*** (1.26)	65.37*** (2.75)	38.47*** (7.44)	58.28*** (.72)	58.44*** (.59)
Marital status	.53 (.60)	.04 (.46)	.03 (.47)	−.99* (.52)	.63 (.80)	−.34 (1.82)	.82** (.40)	.32 (.31)
Sex	.49 (.57)	−.38 (.43)	.33 (.39)	.61 (.37)	1.11** (.55)	.6 9 (1.47)	–	–
White	2.22 (2.90)	.42 (1.60)	−1.37 (1.12)	−.10 (.56)	−.69 (.63)	−.63 (1.76)	2.24*** (.54)	1.99*** (.47)
Coloured	−.35 (2.52)	−3.12** (1.38)	−.28 (1.28)	−.28 (.97)	−1.18 (1.17)	−3.12 (2.53)	−.57 (.89)	.88 (.79)
Age 41–50	−.25 (.17)	−.37*** (.14)	−.39*** (.13)	.01 (.12)	−.28 (.17)	.03 (.41)	−.19 (.11)	−.33*** (.08)
Age 51+	−3.92*** (.76)	−3.66*** (.49)	−1.92*** (.48)	−2.02*** (.52)	−2.45*** (.71)	−4.27** (2.04)	−3.29*** (.43)	−3.45*** (.33)
Urban formal	1.97*** (.58)	1.40*** (.42)	.93** (.40)	1.64*** (.40)	2.69*** (.86)	7.73*** (2.26)	1.96*** (.35)	1.89*** (.28)
TV	2.02*** (.62)	1.83*** (.50)	.80 (.57)	1.32 (.86)	−1.24 (1.88)	15.20 (7.51)	2.43*** (.45)	2.30*** (.39)
Social status	.51 (.47)	.66** (.29)	.81*** (.26)	1.20*** (.27)	.44 (.47)	.55 (1.17)	1.81*** (.23)	1.75*** (.17)
Satisfaction with dwelling	.65*** (.24)	.48*** (.17)	.31* (.17)	−.06 (.17)	.48 (.35)	.20 (.91)	.41*** (.15)	.40*** (.11)
Crime	1.77*** (.35)	.53** (.26)	.65** (.25)	.73*** (.23)	.26 (.37)	2.57*** (.96)	.47** (.20)	1.00*** (.16)
Life satisfaction	1.67*** (.27)	1.84*** (.18)	1.89*** (.18)	1.91*** (.18)	3.17*** (.36)	4.07*** (.86)	2.51*** (.16)	2.19*** (.12)

Table 14 continued

	Income categories						Sex	
	0–800	801–1600	1601–3200	3201–12,800	12,801–51,200	51,201–+	Male	Female
Adjusted R^2	.19	.18	.15	.14	.18	.60	.35	.33

Note The dependent variable in the regression is the newly constructed quality of life index. *** significant at 1 %, ** significant at 5 %, * significant at 10 %. Standard errors are indicated in parentheses

Source Authors' own analysis using GCRO data (GCRO 2009)

Table 15 Regression results pertaining to age

	Age						
	18–20	21–30	31–40	41–50	51–65	66–75	75+
Variables							
(Constant)	59.38*** (1.79)	58.88*** (.85)	58.17*** (.92)	56.60*** (1.10)	53.49*** (1.14)	55.00*** (1.86)	59.34*** (3.73)
Marital status	−1.37* (.78)	.20 (.48)	.19 (.59)	.49 (.61)	1.57** (.59)	1.02 (.85)	3.25** (1.59)
Sex	−.50 (.76)	.41 (.36)	.77* (.41)	1.07* (.51)	.73 (.59)	.20 (.93)	−.12 (1.45)
White	.27 (1.43)	1.99** (.85)	2.33** (.72)	2.09** (.86)	2.49** (.85)	2.90** (1.22)	.26 (1.98)
Coloured	3.70 (2.53)	.59 (1.20)	−.25 (1.13)	1.23 (1.25)	−1.90 (1.57)	.15 (2.79)	−9.35 (7.00)
Age 41–50	–	–	–	–	–	–	–
Age 51+	–	–	–	–	–	–	–
Urban formal	1.57** (.79)	1.88*** (.93)	2.06*** (.31)	1.63** (.54)	2.26*** (.59)	1.33 (.98)	3.95 (1.66)
TV	2.03 (1.29)	3.31*** (.51)	1.06* (.59)	2.84*** (.78)	2.75** (.81)	.37 (1.15)	3.61* (2.12)
Social status	1.84*** (.48)	1.50*** (.25)	2.39*** (.31)	2.13*** (.35)	1.44*** (.38)	.61 (.55)	−.06 (.86)
Satisfaction with dwelling	.41 (.33)	.55*** (.16)	.08 (.18)	.30 (.23)	.52** (.25)	1.27** (.40)	−.23* (.72)
Crime	1.25** (.51)	.74** (.23)	.92*** (.26)	.65** (.31)	.78** (.35)	.95* (.55)	−.77 (1.01)
Life satisfaction	2.61*** (.35)	1.98*** (.17)	2.49*** (.20)	2.29*** (.24)	2.50*** (.26)	2.21*** (.43)	2.24*** (.76)

Table 15 continued

	Age						
	18–20	21–30	31–40	41–50	51–65	66–75	75+
Adjusted R^2	.33	.30	.35	.33	.37	.35	.25

Note The dependent variable in the regression is the newly constructed quality of life index. *** significant at 1 %, ** significant at 5 %, * significant at 10 %. Standard errors are indicated in parentheses

Source Authors' own analysis using GCRO data (GCRO 2009)

References

African National Congress. (1996). *Reconstruction and development programme*. http://www.anc.org.za. Accessed 20 March 2012.

Berenger, V. (2007). Multidimensional measures of well-being: Standard of living and quality of life across countries. *World Development, 35*, 1259–1276.

Bookwalter, J. T., Fuller, B. S., & Dalenberg, D. R. (2006). Do household heads speak for the household? A research note. *Social Indicators Research, 79*(3), 405–419.

Booysen, F. (2002). An overview and evaluation of composite indices of development. *Social Indicators Research, 59*(2), 115–151.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Lawrence Erlbaum Associates, Publishers.

Cummins, R. A. (1996). The domains of life satisfaction: An attempt to order chaos. *Social Indicators Research, 38*(3), 303–328.

Cummins, R. A., Mc Cabe, M. P., Romeo, M. P., & Gullone, E. (1994). The comprehensive quality of life scale instrument development and psychometric evaluation on tertiary staff and students. *Educational and Psychological Measurement, 54*, 372–388.

Dasgupta, P., & Weale, M. (1992). On measuring the quality of life. *World Development, 20*, 119–131.

Development Research Africa. (2009). South Africa—Gauteng City-Region Observatory Quality of Life Survey 2009. https://www.datafirst.uct.ac.za. Accessed 12 Feb 2012.

Diener, E. (1995). A valued based index for measuring national quality of life. *Social Indicators Research, 36*(2), 107–127.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*(2), 276–302.

Dowrick, S., Dunlop, Y., & Quiggin, J. (2003). Social indicators and comparisons of living standards. *Journal of Development Economics, 70*, 501–529.

Easterlin, R. A. (1974). Does economic growth improve the human lot? In P. A. David & M. W. Reder (Eds.), *Nations and households in economic growth: Essays in honor of Moses Abramovitz*. New York: Academic Press Inc.

Economist Intelligence Unit. (2005). *Economist intelligence unit.* http://www.economist.com/topics/economist-intelligence-unit. Accessed 16 Feb 2013.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272–299.

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: Global Print.

Gauteng Planning Commission. (2012). *Gauteng 2055: A discussion document on the long-term development plan for the Gauteng City-Region*. Johannesburg: Gauteng Provincial Government Republic of South Africa.

Gauteng Provincial Government. (2012). *Socio-economic review and outlook*. http://www.gautengonline.gov.za. Accessed 3 Aug 2012.

Gifi, A. (1990). *Nonlinear multivariate analysis*. Chichester, England: Wiley.

Greyling, C. T. (2013). *Measurement and analysis of quality of life of the diverse population of the Gauteng City-Region* (PhD Thesis). Johannesburg. University of Johannesburg.

Groh, A. P., von Liechtenstein, H., & Lieser, K. (2010). The European venture capital and private equity country attractiveness indices. *Journal of Corporate Finance, 16*(2), 20–224.

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, 35*(40), 455–496.

Hansen, T., & Slagsvold, B. (2012). The age and subjective well-being paradox revisited: A multidimensional perspective. *Norsk Epidemiologi, 22*(2), 187–195.

Higgs, T. (2007). Measuring and understanding the well-being of South Africans: Everyday quality of life in South Africa. *Social Indicators Research, 81*(2), 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), 311–336.

Human Sciences Research Council. (2013). *South African Social Attitudes Survey (SASAS).* http://www.hsrc.ac.za/en/departments/sasas. Accessed 10 Oct 2012.

Klasen, S. (2000). Measuring poverty and deprivation in South Africa. *Review of Income and Wealth, 46*(1), 33–58.

Linting, M. (2007). *Nonparametric inference nonlinear principal components analysis: Exploration and beyond.* Doctoral Thesis. University of Leiden. Leiden. https://openaccess.leidenuniv.nl/handle/1887/12386. Accessed Dec 2014

Linting, M., Meulman, J., Groenen, P. J., & Van der Kooij, A. J. (2007). Nonlinear principal components analysis. Introduction and applications. *Psychological methods, 12*(3), 336–358.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*(4), 370–396.

Meulman, J. J., Van der Kooij, A. J., & Heiser, W. J. (2004). Principal components analysis with optimal scaling transformations for ordinal and nominal data. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 49–70). London: Sage.

Møller, V. (2012). South African quality of life trends over three decades, 1980–2010. *Social Indicators Research,*. doi:10.1007/s11205-012-0120-y.

Morris, D. (1979). *Measuring the conditions of the world poor, the physical quality of life index*. New York: Pergamon Press.

National Planning Commission. (2012). *The national planning commission.* http://www.npconline.co.za. Accessed 5 April 2013.

Naudé, W., Rossouw, S., & Krugell, W. (2009). The non-monetary quality of city life in South Africa. *Habitat International, 33*, 310–326.

Nicoletti, G., Scarpetta, S., & Boylaud, O. (2000). *Summary indicators of product and market regulation with an extension to employment protection legislation.* www.oecd.org/eco/eco. Accessed 8 Aug 2013.

OECD. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. Paris: OECD Publishing.

OECD. (2011). OECD Territorial Reviews: The Gauteng City-Region, South Africa 2011. *OECD Publishing,*. doi:10.1787/9789264122840-en.

Osberg, L., & Sharpe, A. (2000). *International comparisons of trends in economic well-being*. New York: Paper presented at the annual meeting of the American Economic Association.

Pallant, J. (2007). *SPSS: survival manual* (3rd ed.). New York: McGrawHill.

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, 104*(2), 195–223.

Ram, R. (1982). Composite indices of physical quality of life, basic needs fulfilment, and income. A Principal Component Representation. *Journal of Development Economics, 11*, 227–247.

Rossouw, S., & Naudé, W. (2008). The non-economic quality of life on a sub-national level in South Africa. *Social Indicators Research, 86*(3), 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, 13*, 439–454.

Rukšėnaitė, J. (2011). Composite indicators as an additional measure for the assessment of the Lithuanian economy. *Lithuanian Journal of Statistics, 49*, 47–51.

Seekings, J. (2001). The uneven development of quantitative social science in South Africa. *Social Dynamics, 27*(1), 1–36.

Sen, A. (1985). Well-Being, agency and freedom: The Dewey lectures. *Journal of Philosophy, 82*(4), 169–221.

South African Audience Research Foundation. (2011). *Development Index.* http://www.saarf.co.za. Accessed 15 Feb 2013.

South African Institute for Race Relationships. (2011). *The South African Development Index.* www.sairr.org.za. Accessed 3 Feb 2013.

Statistics South Africa (Stats SA). 2002. Census 2001 Statistical release—P0301.4/Statistics South Africa. Statistics South Africa. Pretoria.

Stiglitz, J., Sen, A., & Fitoussi, J. P. (2009). *Report by the commission on the measurement of economic performance and social progress.* www.stiglitz-sen-fitoussi.fr/en/index/html. Accessed 12 Nov 2012.

Stroup, M. D., & Stephen, F. (1992). Economic freedom, democracy, and the quality of Life. *World Development, 20*, 119–131.

Szent-Iványi, B., & Vigvári, G. (2012). Spillovers from Foreign Direct Investment in Central and Eastern Europe: An index for measuring a country's potential to benefit from technology spillovers. *Society and Economy, 34*(1), 51–72.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson Education.

United Nations Development Progamme. (2010). *Human Development report*. New York: UNDP.

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.