Westfälische Wilhelms–Universität Münster (WWU)
Institute of Transport Economics*

Smart Cities in Europe: Development of Two Composite Indicators and Econometric Analysis

Markus Overdiek

2019

*The views expressed in this paper do not necessarily reflect those of the Institute of Transport Economics or the WWU.

Contents

List of Figures

List of Tables

List of Abbreviations

AHP Analytic Hierarchy Process

BLUE Best Linear Unbiased Estimator

CI Coverage Index

CR Consistency Ratio

ECC European Capital of Culture

EU European Union

FA Factor Analysis

GDP Gross Domestic Product

GII Global Innovation Index

ICT Information and Communication Technologies

IQR Interquartile Range

KMO Kaiser–Meyer–Olkin Measure of Sampling Adequacy

MAR Missing at Random

MCAR Missing Completely at Random

MI Multiple Imputation

MNAR Missing not at Random

NUTS Nomenclature of Territorial Units for Statistics

OECD Organisation for Economic Co–Operation and Development

OAT One-Factor–at–a–Time

OLS Ordinary Least Squares

PCA Principal Component Analysis

PCC	Pearson Correlation Coefficient
PCFA	Principal Component Factor Analysis
PMM	Predictive Mean Matching
PPS	Purchasing Power Standard
Q–Q	Quantile–Quantile
RCI	Regional Competitive Index
R&D	Research & Development
SA	Sensitivity Analysis
SCCI	Smart City Composite Indicator
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
UA	Uncertainty Analysis
UNESCO	United Nations Educational, Scientific and Cultural Organization
USA	United States of America
VIF	Variance Inflation Factor

1 Introduction

"The 19th century was a century of empires. The 20th century was a century of nation states. The 21st century will be a century of cities."

–

Wellington Web, former Mayor of Denver
(Gassmann, Böhm, & Palmié, 2018, p. 1)

The United Nations (2018) estimates that the percentage of the population living in cities will increase from 55.3 % in 2018 to 68.4 % in 2050 while at the same time, the entire population could increase by another 2.5 billion. Even though that in Europe the population will slightly decrease until 2050, urban areas are also further proliferating with 83.7 % of the population living in cities in 2050.

Those numbers are challenges for the cities as they need to provide their inhabitants with a stable economy, well–functioning mobility and transportation, good governance, sound quality of living, a sustainable environment as well as enabling research & development (R&D), and enhancing people's minds and abilities (Caragliu, Del Bo, & Nijkamp, 2011, p. 65; Lazaroiu & Roscia, 2012, pp. 327–328). Experts from different backgrounds such as large corporations, consultancies, governmental institutions, and researchers try to tackle these complex issues with various approaches and focuses (see for the variety of publications and their origins, for example, Alawadhi et al. (2012), Bosch et al. (2017), IW Consult (2017), J. H. Lee, Hancock, and Hu (2014), and Roland Berger (2017)). Consequently, there is a demand for comprehensive frameworks to assess the cities' current states. Such a framework enables more precise evaluations of their needs as well as the set–up of development plans.

In light of the numerous challenges for cities, the term smart city evolved in a way of understanding which goes beyond a pure focus on technology (Bifulco, Tregua, Amitrano, & D'Auria, 2016, pp. 133–134; Nam & Pardo, 2011, p. 288) and it has become increasingly prevalent in public debate, and academics (Ahvenniemi, Huovila, Pinto–Seppä, & Airaksinen, 2017, p. 235;

Dameri & Cocchia, 2013, p. 4). The popularity of the smart city notion also causes a multitude of definitions (Albino, Berardi, & Dangelico, 2015, p. 4; Grossi & Pianezzi, 2017, p. 79) and sundry criticisms (see Hollands (2008), and Grossi and Pianezzi (2017)), while the field is interdisciplinary and still mostly unexplored (Angelidou, 2014, p. 3). Definitions and borders between disciplines are fuzzy. However, there is some consensus that the smart city concept is inherently holistic (Castelnovo, Misuraca, & Savoldelli, 2016, p. 727; Meijer & Bolívar, 2016, p. 397; Monzon, 2015, p. 19). Ahvenniemi et al. (2017, p. 241) show that some aspects as environmental issues are underrepresented in smart city frameworks. Furthermore, cities often fail to introduce smart strategies in their long–term development plan (Angelidou, 2016, p. 27).

Therefore, the main focus of this thesis aims to introduce a convenient conception of a smart city framework which helps to analyze the state of 65 cities in Europe according to their smartness. This framework persists in the development of an objective and a subjective Smart City Composite Indicator (SCCI) which offer a detailed view on those 65 cities. It is holistic and differentiates at the same time between several aspects to enable close observations of the cities' performances. The approach here is unique in the sense that it distinguishes between objective elements that surround the cities and subjective perceptions of the citizens. Other researchers mix them and rely mostly on the former (see Subsection 2.2). The separate contemplations yield interesting insights because they are benchmarks that allow to examine if inhabitants perceive cities as smart which are smart according to objective criteria. Furthermore, this thesis attempts to identify smart city drivers with the means of an econometric analysis. The idea is to establish the same econometric models for both SCCIs and therefore to investigate if the same variables explain them. Moreover, the econometric models can set a rough agenda for the long–term development of the cities.

The remainder of this thesis is structured as follows. First, a literature review (see Section 2) takes a brief look at composite indicators and discusses their use and implementation generally. After that, seven composite indicators from the literature are examined. Those composite indicators are relevant to the construction of the SCCIs because they consider similar issues and give methodological guidance as they face similar challenges. The literature review closes with a presentation of five contributions that use regression models to have a look at drivers of different urban subject matters. They set the basis for the econometric analysis which is employed to identify the drivers of the cities' smartness.

Section 3 introduces an empirical strategy for the construction of the SCCIs. There are many challenges to a convenient development of the SCCIs. At first, the term smart city needs to be defined. The definition includes four dimensions to portray a detailed view of the city performances according to several aspects. Thereafter, the rationale behind the selection of the 65 European cities is presented, and criteria for appropriate indicators are depicted. The next thing to do is the calculation of the SCCIs, and there are various points to think about in the calculation process. Amongst those are handling of missing data, treatment of outliers, normalization of indicator values, weighting of indicators and dimensions, and aggregation to comprehensive composite indicators. Lastly, uncertainty analysis and sensitivity analysis are done to show that the SCCIs are principally robust to model choices.

The results of the SCCIs are discussed in Section 4. It illustrates the conjunctions within and between the SCCIs, meaning that emphasis is put on aspects which go along within each of the two SCCIs and that the commonalities between both of them are contemplated. Moreover, a descriptive comparison of the SCCIs concentrates on distinctions with respect to population size and between capital and non–capital cities. The descriptive comparison indicates if a city performs better objectively or subjectively according to those aspects.

Section 5 proposes econometric models for the identification of smart city drivers. Thereby, the same models are applied to the objective SCCI and the subjective SCCI, respectively, as dependent variables. After theoretical considerations, the models are specified in line with similar contributions to the literature. Then the models are applied, and the results are presented. The subsection closes with a discussion of the results.

Finally, the thesis ends with a conclusion (see Section 6). It summarizes the main results, points on limitations, and shows avenues for future research.

2 Literature Review

This literature review first takes a closer look at composite indicators in general, mainly concerning their advantages and disadvantages. Thereafter, it contemplates the results of other composite indicators which are in some way or another relevant to the development process of the Smart City Composite Indicators (SCCIs). Lastly, insights from the literature are drawn to build later a regression model which then enables to analyze the SCCIs more closely.

2.1 A General Look on Composite Indicators

Aggregations of underlying performance indicators into a single index are called composite indicators (Jacobs & Goddard, 2007, p. 103). They are useful and widespread tools for policy analysis to compare performances in diverse fields such as environment, economics, society or technology (OECD, 2008, p. 11).[1] Despite their prevalence, composite indicators are controversial (Cherchye, Moesen, Rogge, & van Puyenbroeck, 2007, p. 111). Therefore, this subsection discusses their main advantages and disadvantages.

The idea behind composite indicators is that they are more complex and informative than a number of single indicators (Hagedoorn & Cloodt, 2003, p. 1366). Furthermore, they can assess progress over time and are easier to interpret than many single indicators (OECD, 2008, p. 11), and they can build effective narratives as well as improve public communication (Paruolo, Saisana, & Saltelli, 2013, p. 630; Saltelli, 2007, pp. 72–73).

However, there are several drawbacks which constitute the controversy about composite indicators. Their construction is not straightforward, and a common allegation is that one goal of composite indicators is to gain advocacy for certain narratives due to their subjective nature (Saltelli, 2007, pp. 66–69). Uncertainty is another inherent feature because the knowledge about the underlying structure of the composite indicator is unverifiable and insufficiently precise (Cherchye, Moesen, et al., 2007, p. 135). Furthermore, there is a lack

[1] See Bandura (2008) for a compilation of 178 composite indicators on a country level.

of consensus for theoretical frameworks (Cherchye, Knox Lovell, Moesen, & van Puyenbroeck, 2007, p. 749).

Accordingly, it is necessary to understand and bear in mind potential risks in the construction of composite indicators and in their translation into rankings (Jacobs & Goddard, 2007, p. 109). In the end, the acceptance of a composite indicator is dependent on peer acceptance and negotiation (Saltelli, 2007, p. 70).

2.2 Composite Indicators Relevant to the SCCIs

Most composite indicators focus on comparing country performances in a specific field (OECD, 2008, p. 11). However, composite indicators can similarly be applied to investigate the performances of regions or cities. This subsection considers the main results of seven composite indicators in the literature which use at least a notion similar to holistic smart city approaches and which can give guidance to the construction of the SCCIs due to their methodological remarks. The main aspects considering the methodological approaches of these studies are discussed in Section 3.

A significant contribution comes from Giffinger et al. (2007). They set a focus on medium–sized European cities (between 100,000 to 500,000 inhabitants) since the accumulated population of medium–sized European cities exceeds the accumulated population of large European cities (more than 500,000 inhabitants). This fact underlines their importance. On top of that, challenges for medium–sized cities differ compared to those for large cities. Giffinger et al. (2007) aim to identify perspectives for development, to potentially give strategic advice and to enable benchmarking, lesson–drawing as well as policy transfer via a smart city ranking. However, they also stress possible handicaps of these kinds of rankings which include a threat for long–term development strategies, neglection of complex interrelation in regional development and that poorly ranked cities could simply ignore the results. Their ranking is done with the help of a composite indicator and contains 70 medium–sized European cities. They find that cities from the Benelux, Scandinavia, and Austria are in the top group. Cities in the new EU–member states are at the bottom. However, they differentiate in their study between six characteristics[2] and the results can differ substantially in respect to these characteristics. The characteristics allow obtaining a better idea of the cities' strengths and weaknesses. For example Luxembourg as the first ranked city in total is solely 25$^{\text{th}}$ in the characteristic 'Smart Environment'. Furthermore, Giffinger et al. (2007) divide the six characteristics into 31 factors which are based on 74 mostly objective indicators so that close insights in respect to the performance in those six characteristics can be obtained. For future research, they point on the usefulness of time–series analysis. Therefore, they continue to provide rankings of medium–sized European cities in 2013 and 2014. Furthermore, they also establish a ranking for large European cities in 2015.[3]

Montalto, Jorge Tacao Moura, Langedijk, and Saisana (2018a) do not concentrate on smart cities. What they do is to monitor the performance of 168 cities in 30 European countries for culture and creativity. Similarly to Giffinger et al. (2007), they assign 29 mostly objective indicators to nine dimensions and three major facets.[4] They also provide a well–working online version of their monitor which enables to adapt weights or to simulate the impact of policy actions. With their work, they aim to support policymakers and other actors on a city–level, to emphasize the importance of culture and creativity in respect to resilience and socio–economic perspectives, to benchmark and to inspire new research questions. Montalto et al. (2018a) divide their sample of 168 cities in four groups according to their size. Paris, Copenhagen, Edinburgh, and Eindhoven are at the top of these four groups. Exceptional is that Paris ranks at least second within the three major facets whereas other well–performing cities are not necessarily on top in each major facet. Montalto et al. (2018a) also find that the cultural and creative performance of a city is not determined by population size, but that capitals tend to perform better. When compared to other European cities with at least 50,000 inhabitants, cultural and creative cities have more jobs, younger people, more foreigners, and more human capital. Cities which are culturally and creatively leading are also economically and

[2]Smart Economy, Smart People, Smart Governance, Smart Mobility, Smart Environment, and Smart Living

[3]The rankings are available at http://www.smart-cities.eu.

[4]Cultural Vibrancy, Creative Economy, and Enabling Environment

socially more prosperous which is perhaps mutually reinforced. Lastly, they state that it is critical for the development of low–income cities to be cultural and creative because they otherwise may face a 'low–income trap'.

Another contribution on the city level is from Stanković, Džunić, Džunić, and Marinković (2017). They rank 23 cities from Central and Eastern Europe according to inhabitants subjective perceptions of the cities' smart performance. Twenty–six indicators in five categories[5] rely on the European Urban Audit survey from 2015 (see European Commission (2016b)) to investigate subjectively perceived smart performances. Furthermore, they link the smart performances to the perceived quality of life in those 23 cities. The quality of life is measured by two indicators. In respect to the weights, they determine the category 'Employment and Finance' by far as most important. Stanković et al. (2017) find that there is solely a weak correlation between the ranks of their smart performance model and the ranks according to the perceived quality of life. Only a few cities rank high (e.g., Vilnius) or low (e.g., Miskolc) in both rankings.

Annoni, Dijkstra, and Gargano (2017) contemplate competitiveness of 263 European regions and use 74 mostly objective indicators which are assigned to three dimensions[6] and cover the period from 2012 to 2014. Their Regional Competitive Index (RCI) 2016 follows two previous studies on regional competitiveness from 2010 and 2013 (see Annoni and Dijkstra (2013), and Annoni and Kozovska (2010)). This enables to compare European regional development over time. Furthermore, Annoni et al. (2017) aim to provide a range of regional information which then helps regions to compare themselves with their peers and to plan long–term development. They ascertain that metropolitan and capital regions perform well in many parts of Europe. Solely in Germany, Italy, and the Netherlands are other regions more competitive than the capital region. Moreover, there is a strong positive correlation between the RCI and GDP per capita as well as a moderate positive correlation between the RCI and net migration. Annoni et al. (2017) provide scorecards which point on strengths and weaknesses of a region relative to the 15 regions with the most similar

[5]Infrastructure, Livability and Housing Conditions, Environment, Employment and Finance, and Governance, Urban Safety, and Trust and Social Cohesion

[6]Basic, Efficiency, and Innovation

GDP per capita. A comparison of the three RCI editions over time shows that the regional scores are quite stable. Notable is that in Germany, the RCI scores improved from 2010 to 2013, in France they improved from 2013 to 2016, and in Italy, they deteriorated between 2010 and 2016.

Aiginger and Firgo (2015) introduce a concept of regional competitiveness which takes into account performances of firms and regions as well as 'Beyond GDP' goals.[7] According to 54 mostly objective indicators which are assigned to different outcome and input dimensions,[8] they report results of outcome competitiveness under new perspectives for 229 European regions and try to identify its drivers. Aiginger and Firgo (2015) use data from 2005 as well as from 2011 to show regional developments over time and to identify the new perspectives outcome competitiveness drivers. Their analysis shows that the top regions are in Western and Northern Europe while Southern and Eastern European regions are at the bottom. Despite the neglection of ecological aspects in other studies, their final ranking goes along with similar works such as the RCI 2013. An econometric analysis points on a catching–up process for regions with a low score in 2005. Capabilities and within them 'Innovation & Education' and 'Regional Institutions' are strong new perspectives outcome competitiveness drivers.

Athanasoglou and Dijkstra (2014) measure regional progress in 268 European regions via five indicators, and towards key objectives from Europe 2020.[9] Their regional composite indicator is computed concerning the performance in individual country targets.[10] Generally, more prosperous countries have more ambitious targets than poorer countries. When a common target is used, poorer countries unsurprisingly do worse and richer countries do better.

[7] The 'Beyond GDP' initiative complements GDP with a broader set of indicators which reflect sustainability and well–being in society across social, economic, and environmental dimensions (Stiglitz, Fitoussi, & Durand, 2018, p. 13).

[8] Outcome: Income, Social, and Eco; Input: Cost Competitiveness, Economic Structure, and Capabilities (Innovation & Education, Social System, Ecology, Regional Institutions, and Infrastructure & Amenities)

[9] Europe 2020 is a ten–year growth strategy by the European Commission for smart, sustainable, and inclusive growth (see European Commission (2010)).

[10] Employment, Research and Development, Education, and Fighting Poverty and Social Exclusion

Capital regions frequently outperform non–capital regions within countries and also in comparison to the EU–28 average aggregate score while Brussels is a remarkable exception. There is also a considerable heterogeneity in the scores within a country which emphasizes the importance of disaggregating data from a national to a lower level.

The study by Dutta et al. (2018) is on a country level, but also part of this literature review due to its extensive influence and far–reaching knowledge pool of the persons involved. Dutta et al. (2018) call their composite indicator Global Innovation Index (GII). The GII was done every year since 2008 in a similar fashion. In the GII, Dutta et al. (2018) assess the innovative performance of 126 countries with the help of 80 indicators to promote the importance of innovations. Fifty–seven indicators are objective, 18 are composite indicators from international agencies with a narrow focus, and five indicators are subjective. Solely a few main findings are presented because they are not that relevant for this work here. Those findings include that energy innovations are essential factors for global growth as well as for avoidance of an environmental crisis, and that imbalances in innovation preserve which hamper economic and human development.

In summary, the scope and principal results of seven contributions relevant to the construction of the SCCIs are part of this subsection in the literature review. Those are relevant to the construction of the SCCIs in the sense that they can give guidance for the construction process due to a variety of different approaches and in the sense that they use at least a notion which is similar to smart city concepts, such as competitiveness or innovation. Some studies are quite close to this work in respect to the contemplation of European cities and the smart cities approach (see Giffinger et al. (2007)) while others provide ideas to the construction of the subjective SCCI (see Stanković et al. (2017)) or captivate by their expertise (see Dutta et al. (2018)). Table 1 summarizes key attributes of the seven studies under review in this subsection.

Table 1: Key Attributes of the Studies in Subsection 2.2

Authors	Focus	Level	Observations	Indicators
Giffinger et al. (2007)	Smartness	City	70	74
Montalto et al. (2018a)	Culture & Creativity	City	168	29
Stankovic et al. (2017)	Smartness	City	23	26
Annoni et al. (2017)	Competitiveness	Regional	263	74
Aiginger & Firgo (2015)	Competitiveness	Regional	229	54
Athanasoglou & Dijkstra (2014)	Europe 2020	Regional	268	5
Dutta et al. (2018)	Innovation	Country	126	80

2.3 Regression Models Relevant to the SCCIs

One more goal of this thesis is to investigate which aspects drive the objective and subjective SCCI. This can plausibly be done with the help of regression models. Therefore, the results of five studies are discussed in this subsection. Those studies implement regression models to explain issues on a city level. Necessary for the inclusion of the contributions is that they are relevant in the sense that insights from their models and their use of independent variables can be drawn.

It is evident first to recall the seven contributions in Subsection 2.2 and to see if they also use regression models to explain the outcome of their composite indicators. However, solely Aiginger and Firgo (2015) do so. The basic idea of their regression model is that the elements of their composite indicator at a previous point in time explain the outcome of their composite indicator at a later point in time. Such a set–up is not applicable for the SCCIs because they are measured at a single point in time. On the other hand, some authors in this subsection calculate composite indicators which they then employ as a dependent variable in their regression model. But in all these cases their composite indicators are not helpful to fulfill the needs of the SCCIs, or they solely describe the construction of their composite indicators superficially which makes it inappropriate to include them in Subsection 2.2, and Section 3.

Convenient orientation for the SCCIs regression models offer Neirotti, de Marco, Cagliano, Mangano, and Scorrano (2014). They contemplate 70 cities from all around the world and analyze, if, and how, emerging smart city models differ from the smart city concepts which are developed by stakeholders with expertise such as city planners, technology visionaries, and scientists. For that purpose, they develop a Coverage Index (CI) which represents the number of application domains in which cities have introduced projects. They divide the CI into six categories,[11] and with the help of explanatory factor analysis they assign these six categories to a hard domain and a soft domain.[12] Neirotti et al. (2014) use the CI, the hard domain, the soft domain, and the six categories as dependent variables. Independent variables in their model are regional dummies, GDP per capita, GDP growth, CO_2, internet diffusion, transparency, R&D expenditure, population, and population density. On these dependent and independent variables, they run a multiple linear regression. Overall they find that cities which invest in hard domains are less likely to invest in soft domains, and vice versa. This fact works as evidence that there exists no predominant smart city model, but that there are at least two approaches around. Considering the higher mean in explanatory factor analysis for the hard domain, a further interpretation is that cities' smart strategies primarily focus on technology, and not on people. Furthermore, population density has a significant positive effect on the CI. They conclude that there is still no common smart cities definition

[11] Natural Resources and Energy, Transport and Mobility, Buildings, Living, Government, and Economy and People

[12] Hard Domain: Natural Resources and Energy, Transport and Mobility, and Buildings; Soft Domain: Living, Government, and Economy and People

and that the evolution of a smart city depends mainly on local contextual factors.

Caragliu and Del Bo (2015) have a look at smart specialization strategies and their connection to the smart cities notion in the EU. Regions are smart specializing if they show an above average capability of specialization in industries where they have a positive competitive advantage growth. By this definition, Caragliu and Del Bo (2015) first compute the change in regional smart specialization with respect to sectoral data in value added and labor force. Their results suggest that innovation and science are essential for competitive advantages in the long–run. Second and more relevant to this work, they construct an urban smartness indicator for 309 cities within the EU.[13] They try to find out if the regional smart specialization as an independent variable can explain the smartness indicator. To do so, they employ three models. An Ordinary Least Squares (OLS) regression with robust standard errors when the smartness indicator is a continuous variable and ordered logit as well as probit models when the smartness indicator is displayed as a categorical variable. As control variables, Caragliu and Del Bo (2015) use GDP per capita, regional dummies for urbanization as well as for new members of the EU, the ratio of R&D and GDP, and an indicator for interpersonal trust. Thereafter, they run models with the same model specifications despite that they now assign higher weights to high–tech industries. Altogether they point out that the relationship between regional smart specialization and urban smartness is significant. The basic OLS regression shows that an increase of one unit in the indicator for smart specialization leads to an increase of 0.3 units in the indicator for urban smartness. Counterintuitively, the results of the basic OLS regression also suggest that smart cities are more prevalent in rural areas and new member states of the EU. Caragliu and Del Bo (2015) argue that rural areas are possibly not negatively affected by congestion externalities and that new member states of the EU catch–up. The models which emphasize high–tech industries indicate that regional specialization in innovative sectors does not necessarily drive the emergence of smart cities.

[13] Human Capital, Social Capital, Transport Infrastructure, ICT Infrastructure, Natural Resources, and E–Government

Another contribution comes from Oueslati, Alvanides, and Garrod (2015). They investigate the phenomenon of urban sprawl in 282 European cities and at three points in time (1990, 2000, and 2006). Urban sprawl describes that urban areas take a higher proportion of the land area which is available. Oueslati et al. (2015) analyze developments in respect to urban sprawl with the help of two indices which they use as dependent variables. The indices describe changes in artificial areas and the spatial patterns of residential development. They try to explain these two indices as dependent variables with population and GDP per capita while as control variables they implement the agricultural added value to the area of agricultural land, highway density, the number of rainy days per year, the average temperature of the years' warmest moth, the annual average concentration of NO_2, the altitude of the median city center above sea level, recorded crimes, number of cinema seats and dummies with regard to European sub–regions. Moreover, they choose random and fixed–effects models to take into account the panel characteristic of their data. Their results suggest that increases in income per capita and population growth are linked to the expansion of urban areas, but that the picture for urban fragmentation is less unambiguous.

Węziak–Białowolska (2016) provides an econometric model to investigate the quality of life in 79 European cities. Her model primarily relies on subjective data. The subjective data are from the European Urban Audit survey in 2015 (see European Commission (2016b)). She employs a questionnaire of the European Urban Audit survey in 2015 to measure the satisfaction to live in a specific city as a dependent variable and other questionnaires which capture more precise elements of the city life as well as aspects like unemployment and GDP as explanatory variables. Furthermore, she implements control variables on the citizen and the city level. Those control variables partly have an objective nature. Amongst others, she includes the age group, gender, the population size of the city, and a dummy to distinct Southern European regions from other parts of Europe. Węziak–Białowolska (2016) codes the dependent variable as dichotomous, and therefore she uses a logistic model. Her findings suggest that satisfaction with life in a city varies across European cities and also within them. Furthermore, safety is the essential aspect which contributes to the satisfaction of living in a city.

The article from Senlier, Yildiz, and Aktaş (2009) investigates the quality of life in Kocaeli and compares it with ten European cities that have a similar size.[14] They conduct 300 surveys in Kocaeli with questions generally in line with those from a European Urban Audit survey.[15] Senlier et al. (2009) use factor analysis to assign questions which cover specific issues of the life in Kocaeli to superordinate areas. Those superordinate areas are social and cultural facilities, educational facilities, quality of environment, sufficiency of health services, safety, quality of health services, public transport, neighborhood relations, and overall satisfaction. They constitute the independent variables. The dependent variable is a question concerning the quality of life. Senlier et al. (2009) apply an OLS regression. In accordance with Węziak–Białowolska (2016), Senlier et al. (2009) find that safety is the most crucial aspect for the quality of life in Kocaeli. A descriptive comparison with other European cities which are part of the European Urban Audit survey emphasizes that cities with economic strength also have a high quality of life.

To conclude, the essential features of econometric models and the main results of five contributions relevant to the establishment of a regression model for the SCCIs are part of this subsection in the literature review. The needs of both SCCIs are taken into account because the review includes contributions which rely on objective and on subjective data. This part of the literature review is especially helpful in the sense that it gives guidance for the selection of independent variables. Table 2 summarizes key attributes of the five studies under review in this subsection.

[14] Without clearly stating, Senlier et al. (2009) are probably referring to population size.

[15] The Urban Audit survey which Senlier et al. (2009) refer to is untraceable.

Table 2: Key Attributes of the Studies in Subsection 2.3

Authors	Focus	Level	Observations	Independent Variables
Neirotti et al. (2014)	Smartness	City	70	12
Caragliu & Del Bo (2015)	Smartness	City	309	7
Oueslati et al. (2015)	Urban Sprawl	City	282	15
Weziak-Białowolska (2016)	Quality of Life	City	41,645 Citizens 83 Cities	34
Senlier et al. (2009)	Quality of Life	City	300	8

3 Empirical Strategy for the Smart City Composite Indicators

"All models are wrong; some models are useful."

–

Box, Hunter, and Hunter (2005, p. 440)

The construction of a composite indicator involves several steps and decisions which are based on the personal opinion of the researcher (Saltelli, 2007, pp. 69–70). There are guidelines and handbooks available for the identification of the best–suited steps which point on various aspects to discuss and think about before constructing composite indicators (see European Commission (2016a), Nardo, Saisana, Saltelli, and Tarantola (2005), and OECD (2008)). In accordance with these guidelines, this section provides a nine–step theoretical construction framework for the Smart City Composite Indicators (SCCIs). Application of the theoretical insights and the discussion of interim results are directly part of every step. The theoretical framework for the objective and the subjective SCCI is identical to ensure comparability between both.

Within every step, the approaches from other contributions in the literature are presented, improvements to these are developed and applied on the SCCIs. To do this, three phases are part of every step:

(1) *Literature*: Methodological proposals made by the authors in Subsection 2.2 which are relevant for the respective step in the construction process of the SCCIs.

(2) *Theory*: Theoretical considerations to fit the needs of the SCCIs.[16]

(3) *Application*: Exertion of the developed step in the theoretical framework and discussion of the interim results.

[16]Some theoretical considerations which are made in more detail are not directly necessary for the respective step, but part of the subsequent uncertainty and sensitivity analysis.

These phases and the steps to construct the SCCIs are not wholly self–contained, and sometimes interrelated. Nevertheless, they should give some guidance for better understanding the modeling procedure and modeling choices.

3.1 Definition & Dimensions

Firstly, it is necessary to identify the goal and scope of the SCCIs. The SCCIs aim to measure the smartness of European cities subjectively and objectively. As done by as good as any researcher in the development of composite indicators, discrete and continuous data are implicitly treated equally. Subjective data are those which pertain to subjective feelings quantified through questions, and objective data are those which do not pertain to individual perceptions but comprise tangible aspects (Yuan, Lim, Lan, Yuen, & Low, 1999, p. 8). Therefore, it is possible that objective data are obtained from survey questions if the answers do typically not depend on the individual who answers the question (e.g., indicator B9O measures the average number of rooms per inhabitant).[17]

1. Literature

Giffinger et al. (2007, p. 11) do concentrate on smart cities and define a smart city as "a city well performing in a forward–looking way in [...] six characteristics, built on the 'smart' combination of endowments and activities of self–decisive, independent and aware citizens." The characteristics are: Smart Economy, Smart People, Smart Governance, Smart Mobility, Smart Environment, and Smart Living. Furthermore, Stanković et al. (2017, p. 524) stress five categories (Infrastructure, Livability, Environment, Employment and Finance, and Governance) to capture smart city performances.

Montalto et al. (2018a, p. 44) contemplate cultural and creative cities and expect them "to promote a model of harmonious urban development and wellbeing which is sustainable for both present and future generations." In their opinion, the delimitation to a smart city is that a smart city would put digital and communication technologies at the center while technologies are a

[17] The answer to those questions can also depend on the individual due to imperfect information or intentionally false statements. Those issues are not taken into account here.

complementary for a cultural and creative city (Montalto et al., 2018a, p. 44). Moreover, they try to identify a city as being cultural and creative with the help of three sub–indices (Cultural Vibrancy, Creative Economy, and Enabling Environment) and nine dimensions (Montalto, Jorge Tacao Moura, Langedijk, & Saisana, 2018b, p. 1).

Other works considered here have a slightly different focus. They emphasize regional competitiveness in Europe (Aiginger & Firgo, 2015; Annoni et al., 2017; Athanasoglou & Dijkstra, 2014) as well as global country competitiveness (Dutta et al., 2018). For example, Annoni et al. (2017, p. 2) define regional competitiveness as "the ability of a region to offer an attractive and sustainable environment for firms and residents to live and work." Even though those notions on competitiveness are often closely related to smart city definitions, they have different focal points and thus, are not contemplated in more detail within this subsection.

2. Theory

There is no universal definition for the term smart city (Neirotti et al., 2014, p. 35; Stanković et al., 2017, p. 521). Various disciplines and scholars have controversial debates about the term in general and the aspects that should be considered when talking about smart cities. However, for the research purpose here it is most convenient to follow a strand of literature which understands the smart city notion as an inherently holistic concept (Ahvenniemi et al., 2017, p. 236). A holistic understanding requires the identification of different dimensions in which a city needs to perform well so that it is labeled smart. The International Telecommunication Union (2014) provide a detailed analysis of 116 smart city definitions.[18] The definitions come from leading stakeholders in this area such as researchers, corporations or governmental institutions. They identify 50 keywords in these 116 definitions and group them logically into key dimensions.[19] (International Telecommunication Union, 2014, pp. 7–12) The dimensions are:

[18]More precisely of smart sustainable city definitions.
[19]They use the term categories instead of dimensions.

1. ICT/Communication/Intelligence/Information
2. Infrastructure/Services
3. Environment/Sustainability
4. People/Citizens/Society
5. Quality of Life/Lifestyle
6. Governance/Management/Administration
7. Economy/Resources
8. Mobility

A holistic definition of the term smart city ideally takes into account these eight dimensions because they represent the numerous parts of a smart city which many experts perceive as important.

3. Application

The dimensions according to the International Telecommunication Union (2014) are merged here due to data availability. Indicators of the subjective and the objective SCCI are obliged to provide revealing results for every dimension. This requirement is solely fulfilled by a substantial number of high–quality indicators for each dimension and to obtain those is easier to manage with a smaller number of dimensions. The merged dimensions are:[20]

A: Infrastructure & Mobility

B: Living & Social Cohesion

C: Economy & Governance

D: Environment & Sustainability

Dimension A emphasizes primarily the capability of a city to move people and to provide physical structures. Dimension B highlights aspects which are

[20] The dimension '1: ICT/Communication/Intelligence/Information' is not directly included in any of the merged dimensions. Nevertheless, the dimension is indirectly part of every merged dimension as an enabler.

related to the overall living conditions in a particular city and accounts for aspects as health, education, openness, and equality. Dimension C measures the success in economic competition and the quality as well as the efficiency of public services. Dimension D points on the soundness of environmental conditions within a city.

Consequently, a ’Smart City’ is defined here as a city which performs well in the four dimensions ’Infrastructure & Mobility’, ’Living & Social Cohesion’, ’Economy & Governance’, and ’Environment & Sustainability’. The way the term smart city is defined is consistent with similar approaches in the literature (see Giffinger et al. (2007), Montalto et al. (2018a), and Stanković et al. (2017)).

3.2 Sample

As a next step, the cities contemplated in the SCCIs are selected. This subsection is closely related to Subsection 3.3. Some aspects could be discussed in both subsections. Nevertheless, they are split for understandability, but their interrelatedness should be kept in mind.

1. Literature

Giffinger et al. (2007, p. 13) select 70 medium–sized European cities. They start with 244 functional urban areas covered by European Urban Audit. They exclude 150 cities following three criteria. The first criterion is that they require a population size of 100,000 to 500,000 inhabitants. Second, they ask for at least one university in the city. Third, they want the catchment area of the cities to be less than 1,500,000 inhabitants. Furthermore, Giffinger et al. (2007, p. 13) incorporate the fact that data availability for some cities is low and reduce their sample size further but do not report any precise requirement.

Stanković et al. (2017) conduct their analysis on 23 Eastern European cities. Their criteria are twofold. A city needs to be part of the European Urban Audit survey from 2015 (see European Commission (2016b)) and it has to be from Central or Eastern Europe (Stanković et al., 2017, pp. 529–530).

Montalto et al. (2018a, p. 54) select 168 European cities from about 1,000 cities in Eurostat's Urban Audit database according to three criteria: 1. Cities which have been or will be European Capital of Culture (ECC) up to 2019, or which have been shortlisted to become an ECC up to 2021; 2. UNESCO Creative Cities; 3. Cities are hosting at least two regular international cultural festivals. They further exclude 13 of the 168 cities because they demand a minimum of 45 % data coverage at the index level and 33 % for the 'Cultural Vibrancy' and 'Creative Economy' sub–indices, or because cities are located outside the EU (Montalto et al., 2018b, p. 2).

The samples from other works in consideration here employ regions or countries, and therefore they are not that helpful for the sample selection. However, worth of mention is that Dutta et al. (2018, p. 370) require the economies to cover a minimum of 66 % of data points in both of their sub–indices, 'Input' and 'Output'.

2. Theory

Appropriate databases need to be identified for the sample selection of the SCCIs. For the subjective SCCI, data can be drawn from a survey which was conducted in the context of the European Urban Audit in 2015 (European Commission, 2016b, p. 2). The European Urban Audit survey 2015 was done on behalf of the Directorate–General for Regional and Urban Policy in 79 European cities to get insights about the opinions of cities' inhabitants on various urban issues.[21] During the fieldwork 40,798 citizens with different demographic and social characteristics were interviewed via telephone in the 28 States of the European Union, Iceland, Norway, Switzerland, and Turkey between 21st of May and 9th of June 2015. The interviews took place in the respective mother tongues.[22] Previous European Urban Audit surveys are from

[21] Strictly speaking, the European Urban Audit survey 2015 was done in 79 European cities and four Greater cities. The cities which are additionally considered as Greater cities are Athens, Lisbon, Manchester, and Paris. However, the Greater cities are directly excluded from the discussion here.

[22] A challenge of cross–national research is that language biases (see Harzing et al. (2009)) and cultural biases (see van de Vijver and Tanzer (2004)) are an issue. But there is no convenient study which provides information for those biases on all or at least almost all of the diverse regions in the sample.

2006, 2009, and 2012 (European Commission, 2016b, p. 8).[23]

The European Urban Audit survey is the most compelling database available for the subjective SCCI because there exists no other survey which covers a similar number of cities, questions, and respondents. Using data from other surveys to increase the sample size is inappropriate because it would lead to massive inconsistencies. Other surveys implement different methodologies, contemplate entirely dissimilar samples of cities, and do not provide the quality as the European Urban Audit survey. Additionally, framing effects (see Kahneman (2011)) could be further amplified when data from other surveys are implemented.

Contrary to the subjective SCCI, there are many different datasets available for the objective SCCI. Those datasets are mainly from the OECD, Eurostat or other EU institutions.[24] But not every dataset is complete. Therefore, minimum data coverage needs to be set. The determination of an appropriate threshold for the cities' data coverage is a trade–off between the quality as well as the accuracy of the indicators and the number of indicators that are part of the objective SCCI and thus, to some extent arbitrary. For the objective SCCI, a stricter threshold than in Montalto et al. (2018a) and Dutta et al. (2018) is plausible because their threshold is quite low. The requirement for the objective SCCI is a minimum of 75 % indicator values for each city within the sample in every dimension and 85 % indicator values for each city across all dimensions.

3. Application

The first and at the same time the primary requirement for the SCCIs samples is that a city needs to be part of the European Urban Audit survey. This requirement limits the sample size to 79 cities.

[23] Unfortunately, the European Urban Audit survey provides no information about the criteria whereby the cities were selected. However, the capital city of every country is part of the survey. Generally speaking, more cities of a country are in the sample when the country has more citizens, and within a country, regional diversity of the cities in the sample can be observed.

[24] Extensive information about the data such as sources, descriptions as well as further remarks are available online: https://drive.google.com/drive/folders/1cyG9ZpZm-3BakRKBRrDlnxOvJVgQuj8t?usp=sharing.

The second requirement is that each city needs to provide a minimum of 75 % indicator values in every dimension and 85 % indicator values overall. This threshold leads to the exclusion of 14 from the 79 cities so that a sample of 65 cities remains (see Table A1).[25] Those 65 cities include the capitals from all 28 countries of the EU.

3.3 Indicators

Now the indicators of the SCCIs are discussed more closely. As also stated above, this subsection is strongly interrelated to Subsection 3.2.

1. Literature

Montalto et al. (2018b, p. 1) select their 29 indicators in respect to the following five criteria: Coverage, Relevance, Accessibility, Quality, and Timeliness. Dutta et al. (2018, p. 76) state that expert opinion and statistical analysis are behind the indicator selection process of their 80 indicators. They also provide a detailed description of every single indicator (Dutta et al., 2018, pp. 352–365).

Giffinger et al. (2007, pp. 22–23) include 74 indicators from different regional levels in their analysis. They do not state a minimum data availability threshold which any indicator has to satisfy, but they have a coverage rate of 87 % for their 74 indicators (Giffinger et al., 2007, p. 14). Annoni et al. (2017, p. 19) set a maximum of missing values for each indicator around 10 to 15 %. Stanković et al. (2017) select 26 indicators for their smartness ranking and two indicators to measure life satisfaction. All of their indicators are drawn from the European Urban Audit survey from 2015 (see European Commission (2016b)).

2. Theory

The reasoning which researchers give for selecting their indicators in the construction process of their composite indicators is often sparse even though there

[25] Ankara, Antalya, Braga, Diyarbakir, Dortmund, Essen, Geneva, Irakleio, Istanbul, Oslo, Oulu, Piatra Neamt, Reykjavik, and Zurich were excluded after the list of indicators was finalized.

are criteria available. The OECD (2008, pp. 46–48) points on six criteria for selecting indicators:

1. Relevance
2. Accuracy
3. Timeliness
4. Accessibility
5. Interpretability
6. Coherence

Relevance as the first criterion for selecting the indicator says that an indicator has to be relevant in the context of the composite indicator. Accuracy alludes to the importance of credible sources and trustworthy data which are not distorted by interests or unprofessional collecting practices. Timeliness says that consideration of the indicators' vintages is necessary. Accessibility focusses on the replicability and costs for updating the indicator. Interpretability implies that every indicator is interpretable and has meaning for itself. Coherence means that the indicators are collected with the same concepts, definitions, classifications and methodologies over time and across countries.

The indicators of the objective and the subjective SCCI should also correspond to each other to ensure comparability between both, and it needs to be possible that they can be assigned meaningfully to the four dimensions 'Infrastructure & Mobility', 'Living & Social Cohesion', 'Economy & Governance', and 'Environment & Sustainability'. In addition, the SCCIs set a maximum of missing values within each indicator. If more values are missing, the indicator is excluded. The maximum of missing values for each indicator is set to a quite modest level of 30 % for the SCCIs because it enables to include some interesting indicators. Furthermore, another requirement is data availability for each city in the sample and that requirement is quite strict (see Subsection 3.2).

Another issue in selecting appropriate indicators is that the availability of those indicators on a city level is limited. Therefore, indicators are from a regional or national level when it can be assumed that they still have meaning on a city level. This is a standard procedure in the literature (see Giffinger et al. (2007)). Moreover, sometimes it is difficult to define which data can be assigned to the city level. There are different typologies for European cities and metropolitan regions (see Eurostat (2018)). For simplicity, metropolitan and NUTS 3 regions as mentioned by Eurostat (2018) are assumed to incorporate the city level. NUTS 2 regions depict the regional level, and NUTS 0 regions constitute the country level.

3. Application

Tables A2 and A3 display the indicators of the SCCIs.[26] They satisfy all six requirements as outlined above:

1. Relevance means here that the indicators have to fit the smart cities' definition. Since a smart city is defined herein with respect to its performance in the four dimensions 'Infrastructure & Mobility', 'Living & Social Cohesion', 'Economy & Governance', and 'Environment & Sustainability', indicators are meaningful for these dimensions.

2. The indicators are accurate. Indicators for the subjective SCCI are from the European Urban Audit survey 2015, and the indicators for the objective SCCI are mostly from the OECD, Eurostat or other EU institutions, and thus, they are sufficiently accurate.

3. Timeliness is also satisfied. The fieldwork for the subjective SCCI was done in 2015. Hence, the indicators for the objective SCCI are optimally also from 2015. Due to limitations in data availability, some indicators are from other years.[27] But they are as close as possible to the year 2015 and no more distant than four years.

4. The replication of the SCCIs is easily possible since every indicator is downloadable free of charge. Furthermore, an online database is part

[26]Recall that more information about the indicators such as sources, descriptions as well as further remarks are available online: https://drive.google.com/drive/folders/1cyG9ZpZm-3BakRKBRrDlnxOvJVgQuj8t?usp=sharing.

[27]41 of the 73 indicators for the objective SCCI are from 2015.

of this work which contains more information about the sources for the indicators as well as more detailed descriptions which are important for the accessibility.

5. Every indicator is interpretable and has meaning for itself and is, therefore, interpretable.

6. The coherence over time is not an issue as the SCCIs do not include time series. Furthermore, the indicators for the subjective SCCI are from the same source and use the same methodology across countries. The indicators for the objective SCCI are mostly from governmental bodies such as the EU or the OECD.

The maximum missing values for each indicator is 30 % for the SCCIs. There are no missing values for the subjective SCCI. The coverage rate with this requirement for the objective SCCI is around 90 % when taking into account the cities in the sample and thus, better than in Giffinger et al. (2007).

For the subjective SCCI, all 19 indicators are on a city level. For the objective SCCI, there are 73 indicators.[28] 30 indicators are on a city level, 35 indicators are on a regional level, and eight indicators are on a national level.

3.4 Missing Data

Missing data are an essential aspect in the construction of composite indicators because their treatment can have a severe impact on the final results. It is impossible to simply neglect missing values. All indicator values (missing or not) are by definition always part of composite indicators. Despite the importance of the problem, many researchers use inadequate methods (Baraldi & Enders, 2010, pp. 5–6). Therefore, a detailed discussion of proper missing data treatment is expedient. Note that there are no missing values in the data for the subjective SCCI and thus, this subjection is solely relevant for the objective SCCI.

[28] A few of these 73 indicators are also composite indicators.

Giffinger et al. (2007) insert values from previous years and values from other regional levels. Besides, they use z–scores to normalize their indicators and divide them by the number of values added to obtain imputations (Giffinger et al., 2007, p. 14).[29] Similarly to this is the approach by Annoni et al. (2017). Annoni et al. (2017, p. 15) state that they do not take into account the values of one of their indicators for some regions due to regional specificities. This statement underlines that many researchers think it could be possible to ignore missing values. However, what Giffinger et al. (2007) and Annoni et al. (2017) do is that they assign a weight of zero to an indicator value which is missing in a specific region whereas a different weight is assigned to a non–missing value of the same indicator in another region. Moreover, if missing indicator values of a particular region are assigned a zero weight, it subsequently implies that the weights for non–missing values of this region increase.[30] The consequences are sometimes odd. Annoni et al. (2017) provide various data sheets online which entails the feasibility to reproduce their work. Their pillar 'Higher Education and Lifelong Learning' consists of three indicators. For the French region Guyane, there is solely the indicator 'Early School Leavers' available. The z–score for this indicator in Guyane is the worst z–score among all indicators and regions. Nevertheless, it is used to compute the entire pillar 'Higher Education and Lifelong Learning' which is then used to generate ranks after conducting a min–max normalization. Accordingly, this contributes to the fact that Guyane is on the last rank in the corresponding dimension 'Efficiency', and also in total.

Aiginger and Firgo (2015) do not provide verifiable information due to the imputation of missing values. All they indicate about this issue is one sentence in a footnote in their appendix. In the footnote, Aiginger and Firgo (2015, p. 42) say that they employed a few econometric imputations for some variables.

[29] It does not become entirely clear, but it is plausible that they do this within each of their six dimensions.

[30] An alternative interpretation which would lead to the same results is that mean values are computed for each region (not for each indicator) in each dimension, and that those mean values are then imputed for the missing values. This would imply that the weights for indicator values from the same indicator are always the same among regions.

Athanasoglou and Dijkstra (2014) provide more information. In their index, they calculate regional target values. If the target value is not viable, they draw on national target values from a previous year and compare them to target values from other countries in that year to estimate the corresponding target values for their index (Athanasoglou & Dijkstra, 2014, pp. 22–23). Dutta et al. (2018, p. 370) impute values from other years with a cut–off year in 2007. Besides, they do not impute any values which in fact means that they impute the respective sub–pillar score (Dutta et al., 2018, p. 76). Montalto et al. (2018a) use the most sophisticated imputation methods of all contributions in contemplation. Montalto et al. (2018b, pp. 2–3) impute missing values first by taking the national average if available. Second, they group the cities according to the triplet population–GDP–employment. Then they estimate and impute values based on the values of peer cities. Third, they impute the remaining missing values by using the average of the three nearest neighbors.

2. Theory

Despite the OECD (2008, pp. 55–62) discussing the issue of missing data in its handbook on composite indicators in detail, other contributions in the literature do not contemplate it sufficiently. They directly describe the methods which they use to deal with missing values without investigating the structure of the data. Sometimes this leads to poorly reasoned procedures. Thus, a more sophisticated framework to handle missing data is expedient.

Imputations are in a first step obtained by plausible approximations. These approximations use available data of the year which is closest to the year 2015 or of the regional level which is closest to the city level (Aiginger & Firgo, 2015, p. 42; Athanasoglou & Dijkstra, 2014, p. 22). The method draws its convenience from the fact that it relies on existing data which intuitively share a great deal of correspondence with the true data points for the relevant cities in each indicator.

Since this evident missing data imputation technique is not able to account for all missing values, proceeding considerations are necessary. The most common method to deal with missing data is to solely consider the cases with complete information (Pigott, 2001, p. 354). This procedure is not applicable here because it would reduce the number of indicators and particularly the sample size in an extent which is unjustifiable.[31] Nevertheless, minimum requirements for data availability are set (see Subsections 3.2 and 3.3).

Furthermore, various approaches are applicable to impute indicator values based on mathematical properties. For the selection of an appropriate imputation method, there exists the distinction between data which are missing completely at random (MCAR), missing at random (MAR), missing not at random (MNAR) (Baraldi & Enders, 2010, p. 6) based on the work of Rubin (1976). Data are MCAR if they are absent in an utterly unsystematic way which does not relate to other values of the indicator or to values of other indicators under study.[32] Data are MAR if they are absent in a way which does not relate to other values of the indicator, but which does relate to values of other indicators.[33] Data are MNAR if their missingness is systematic and depends on the indicator values themselves (Baraldi & Enders, 2010, pp. 7–8).[34]

The MCAR mechanism can be examined with the help of a test proposed by Little (1988). This likelihood ratio test statistic assumes multivariate normality but also works for non–normal data (Li, 2013, p. 797; Little, 1988, p. 1201). Let $\boldsymbol{x}_j$ be the matrix of the indicator values (j=1,2,...,n) with p dimensions, J the indicators with missing value patterns, $\hat{\boldsymbol{\mu}}$ the maximum likelihood population mean vector and $\tilde{\boldsymbol{\Sigma}}$ maximum likelihood covariance matrix (Li, 2013, p. 797). The formula for the application of the test by Little (1988) is then given by:

$$d^2 = \sum_{j=1}^{J} n_j(\overline{\boldsymbol{x}}_j - \hat{\boldsymbol{\mu}}_j)^\intercal \tilde{\boldsymbol{\Sigma}}_j^{-1}(\overline{\boldsymbol{x}}_j - \hat{\boldsymbol{\mu}}_j) \sim \chi^2 \tag{1}$$

[31] There are some cities which solely contain one missing value among the 75 objective indicators. However, solely Vienna lacks no data at all.

[32] If the tendency of individuals to report their income is purely by chance (Sinharay, Stern, & Russell, 2001, p. 318).

[33] If the tendency of individuals to report their income is related to other indicators such as education (Sinharay et al., 2001, p. 318).

[34] If individuals with a high or a low income tend not to report it (Sinharay et al., 2001, p. 318).

If H_0 is not rejected for a chosen significance level and with $\sum_{j=1}^{J} p_j - p$ degrees of freedom (where p_j are the number of observed components in pattern j), MCAR is assumed:

$$H_0 : d^2 \sim \chi^2$$

After investigating the structure of the data closely, the mean imputation as the most commonly employed imputation method as well as a specified multiple imputation (MI) method as a sophisticated imputation approach are discussed.

The mean imputation is a simple and commonly used approach for single imputations (Song & Shepperd, 2007, pp. 269–270). It replaces missing values of an indicator with the mean of the observed values of this indicator. This procedure preserves the mean of the indicator, but often the variance and covariance get severely biased (Huisman, 2009, p. 4). Another property is that the imputed values have a correlation of zero with values of other indicators (Baraldi & Enders, 2010, p. 12). Mean imputation also assumes that the missing values are MCAR which is in reality rarely the case. It solely works fine when a few values are missing (Saunders et al., 2006, p. 22). Several authors emphasize that it "is the worst missing data handling method available [and] in no situation [...] defensible" (Enders, 2010, p. 43). Hence, the imputation of missing values for the objective SCCI has to rely on a more ambitious method.

Amongst imputation methods, maximum likelihood and MI represent the "state of the art" (Schafer & Graham, 2002, p. 173). The choice for a convenient specification of one of these state of the art methods depends on the fulfillment of multivariate normality. Multivariate normality is examinable with the help of an omnibus test for multivariate normality proposed by Doornik and Hansen (1994) as it generally shows good performance (see Farrell, Salibian–Barrera, and Naczk (2007)). Let $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ be approximate normal variates which are transformed from the sample skewness and kurtosis (C. Lee, Park, & Jeong, 2016, p. 1403). Then the test statistic is defined as:

$$DH = \boldsymbol{Z}_1^{'}\boldsymbol{Z}_1 + \boldsymbol{Z}_2^{'}\boldsymbol{Z}_2 \qquad (2)$$

If H_0 is not rejected for a chosen significance level and with $2k$ degrees of freedom, multivariate normality is assumed:

$$H_0 : DH \sim \chi^2$$

Without multivariate normality at hand as the results of the test statistics will show, predictive mean matching (PMM) as an option for MIs is appropriate for the estimation of imputations. MIs are sets of multiple plausible values for the missing values (Rubin, 1996, p. 476). PMM "imputes missing values by means of the nearest–neighbor donor with distance based on the expected values of the missing variables conditional on the observed covariates" (Vink, Frank, Pannekoek, & van Buuren, 2014, p. 62). It yields several advantages for the purpose here. First, PMM performs well for data which are multivariate normal and for data which are not multivariate normal (Vink et al., 2014, pp. 80–84; Wulff & Ejlskov, 2017, p. 47). Second, PMM preserves the distributional pattern of the indicator (Vink et al., 2014, p. 78; Wulff & Ejlskov, 2017, p. 47). Third, PMM is quite robust against model misspecifications (Morris, White, & Royston, 2014, p. 4). Fourth, no distributional assumptions are necessary (Kleinke, 2017, p. 372). Fifth, PMM leads to plausible imputations because they can not fall outside the range of the observed values (Vink et al., 2014, p. 78).[35]

After the data sets are compiled by using PMM, point estimates are inserted for the missing values according to Rubin's rule of combination (Carlin, Li, Greenwood, & Coffey, 2003, p. 3). Let $\hat{U}_i$ represent the imputed value for the i^{th} dataset and m the number of complete dataset estimates. Then the point estimate imputation for each missing value is:

$$\overline{U} = \frac{1}{m} \sum_{i=1}^{m} \hat{U}_i \tag{3}$$

[35] Multivariate Imputation by Chained Equations is also often used as a practical tool for multivariate imputations. It also offers the possibility to assume different distributions for each indicator. However, some attempts in this direction led to inconvenient and illogical results for a small number of values.

Before proceeding, all indicators in which a lower value implies a better performance are multiplied by −1 to warrant a consistent statistical analysis. Otherwise, test results and estimation of imputations are distorted due to a misleading compilation of indicator properties.

The investigation of the data structure yields ambiguous results. For dimensions A and B, the MCAR assumption is rejected at a standard significance level of 0.05. For dimensions C and D, the MCAR assumption is not rejected at a significance level of 0.05 (see Table A4). Since the MCAR assumption is rejected for two dimensions of the objective SCCI at a significance level of 0.05, the mean imputation is not an appropriate method.

The results of the multivariate normality test by Doornik and Hansen (1994) show that multivariate normality does not prevail at a standard significance level of 0.05 in neither dimension. These results cancel out a great deal of methods since most dimensions are not assumed to be multivariate normal. But as outlined within the theoretical considerations, PMM as a way to deal with MIs is a reasonable alternative for the objective SCCI.

The application of MI is ambiguous about the ideal number of iterations. However, the imputations for the objective SCCI follow a rule of thumb which proposes to create 20 data sets (Baraldi & Enders, 2010, p. 15). In every dimension all of the indicators with complete data are implemented as auxiliary variables for practical reasons and because researchers recommend to use a large set of auxiliaries in the context of MIs (Enders, 2010, p. 133).

3.5 Outliers

Outliers are a controversial subject among scholars (e.g., 'The Black Swan: The Impact of the Highly' (Taleb, 2007)). There is no consensus in respect to the *if* and the *how* of outlier treatment. However, the treatment of outliers is plausible for the SCCIs because some values would influence the results in a profuse way which is most likely inconvenient. Furthermore, sometimes it is possible that extreme values occur due to problems within the collection process. This is why the treatment of outliers is discussed closer. An outlier is defined here as “a data object that deviates significantly from the rest of the objects” (Han, Kamber, & Pei, 2012, p. 544).

1. Literature

Montalto et al. (2018b, p. 2) and Dutta et al. (2018, p. 371) follow similar approaches. They identify the existence of outliers by assessing the values for the skewness and kurtosis. If values for the skewness and/or the kurtosis are quite high, they treat outliers because a significant deviation from the normal distribution is at hand.[36] Skewness is informative about symmetry (Krishnamoorthy, 2006, p. 12) while kurtosis is informative about the tails (and not as often misspecified about the peak) (Westfall, 2014, pp. 191–192). But assessing the normality assumption with skewness and kurtosis in small samples can be inadequate (Hain, 2010, pp. 46–57; A. R. Henderson, 2006, pp. 115–119; Kim, 2013, pp. 52–53; McNeese, 2016; Razali & Wah, 2011, p. 32) and is also quite arbitrary as recommendations for rule–of–thumbs can vary substantially.[37] After identifying the indicators with outliers, both contributions apply winsorization to those indicators. Montalto et al. (2018a) use the interquartile range (IQR) and assign the next highest values to the outliers within each indicator. Dutta et al. (2018) employ a natural log formula within a range given by the minimal and maximal value for the indicator.[38]

Annoni et al. (2017, p. 19) apply a Box–Cox transformation to adjust for outliers when necessary.[39] The Box–Cox transformation is a power transformation while the specification of the transformation depends on a chosen parameter (Osborne,

[36]Montalto et al. (2018a) require the skewness of an indicator to be greater than two and/or the kurtosis to be greater than 3.5. Dutta et al. (2018) use values of 2.25 for the skewness and 3.5 for the kurtosis as boundaries.

[37]Most handbooks recommend ± 2 as boundaries for skewness and kurtosis. More conservative suggestions are ± 1, and more liberal suggestions are ± 3.

[38]Interestingly and despite the similarities in their approaches, Montalto et al. (2018a) treat outliers before imputing missing values, and Dutta et al. (2018) do it vice versa. Theoretical contributions on this issue are rare. However, it seems adequate to use the original data for imputations and as a next step to treat the outliers.

[39]They do not state what precisely *necessary* means. However, it is likely that they also assess if the data of an indicator can be assumed to be normally distributed as the Box–Cox transformation accounts for this.

2010, p. 4). Annoni et al. (2017, p. 24) set those parameters either to 0.3, 0.5 or 0.8. Despite numerous advantages of the Box-Cox transformation such as meeting the normality assumption (Osborne, 2010, p. 6), there is also a drawback because power transformations affect the data normalization process (OECD, 2008, p. 84). Athanasoglou and Dijkstra (2014, pp. 19–20) do not discuss the treatment of outliers in much detail. However, outliers implicitly do not play an important role in their framework since there are target values for each observation which means that regions are assessed against themselves. This generally diminishes the occurrence of extremely good or bad performances. Aiginger and Firgo (2015), Giffinger et al. (2007), and Stanković et al. (2017) do not state that they treat outliers.

2. Theory

There are various methods to assess the prevalence of outliers and each edit of these outliers has to be done with much caution. As described above, some authors use skewness and kurtosis to detect the indicators with suspicious values. Their idea to first evaluate normality is also applied for the SCCIs because it warrants that not more indicators and values are treated than necessary. But due to the mentioned weaknesses, a more sophisticated approach to find these indicators is employed. The approach here follows a recommendation which suggests using a normality test in conjunction with a visual assessment (Ghasemi & Zahediasl, 2012, p. 489) because solely relying on formal tests can be insufficient (Razali & Wah, 2011, p. 32).

The sample for the SCCIs is between small and medium size and they each contain the same 65 European cities. For a sample size of this magnitude, the power of the Shapiro–Wilk test to examine the normality assumption is appropriate (Adefisoye, Golam Kibria B.M., & George, 2016, p. 7; Ahad, Yin, Othman, & Yaacob, 2011, p. 641; Ghasemi & Zahediasl, 2012, p. 489; Razali & Wah, 2011, p. 32).

The normality test by Shapiro and Wilk (1965) was originally developed for sample sizes which are no larger than 50. Meanwhile, it is possible to conduct the test for up to 2,000 or even 5,000 observations (Razali & Wah, 2011, p. 25). It is an omnibus test which enables to detect departures from normality either due to skewness, kurtosis, or both (Althouse, Ware, & Ferron, 1998, p. 2). Let $x_{(i)}$ be the i^{th} order statistic, x_i the indicator values, and $\overline{x}$ the mean of these indicator values. Furthermore, a_i contains the expected values of the order statistics of independent and identically distributed random variables sampled from the standard normal distribution as well as the covariance matrix of those order statistics (Razali & Wah, 2011, p. 25).

Then the test statistic is given by:

$$W = \frac{(\sum_{i=1}^{n} a_i x_{(i)})^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2} \tag{4}$$

If H_0 is not rejected for a chosen significance level, normality is assumed:

$$H_0 : X \sim N(\mu, \sigma^2)$$

As a next step, a quantile–quantile (Q–Q) plot as a visualization method is conducted to show whether there are outliers even though the Shapiro-Wilk test would indicate otherwise. A Q–Q plot draws the quantiles of the actual values against those of the expected values for a normal distribution (Wang & Bushman, 1998, pp. 49–40). For the sake of completeness, values for skewness and kurtosis as in Dutta et al. (2018) and Montalto et al. (2018a) are also calculated and compared to the results of the Shapiro–Wilk test.

Now, as the indicators which possibly contain outliers are determined, these outliers need to be detected. There are a lot of options available, and many researchers rely on the standard deviation (e.g., outliers are those values which are more than ±2 standard deviations distant from the mean) (Leys, Ley, Klein, Bernard, & Licata, 2013, p. 765). This is problematic because the outlying values have a substantial influence on the standard deviation (Leys et al., 2013, p. 764), especially in small samples (Cousineau & Chartier, 2010, p. 60). Hence, a robust measure is preferred to detect outliers. As a robust measure, the IQR is selected as it is also the case in the work by Montalto et al. (2018a). IQR has a breakdown point of 25 % (Rousseeuw & Croux, 1993, p. 1273).[40] Let Q_3

[40] The median absolute deviation as another robust method has a breakdown point of 50 %. But 25 % is perfectly enough for the SCCIs.

be the 75th percentile and Q_1 the 25th percentile. Then the IQR is denoted by:

$$IQR = Q_3 - Q_1 \tag{5}$$

A value is perceived as an outlier if it is below $Q_1 - 1.5IQR$ or above $Q_3 + 1.5IQR$ (Calero, Moraga, & Piattini, 2008, p. 42).

The outliers are then winsorized after the determination non–normality. Winsorization converts outlying high (low) values to the next value which is high (low), but which is not an outlier anymore (Salkind, 2010, pp. 1636–1637). It implies the intuitive advantage that after winsorization, some values which are extremely good (bad), are still good (bad) without being so extreme that they distort the final results.

3. Application

Essential for the interpretation of test statistics is the determination of an appropriate significance level. Often a significance level of 0.05 is used. Nevertheless, values should only be treated when necessary. The results for the Shapiro–Wilk test (see Tables A6 and A8) in line with the Q–Q plot and with values for the skewness and kurtosis yield that a significance level of 0.01 can be deemed as appropriate. With this significance level, there are no indicators determined as from a normal distribution which contain too extreme values so that they would distort the final results. Furthermore, every indicator in the SCCI which is from a normal distribution according to the Shapiro–Wilk test is also from a normal distribution according to the skewness and kurtosis values applied by Montalto et al. (2018b, p. 2) and Dutta et al. (2018, p. 371). Those skewness and kurtosis requirements would point on more indicators to be from a normal distribution (see Tables A6,A7,A8, and A9).

The IQR detects 186 outliers for the objective SCCI and eight outliers for the subjective SCCI in respect to the non–normal indicators. The indicator with the most outliers is C15O. It contains ten outliers.[41] The indicator A1S alone

[41]The number is that high because the data for the indicator is on the national level. If a country has too extreme values, it is the same for every city within this country. According to that, already six cities from the United Kingdom contain outlying values.

accounts for half of the eight outliers of the subjective SCCI. The city with the most outliers within the non–normal indicators according to the IQR is by far Nicosia in the objective SCCI with twelve outliers and Naples as well as Palermo in the subjective SCCI with two outliers for each of these two cities.

Winsorization is conducted for all these 194 outliers of the subjective and objective SCCI. An alternative to winsorization are power transformations. They are possibly more adequate when the normality assumption is an absolute condition. But as already noted, they affect the (obligatory) normalization process and although it can be beneficial for the overall analysis to fulfill the normality assumption for all (or at least many) indicators, it is not a necessary requirement. Winsorization does also affect the normalization process. Nevertheless, running some trials with diverse power transformations techniques and comparing the values normalized by z–scores as introduced in the next subsection yields that winsorization affects the normalized values less because its influence on mean and standard deviation is rather small.

3.6 Normalization

Normalization is an obligatory procedure for the objective SCCI because the indicators are expressed in various units and scales. For the subjective SCCI this is not the case, and hence normalization is not a necessary requirement. However, for the sake of consistency and better comparability between both composite indicators, the subjective SCCI is also normalized.

1. Literature

Giffinger et al. (2007, p. 14) use z–scores. Z–scores divide the distance from an indicator value to the mean of this indicator by its standard deviation. Annoni et al. (2017, p. 19) employ weighted z–scores with population sizes as weights.

Another commonly used procedure is the max–min method. Dutta et al. (2018, p. 371) stretch a range between the maximum and minimum value of the relevant indicator. The actual indicator value is put in relation to this range by taking into account the minimum (maximum) value when more (less) is better. For index data, they use the range of the index instead of the maximum and minimum values. The normalized values can range between 0 and 100. Montalto et al. (2018b, p. 3) pursue a similar approach. Athanasoglou and Dijkstra (2014, pp. 19–20) calculate distances of actual performance values to target values. The performance relative to the target can range between a minimum of zero and a maximum of one. In Aiginger and Firgo (2015, pp. 16–17) the normalized values have a maximum of one and a minimum of zero after conducting the max–min method.

2. Theory

Talukder, Hipel, and vanLoon (2017, pp. 5–7) provide a helpful overview of normalization. They point on five commonly used normalization techniques:

- Ranking
- Distance to Target
- Min–Max
- Proportionate
- Z–score

Ranking normalization assigns a higher rank to a better indicator value. It leads to the loss of a great deal of information because the distances between indicator values become unobservable. Distance to target normalization is not applicable for the SCCI because there are no target values to rely on. The min–max normalization sets boundaries with the help of maximum and minimum values for each indicator. It is heavily dependent on those values, and another disadvantageous property is that it does not eliminate differences in variance between the indicators (Talukder et al., 2017, p. 7). The proportional normalization divides an indicator value by the sum of all indicator values. Despite some advantageous features of this method, it is unsuitable for the objective SCCI because in the objective SCCI are two indicators (B13O and B18O) which contain both, positive and negative values. The issue is that the natural boundary between zero and one would not be at hand for these indicators. It implies that the negative values which indicate a bad performance would be worse when they are transformed compared to indicators with solely positive or negative values, and vice versa.

Z–score normalization remains as another widely used alternative. Let x_i be an indicator value, $\overline{x}$ the sample mean of an indicator and σ the standard deviation of this indicator. Then z–scores are obtained by:

$$z_i = \frac{x_i - \bar{x}}{\sigma} \tag{6}$$

Z–scores have a mean of zero and a standard deviation of one. The method yields that scores from different distributions are directly comparable, it adjusts for different scales and variances, it maintains relative differences due to the linear transformation, and it diminishes the influence of extreme values (Talukder et al., 2017, p. 6). Due to major drawbacks of other methods and the convenient features of the z–score normalization, it seems most appropriate to rely on z–scores for the SCCIs.

3. Application

Critical in the application of z–score normalization is to differentiate between indicators in which a higher value displays a better performance and indicators in which a lower value displays a better performance. The latter need to be multiplied by –1 in a step prior (as it is already done in case of the SCCIs) or directly after z–score normalization so that a better z–score within each indicator points on a better performance. Furthermore, outliers have to be treated before normalization. Treating outliers after calculating the z–scores would distort their mean and standard deviation. Weighted z–scores as in Annoni et al. (2017, p. 19) are not used because the influence of the population size or other factors is unclear due to the novelty of this research project.

3.7 Weighting

By far most of the composite indicators employ equal weights for every indicator (Tate, 2012, p. 330). But apart from its simplicity and replicability, there is hardly a good reason why equal weights for every indicator are the best choice. The situation that every indicator has the same importance is in most cases rather unlikely. Hence, this subsection discusses a thoughtful weighting scheme for the SCCIs.

1. Literature

Giffinger et al. (2007, p. 14) state that they aggregate their indicators besides a small correction due to the coverage rate of an indicator “on all levels without any weighting.” It is not entirely clear what they mean as it is impossible not to apply any weights. If they use the same weights for each dimension and the same weights for each indicator within a particular dimension it implies that higher weights are assigned to indicators which are in a dimension with a lower number of indicators. Following this plausible interpretation of their imprecise methodological description yields partly substantial differences in indicator weights without a proper reason. Similarly, Athanasoglou and Dijkstra (2014, p. 23) use equal weights for each of their four objectives. Solely one objective consists of two indicators instead of one, and within this objective, the indicator weights are also equal.

Dutta et al. (2018, p. 370) assign weights of either 0.5 or 1.0 to each component in their composite indicator in a way to ensure that they obtain the highest correlation between them. Alongside, they use equal weights within each subdivision of their composite indicator. This again entails that a lower number of elements in a subdivision leads automatically to higher weights. Montalto et al. (2018b, p. 3) proceed similarly to obtain local indicator weights. Apart from that, they consult experts so that they can implement the budget allocation method to weigh their sub–indices and dimensions. Within the budget allocation process, experts assign more of a predetermined budget to elements which importance they want to stress (OECD, 2008, p. 32).

The weighting scheme in Annoni et al. (2017, pp. 16–18) implies that weights differ between observed regions as it takes into account five different regional development stages. Less developed regions get a higher weight for the ’Basic’ pillar, and vice versa. More developed region gets a higher weight for the ’Innovation’ pillar, and vice versa. Furthermore, they use weighted z–scores with the regions' population size as weights (Annoni et al., 2017, p. 19).

Aiginger and Firgo (2015, pp. 14–15) apply equal weights for each of their three pillars and use Principal Component Factor Analysis (PCFA) to determine indicator weights. PCFA groups indicators and can estimate weights according to correlations between them (OECD, 2008, p. 89).

Stanković et al. (2017, pp. 526–535) use the Analytic Hierarchy Process (AHP) from the field of multi–criteria analysis. The AHP offers a framework to determine weights based on reciprocal pairwise comparisons between different alternatives (Saaty, 2008b, pp. 1–2). However, Stanković et al. (2017) are uncritical about the large gap in indicator weights. The highest indicator weight is more than 17 % while some indicators weigh less than 1 %.

2. Theory

Weights for the indicators of the SCCIs are determinable based on different methods. Those methods are categorizable by three main groups which are equal weighting, statistic–based weighting, and public/expert opinion–based weighting (Gan et al., 2017, p. 493).[42] Equal weights are a proper choice when a statistical or empirical base is absent (OECD, 2008, p. 31). Moreover, they are simple and straightforward (Gan et al., 2017, p. 495). Let n be the number of indicators within the dimensions and the number of dimensions, respectively. Then the i^{th} local weight of an indicator is given by:

$$\eta_i = \frac{1}{n} \tag{7}$$

While the i^{th} weight of a dimension is given by:

$$\psi_i = \frac{1}{n} \tag{8}$$

However, there are major drawbacks of equal weighting due to its arbitrariness and non–transparency (Rowley, Peters, Lundie, & Moore, 2012, p. 29). Those

[42] Equal weights are strictly speaking also a form of opinion–based weighting (Mikulić, Kožić, & Krešić, 2015, p. 313).

drawbacks are immense and not justifiable for the SCCIs. Henceforth, a weighting scheme which combines statistic–based weighting and public/expert opinion–based weighting is proposed. The weighting scheme uses PCFA to assign weights to the indicators within each dimension and the AHP to assign weights to each dimension.

The usage of two widespread data structuring techniques, Principal Component Analysis (PCA) and Factor Analysis (FA) is sometimes confusing due to their similarities (Glorfeld, 1995, pp. 377–378).[43] Hence, the term PCFA describes here the conventional approach that PCA is employed to extract principal components and to consider those as factors for conducting a subsequent factor analysis (OECD, 2008, pp. 69, 89).

PCFA uses linear transformation techniques to reduce data dimensionality without losing a significant amount of information (Gan et al., 2017, p. 493). The idea of PCFA as a weighting procedure is to account for a considerable variation in the data by a small number of factors and to adjust for overlapping information (Hermans, van den Bossche, & Wets, 2008, p. 1338).[44] Factor loadings are assigned to each indicator. The indicators with the largest variation across the observations have the largest factor loadings, and thus, it is vital that there are no outliers anymore as they would have a considerable influence on the factor loadings and accordingly on the weights (OECD, 2008, p. 26). The main advantage of this statistic–based procedure is that it reduces the risk of double weighting, but on the other hand, the results are possibly illogical and differ from reality (Gan et al., 2017, p. 495).

An essential condition to conduct PCFA is that the indicators are sufficiently intercorrelated. A measure to examine an adequate fulfillment of the intercorrelation is the Kaiser–Meyer–Olkin Measure of Sampling Adequacy (KMO)

[43]See Johnson and Wichern (2007, pp. 430–538) for a detailed theoretical discussion on PCA and FA.

[44]PCFA could also be used to determine the design of the SCCIs dimensions with indicators assigned to a small number of factors which then represent the dimensions. However, the assignment of indicators to dimensions relies on discrete decisions as the ratio of observations to indicators is too low for the objective SCCI (OECD, 2008, p. 66), and indicators with the highest correlation are not necessarily those which logically represent a dimension of the SCCIs.

(Kaiser, 1970; Kaiser & Rice, 1974). Let r_{jk} be an original correlation and q_{jk} an anti–image correlation. Then the KMO is denoted by:

$$KMO = \frac{\sum \sum_{k \neq j} r_{jk}^2}{\sum \sum_{k \neq j} r_{jk}^2 + \sum \sum_{k \neq j} q_{jk}^2} \tag{9}$$

The KMO can have a range between zero and one, and a value above 0.5 is acceptable (Kaiser & Rice, 1974, pp. 112–113).[45]

If the sample is adequate (i.e., the KMO within a dimension is above 0.5), then the next step is to determine the number of factors to retain in the PCFA. The determination of the number of factors to retain follows the suggestion from the OECD (2008, p. 89). A factor is retained when

i) its eigenvalue is larger than one.

ii) it contributes more than 10 % to the explanation of overall variance.

iii) it contributes to the explanation of the overall variance so that it exceeds 60 %.

The retained factors are then rotated by using varimax rotation as it is common practice (OECD, 2008, p. 90). The handbook of the OECD (2008, p. 90) next recommends calculating the weights by grouping the indicators with the highest factor loadings, squaring and scaling them to unity sum and then multiplying those values by the proportion of the variance which the factor explains in respect to the total variance explained by the retained factors.[46] This approach is appropriate when another aim is to construct intermediate composites, and also, it is adopted in the literature (see Aiginger and Firgo (2015), and Sharpe and Andrews (2012)). However, as the construction of intermediate composites is not an aim of the SCCIs, it would neglect some information because it

[45]The handbook from the OECD (2008, p. 67) refers incorrectly to Kaiser and Rice (1974) and says that they propose a value of 0.6.

[46]The description of the OECD (2008, p. 90) is not entirely clear as the treatment of the indicator 'Tech Exports' shows. They follow Nicoletti, Scarpetta, and Boylaud (2000) and state to group the indicators with the highest factor loadings. Either they refer solely to the non–negative factor loadings or to the squared factor loadings scaled to unity sum. This is not consistent with Nicoletti et al. (2000, p. 27) as their indicator 'Price controls' shows.

ignores a lot of factor loadings. Hence, the SCCIs follow other approaches and descriptions which take into account all factor loadings of the retained factors (see Feroze and Chauhan (2010), Raihan (2011), Salvati and Carlucci (2014), and Sands and Podmore (2000)). Let δ_j be the proportion of the explained variance of factor j in respect to all retained factors J, l_{ij} the factor loading of the i^{th} indicator on factor j and E_j the variance explained by the factor j (Gan et al., 2017, p. 495). Then the local weight for each indicator is calculated by:

$$\eta_i = \delta_j \frac{l_{ij}^2}{E_j} \tag{10}$$

Weights for the dimensions are obtained next via the Analytic Hierarchy Process (AHP).[47] The fundamentals of the AHP were developed by Saaty (1977). Instead of judging the relative importance of indicators and dimensions intuitively, the AHP offers a scientific design in which reciprocal pairwise comparisons take place (Saaty, 2008b, p. 2). The design requests a problem definition and a hierarchy structure which considers goals, criteria, sub–criteria, and alternatives.

The hierarchy structure is vital to conduct the reciprocal pairwise comparisons in which matrices that contain all the relevant elements and a scale with values from one to nine are used to point out “how many times more important or dominant one element is over another element with respect to the criterion or property with respect to which they are compared” (Saaty, 2008a, p. 85). The reciprocal automatically enters the matrix on the opposing side of the diagonal. Table 3 shows the scale of dominance and importance based on Saaty (1990) in which the intensity of importance is assessed with numbers between one and nine. One indicates that both elements are equally important and nine points on the extreme importance of one element over another.

The weights are obtained by standardizing the matrices. These weights can be checked for consistency. They are consistent when the consistency ratio (CR)

[47] There are more extensive frameworks to apply the AHP available than it is expedient to carry out here (see Maletič, Maletič, Lovrenčić, Al–Najjar, and Gomišček (2014), and Rangone (1996)).

Table 3: Scale of Dominance and Importance

Intensity of Importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective.
3	Moderate importance	Experience and judgment slightly favor one activity over another.
5	Strong importance	Experience and judgment strongly favor one activity over another.
7	Very strong importance	An activity is favored, and its dominance demonstrated in practice.
9	Extreme importance	Favor of one activity over another is the highest possible.
2,4,6,8	Intermediate values	When compromise is needed.

is less than or equal to 0.1 (Saaty, 1990, p. 13).[48]

The AHP provides credence and transparency. Nevertheless, an issue is the identification of proper values for the reciprocal pairwise comparisons. These values will always remain subjective (Saaty, 2008a, pp. 85, 95–97).[49] After the weights for the indicators within a dimension and the weights of the dimensions are calculated, the overall indicator weight is computed by simply multiplying the local indicator weights with the weights of the dimensions. Let η_i be the weight of the i^{th} indicator within the dimensions and ψ_i the weight of the i^{th} dimension. Then the global weight for each indicator is given by:

$$\omega_i = \eta_i \psi_i \tag{11}$$

[48]For a detailed discussion on the calculation of the CR, see Alonso and Lamata (2006).

[49]For more criticism on the AHP, see Department for Communities and Local Government (2009).

As a first step, the KMO is computed (see Tables A10 and A11). The overall values for the KMO are above 0.5 in each dimension of the SCCIs, and accordingly, it is adequate to conduct the PCFA. The values for the subjective SCCI are generally better than for the objective SCCI. They can be improved by excluding the individual indicators with the lowest KMO. However, there is a trade–off because every indicator is selected carefully and its exclusion comes solely at other disadvantages which is why the KMO values are not further improved here. Furthermore, Aiginger and Firgo (2015, p. 52) report lower KMO values.

After demonstrating that each dimension is adequate to conduct the PCFA, the number of retained factors is determined (see Table A12 and A13). The number of retained factors in the subjective SCCI is naturally lower as it contains fewer indicators in each dimension. Solely one factor remains in dimension D of the subjective SCCI whereas six factors remain in every dimension of the objective SCCI.

The factor loadings of the retained factors (see Table A14 and A15) are then squared and scaled to unity sum. The normalized factor loadings are next set in relation to the proportion of the explained variance by each retained factor to determine the local weights of each indicator.

The weights for the dimensions are obtained via the Analytic Hierarchy Process (AHP). First, a hierarchy structure is established with which help the intensities of importance are assessed (see Figure 1).

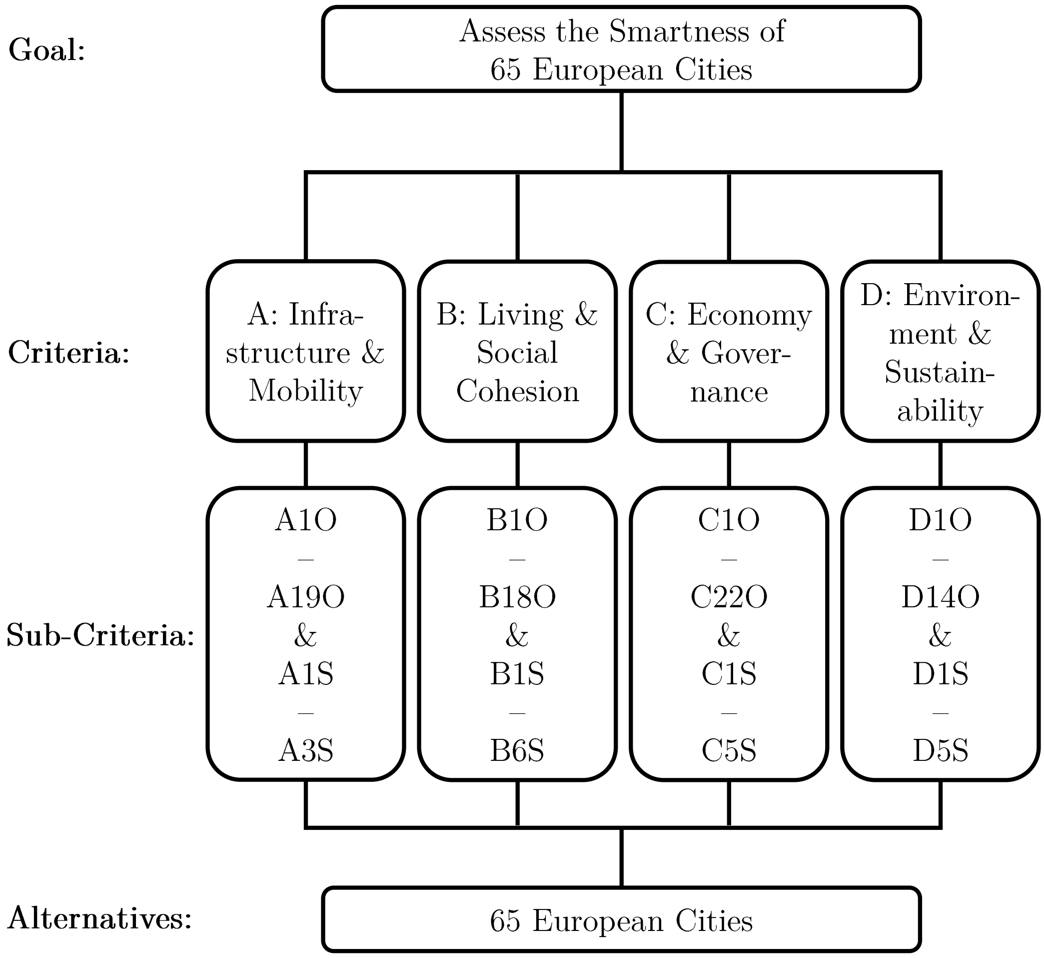

Figure 1: Hierarchy Structure for the SCCIs

The intensities of importance for the pairwise comparisons within the AHP (see Table A16) represent the importance of one dimension compared to another and take several aspects into account. One aspect is that the European Urban Audit 2015 asks the respondents which issues they perceive as important. Other aspects for the decision about the intensities of importance consider a broad knowledge about these kinds of composite indicators which are primarily based on objective data. Furthermore, the quality of the indicators in every dimension is borne in mind. All in all, the CR (0.045) indicates that the intensities of importance are consistent.

As the last step, the global weights are calculated by multiplying the local weights for each indicator with weights for the dimensions. Tables 4 and 5 show the results. The local weights do not vary so much within each dimension of the subjective SCCI compared to the objective SCCI. This is because the indicator values of the subjective SCCI are more highly correlated and no indicator provides a lot of different values compared to others.

Table 4: Indicator Weights of the Objective SCCI

	Weight in %		
Indicator	Dim.	Local	Global
A1O	19.81	4.77	0.94
A2O	19.81	6.24	1.24
A3O	19.81	5.21	1.03
A4O	19.81	5.04	1.00
A5O	19.81	6.23	1.23
A6O	19.81	5.15	1.02
A7O	19.81	6.27	1.24
A8O	19.81	5.86	1.16
A9O	19.81	5.11	1.01
A10O	19.81	5.01	0.99
A11O	19.81	4.97	0.99
A12O	19.81	5.21	1.03
A13O	19.81	5.49	1.09
A14O	19.81	5.19	1.03
A15O	19.81	4.20	0.83
A16O	19.81	4.01	0.79
A17O	19.81	5.72	1.13
A18O	19.81	5.00	0.99
A19O	19.81	5.32	1.05
B1O	27.48	5.30	1.46
B2O	27.48	6.17	1.70
B3O	27.48	4.10	1.13
B4O	27.48	5.30	1.46
B5O	27.48	6.14	1.69
B6O	27.48	4.81	1.32
B7O	27.48	6.42	1.76
B8O	27.48	6.26	1.72
B9O	27.48	6.18	1.70
B10O	27.48	5.68	1.56
B11O	27.48	5.56	1.53
B12O	27.48	6.42	1.76
B13O	27.48	3.23	0.89
B14O	27.48	5.54	1.52
B15O	27.48	5.66	1.56
B16O	27.48	6.29	1.73
B17O	27.48	5.17	1.42
B18O	27.48	5.77	1.59

	Weight in %		
Indicator	Dim.	Local	Global
C1O	38.73	4.75	1.84
C2O	38.73	5.09	1.97
C3O	38.73	4.23	1.64
C4O	38.73	4.66	1.80
C5O	38.73	3.97	1.54
C6O	38.73	5.14	1.99
C7O	38.73	4.83	1.87
C8O	38.73	4.66	1.80
C9O	38.73	3.46	1.34
C10O	38.73	5.11	1.98
C11O	38.73	3.00	1.16
C12O	38.73	3.50	1.16
C13O	38.73	4.69	1.82
C14O	38.73	5.49	2.13
C15O	38.73	5.06	1.96
C16O	38.73	4.45	1.72
C17O	38.73	4.04	1.57
C18O	38.73	4.20	1.63
C19O	38.73	4.71	1.82
C20O	38.73	5.21	2.02
C21O	38.73	4.68	1.81
C22O	38.73	5.07	1.96
D1O	13.97	7.60	1.06
D2O	13.97	7.34	1.03
D3O	13.97	6.83	0.95
D4O	13.97	8.35	1.17
D5O	13.97	7.18	1.00
D6O	13.97	8.46	1.18
D7O	13.97	8.45	1.18
D8O	13.97	6.99	0.98
D9O	13.97	6.78	0.95
D10O	13.97	6.49	0.91
D11O	13.97	7.13	1.00
D12O	13.97	6.42	0.90
D13O	13.97	6.27	0.88
D14O	13.97	5.68	0.79

Table 5: Indicator Weights of the Subjective SCCI

	Weight in %				Weight in %		
Indicator	Dim.	Local	Global	Indicator	Dim.	Local	Global
A1S	19.81	35.53	7.04	C1S	38.73	18.94	7.34
A2S	19.81	32.23	6.39	C2S	38.73	20.11	7.79
A3S	19.81	32.24	6.39	C3S	38.73	20.15	7.80
				C4S	38.73	19.95	7.73
B1S	27.48	16.50	4.54	C5S	38.73	20.85	8.08
B2S	27.48	16.23	4.46				
B3S	27.48	13.01	3.58	D1S	13.97	20.94	2.93
B4S	27.48	18.53	5.09	D2S	13.97	18.89	2.64
B5S	27.48	18.20	5.00	D3S	13.97	22.75	3.18
B6S	27.48	17.54	4.82	D4S	13.97	20.29	2.83
				D5S	13.97	17.13	2.39

3.8 Aggregation

Aggregation is the last step to construct the SCCIs. Alongside with weighting, aggregation generally has the largest impact on the outcome of composite indicators (Athanasoglou & Dijkstra, 2014, p. 29). Thus, a closer look at this aspect follows.

1. Literature

Most contributions which are considered here use the simple additive aggregation method to sum the normalized and weighted individual indicators (Aiginger & Firgo, 2015, p. 14; Athanasoglou & Dijkstra, 2014, p. 20; Dutta et al., 2018, p. 371; Giffinger et al., 2007, p. 14; Montalto et al., 2018b, p. 3).[50] Solely Stanković et al. (2017, p. 535) take a different approach by applying the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). The TOPSIS assess those alternatives best which have the shortest distance to a positive–ideal solution and the farthest distance to a negative–ideal solution (Opricovic & Tzeng, 2004, p. 448).

[50] Annoni et al. (2017) do not explicitly state their aggregation method. However, as they provide their data online, it is possible to reconstruct their index which then shows that they also use the simple additive aggregation method.

Non–compensatory and compensatory aggregation methods are distinguishable (Gan et al., 2017, pp. 497–498). Non–compensatory methods are either based on the properties of aggregation functions or the perspective of multi–criteria decision making (Gan et al., 2017, p. 498). The idea behind non–compensatory aggregation methods is that increases in values of one indicator cannot offset decreases in values of another indicator (R. Greene, Devillers, Luther, & Eddy, 2011, p. 413). In some situations this property is preferable. However, non–compensatory aggregation methods come at the price of a huge loss of information because they do not consider the intensity of preferences (Munda & Nardo Michaela, 2005, p. 15). For the SCCIs this would mean that it is not important how much better or worse a city performs in the indicators. Due to this major drawback solely compensatory methods are taken into account for the SCCIs.

Despite the theoretical assumption of preferential independence which is unlikely to meet (Gan et al., 2017, p. 497), compensatory aggregation methods are more realistic because they imply a trade–off between indicator values (R. Greene et al., 2011, p. 416). The most prominent compensatory aggregation methods are the simple additive aggregation and the simple geometric aggregation (Zhou, Fan, & Zhou, 2010, p. 361). TOPSIS is also a compensatory aggregation method but is less common (Zhou & Ang, 2009, p. 88). Zhou and Ang (2009) provide a measure to compare compensatory methods. Based on an information loss concept (Zhou, Ang, & Poh, 2006), their results suggest that the simple geometric aggregation leads to the least loss of information followed by the simple additive aggregation while TOPSIS performs worse (Zhou & Ang, 2009, p. 93). TOPSIS is therefore not used as an aggregation method for the SCCIs. Central to the usage of the simple geometric aggregation is that it allows compensation between indicators solely within limitations as very low values in an indicator can not that easily be compensated by high values in other indicators (Gan et al., 2017, p. 497). Let ω_i be the i^{th} indicator weight and x_i the i^{th} indicator value. Then the score for the i^{th} city in the objective as well as in the subjective SCCI is given by:

$$SCCI_i = \prod_i^n x_i^{\omega_i} \tag{12}$$

Instead, the simple additive aggregation implies a full trade–off among indicator values (R. Greene et al., 2011, p. 420). This property makes the simple additive aggregation the best suited for the SCCIs because there is no reason apparent why very bad values should be more important in the calculation than very good values. Let ω_i be the i^{th} indicator weight and x_i the i^{th} indicator value. Then the score for the i^{th} city in the objective as well as in the subjective SCCI is given by:

$$SCCI_i = \sum_i^n x_i \omega_i \tag{13}$$

3. Application

The simple additive aggregation is used to calculate the scores for the objective and subjective SCCI. Tables 6 and 7 show the results. Additionally, the results for each dimension are in the appendix (see Tables A17 and A18).

Table 6: Results of the Objective SCCI

#	City	Score	#	City	Score
1	Prague	0.5485	34	Vilnius	–0.0260
2	Stockholm	0.5378	35	Rennes	–0.0400
3	Munich	0.5250	36	Bologna	–0.0423
4	Luxembourg	0.5057	37	Kraków	–0.0533
5	Graz	0.5002	38	Budapest	–0.0613
6	Bratislava	0.4928	39	Ostrava	–0.0928
7	Helsinki	0.4566	40	Glasgow	–0.1034
8	Vienna	0.4538	41	Nicosia	–0.1104
9	Amsterdam	0.4264	42	Newcastle	–0.1224
10	Paris	0.4229	43	Riga	–0.1245
11	Copenhagen	0.4177	44	Barcelona	–0.1370
12	Hamburg	0.3736	45	Manchester	–0.1450
13	Berlin	0.3057	46	Rome	–0.1743
14	Leipzig	0.2753	47	Sofia	–0.1808
15	Ljubljana	0.2609	48	Liège	–0.1818
16	Dublin	0.2280	49	Gdańsk	–0.1877
17	Rostock	0.2179	50	Belfast	–0.1974
18	Antwerp	0.2132	51	Valletta	–0.2083
19	London	0.2121	52	Lille	–0.2170
20	Warsaw	0.1274	53	Athens	–0.2256
21	Brussels	0.1067	54	Oviedo	–0.2440
22	Bordeaux	0.0948	55	Zagreb	–0.2512
23	Rotterdam	0.0774	56	Turin	–0.3141
24	Malmö	0.0752	57	Białystok	–0.3336
25	Groningen	0.0693	58	Málaga	–0.3347
26	Aalborg	0.0513	59	Košice	–0.3632
27	Lisbon	0.0378	60	Bucharest	–0.4275
28	Madrid	0.0371	61	Miskolc	–0.4997
29	Strasbourg	0.0141	62	Cluj–Napoca	–0.5659
30	Tallinn	–0.0046	63	Naples	–0.6251
31	Verona	–0.0062	64	Burgas	–0.6574
32	Marseille	–0.0139	65	Palermo	–0.7761
33	Cardiff	–0.0169			

Table 7: Results of the Subjective SCCI

#	City	Score	#	City	Score
1	Aalborg	1.2977	34	Prague	0.0452
2	Graz	0.9329	35	Dublin	0.0205
3	Munich	0.9171	36	Ostrava	–0.0713
4	Vienna	0.9029	37	Berlin	–0.0888
5	Luxembourg	0.8696	38	Kraków	–0.0980
6	Groningen	0.8628	39	Zagreb	–0.1312
7	Cardiff	0.7974	40	Valletta	–0.1619
8	Copenhagen	0.7410	41	Lille	–0.1771
9	Belfast	0.7299	42	Málaga	–0.2285
10	Stockholm	0.7227	43	Paris	–0.2482
11	Newcastle	0.6829	44	Riga	–0.2515
12	Helsinki	0.6771	45	Warsaw	–0.2925
13	Glasgow	0.6563	46	Verona	–0.3047
14	Malmö	0.5755	47	Brussels	–0.3064
15	Rennes	0.5684	48	Budapest	–0.3574
16	Hamburg	0.5645	49	Barcelona	–0.3956
17	Rostock	0.5424	50	Košice	–0.4162
18	Leipzig	0.5075	51	Liège	–0.4312
19	Manchester	0.4905	52	Bologna	–0.5068
20	Antwerp	0.4731	53	Miskolc	–0.5809
21	Rotterdam	0.4293	54	Nicosia	–0.6364
22	Bordeaux	0.4181	55	Bratislava	–0.7289
23	Ljubljana	0.3686	56	Turin	–0.7830
24	Cluj–Napoca	0.3624	57	Marseille	–0.8192
25	Amsterdam	0.3377	58	Madrid	–0.8647
26	London	0.3339	59	Lisbon	–0.8851
27	Strasbourg	0.3054	60	Bucharest	–0.9239
28	Białystok	0.2854	61	Sofia	–0.9674
29	Oviedo	0.2794	62	Rome	–1.6344
30	Tallinn	0.2142	63	Naples	–1.6372
31	Burgas	0.1927	64	Athens	–1.6602
32	Vilnius	0.1434	65	Palermo	–1.7991
33	Gdańsk	0.1395			

3.9 Uncertainty and Sensitivity Analysis

There exists a consensus on the importance of uncertainty analysis (UA) and sensitivity analysis (SA), but often models lack this step (Hermans, van den Bossche, & Wets, 2009, p. 1223). Therefore, this subsection SCCIs provides uncertainty and sensitivity analysis results.

1. Literature

Athanasoglou and Dijkstra (2014, p. 29) employ an UA and a SA in which they use Monte Carlo experiments to test the robustness of their modeling choices with respect to weights and aggregation. Dutta et al. (2018, p. 76) also use Monte Carlo simulations, and besides weights and aggregation, they additionally assess the impact of the treatment of missing values on their composite indicator. Furthermore, Montalto, Jorge Tacao Moura, Langedijk, and Saisana (2018c, p. 6) apply Monte Carlo simulations. They run it on their normalization and weighting choices.

Aiginger and Firgo (2015), Annoni et al. (2017), Giffinger et al. (2007) and Stanković et al. (2017) do not offer an UA and a SA.

2. Theory

UA and SA are related and should be run in tandem (Saltelli et al., 2008, p. 1). UA concentrates on uncertainties in the inputs (i.e., the different stages to construct the composite indicator) and how they conjunctively affect the values of the composite indicator, whereas SA identifies how much each source of uncertainty contributes to the variance of the composite indicator values (Greco, Ishizaka, Tasiou, & Torrisi, 2018, p. 21; Saisana, Saltelli, & Tarantola, 2005, p. 308). Uncertainties arise from all steps in the construction line of the composite indicator (Saisana et al., 2005, p. 309). Uncertainty and sensitivity analysis should thus ideally concentrate on all of these steps. However, the choices of a weighting scheme and an aggregation method generally have the largest impact on composite indicators (Athanasoglou & Dijkstra, 2014, p. 29). Therefore, an UA and a SA on those stages in the construction of the SCCIs are conducted.

Even though other authors in a similar strand of literature rely on Monte Carlo simulations (Athanasoglou & Dijkstra, 2014, p. 29; Dutta et al., 2018, p. 76; Montalto et al., 2018c, p. 6), the UA and SA here use one–factor–at–a–time (OAT). The idea behind OAT is to change one option at a time within a particular construction stage while other stages are held constant (Tate, 2012, p. 331). In general, OAT is the most popular SA technique (Saltelli & Annoni, 2010, p. 1509). Despite several drawbacks in respect to the theoretical groundings and meaningfulness of this approach (Saltelli & Annoni, 2010, pp. 1508–1510; Saltelli et al., 2008, pp. 66–76; Tate, 2012, pp. 331–332), it has some endorsing features. OAT enables to precisely determine the effect of changing a specific option at the construction stage (Saltelli et al., 2008, p. 75). Furthermore, OAT does not involve any noise when there is no stochastic term, the sensitivities refer to the same starting point, and it never detects irrelevant changes as influential (Saltelli & Annoni, 2010, p. 1510).

With OAT, UA and SA can be done at the same time as in Hudrliková (2013, pp. 470–471). UA is done with OAT in the sense that there are in the end several different scenarios which are observable. Moreover, descriptive statistics with respect to deviations from the baseline scenario can be reported. SA is done in the sense that there is always a scenario which is the same despite one change. This offers the possibility to investigate the influence of specific changes closely.

Lastly, the validation of the modeling choices is done via correlation coefficients. Therefore, Pearson correlation coefficients (PCCs) need to be calculated (see Pearson (1895)). Correlation coefficients enable to get insights about the strength and direction of the linear relationship between pairs of variables (Mukaka, 2012, p. 71). If the correlation coefficients are high, the SCCIs are insensitive to uncertainties in modeling choices (Tate, 2012, p. 331).

3. Application

As discussed, the OAT is applied to weighting and aggregation as the most influential construction stages of composite indicators. In OAT it is crucial to limit the focus on some construction choices against which the SCCIs as baseline scenarios are tested. Without a limitation on construction stages as well as on choices within the construction stage, OAT would lead to a few hundred or even a thousand cases. The examination of a hundred or a thousand of cases would be confusing and is also beyond the scope here in respect to computational costs.

The UA and SA for the SCCIs contemplate twelve different cases with respect to three construction stages, including the baseline scenario which represents the choices made for the SCCIs. The construction stage weighting is separated in two steps because the weighting scheme for the SCCIs implies PCFA to weigh indicators within the four different dimensions and the AHP to weigh dimensions.[51] In each of the three construction stages, different modeling choices are considered beneath the decisions for the SCCIs. Those alternative modeling choices are now introduced.[52]

In the first weighting stage, the factor analysis approach for weighting proposed by the OECD (2008, pp. 89–91) is contemplated due to the widespread influence of its handbook on the development of composite indicators. Furthermore, the option that equal weights are assigned to every indicator within each dimension is examined during this stage as it is the most commonly weighting procedure. In the second weighting stage, again, equal weights are looked at among the baseline scenario as the most common weighting procedure. Recall here that equal weights for each dimension imply higher weights for indicators within a dimension with a smaller number of indicators. In the aggregation stage, the simple geometric aggregation is considered as a different option compared to the simple additive aggregation because it provides interesting insights about the outcome for the SCCIs when the idea is that low indicator values can not be compensated so easily by higher values in other indicators. The simple geometric aggregation also has the best performance of common compensatory aggregation methods in terms of information loss. Furthermore, the mentioned ineptitude of non–compensatory methods for the SCCIs excludes those aggre-

[51] Aggregation could also be divided into two steps. One step for aggregating the indicators and the other for weighting the dimensions. However, due to the limitations of the number of scenarios, the aggregation method for the first step in weighting is here always the same as for the second step in weighting.

[52] See the theoretical parts in Subsection 3.7 and Subsection 3.8 for more details on the chosen alternatives.

gation procedures. The application of the simple geometric aggregation is not directly possibly because z–scores are used for normalization and thus, negative values would distort the results. Therefore, it is first necessary to get positive values. Since the lowest z–score in the SCCIs is –3.34, the value five is simply added to every single z–score as suggested by Saisana (2012, p. 6).

Figure 2 displays the different modeling choices for the UA and SA which leads to twelve scenarios. ∅ describes the equal weighting procedure, Σ stands for the simple additive aggregation method, and Π for the simple geometric aggregation method. Furthermore, FA (OECD) points on the factor analysis weighting as suggested by the OECD (2008) while PCFA and AHP denote the two weighting procedures implemented in the SCCIs. Scenario 1 as the baseline scenario is bold and represents the choices made for the SCCIs.

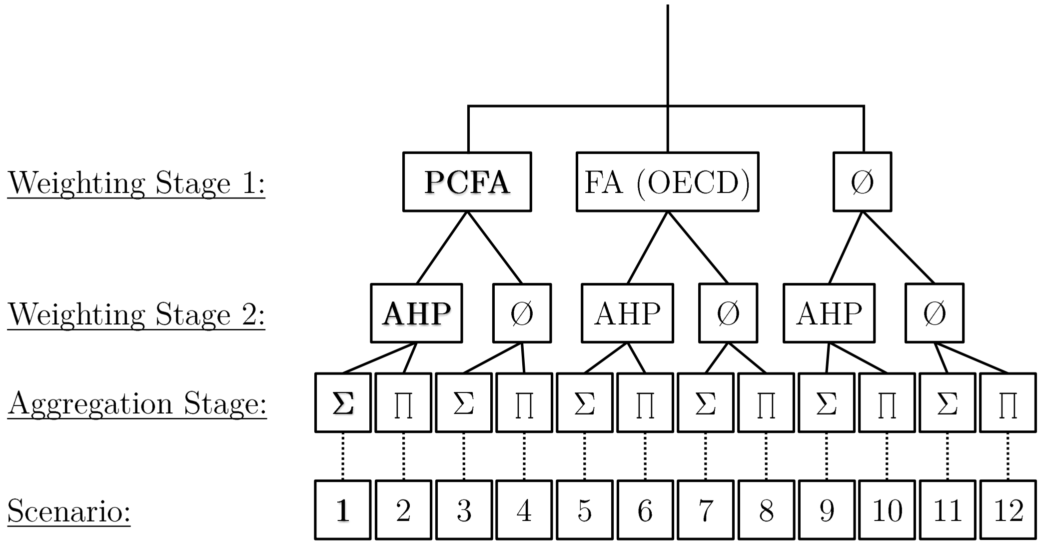

Figure 2: Scenarios of the UA and SA

Figures 3 and 4 portray the results of the UA and SA graphically to facilitate its evaluation. They display the range between the minimum and the maximum rank for each city in the twelve scenarios while the triangles point on the scenarios' mean rank. Figures 3 and 4 are both ordered according to their outcome in the baseline scenario, starting with Prague on the top left in the case of the objective SCCI, and with Aalborg in case of the subjective SCCI.

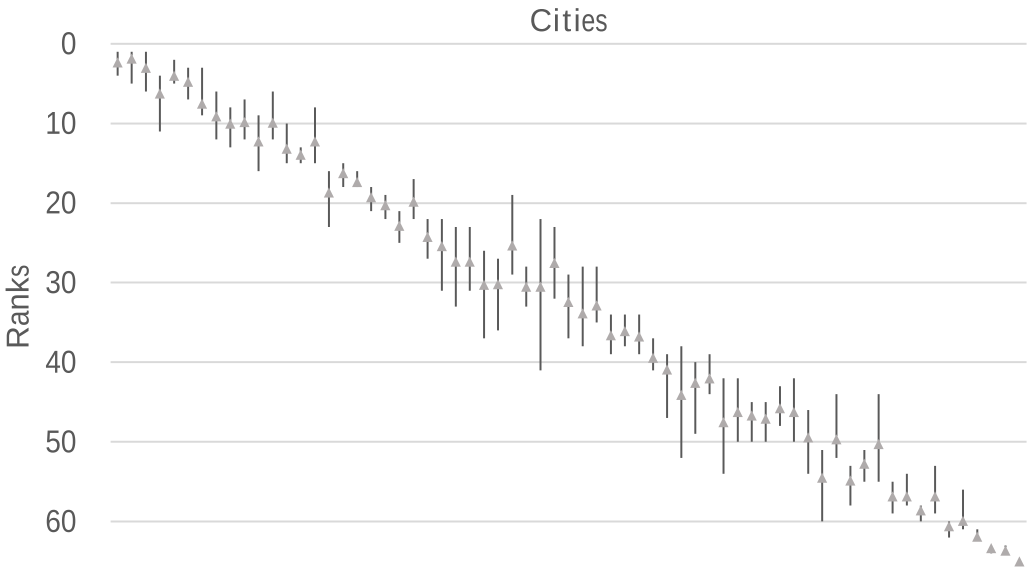

Figure 3: UA and SA Results of the Objective SCCI

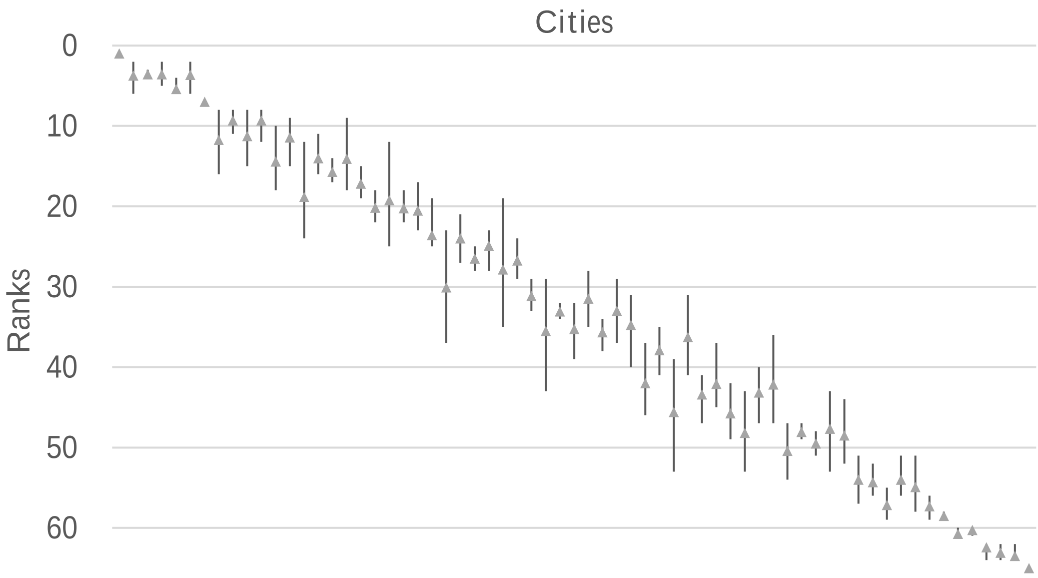

Figure 4: UA and SA Results of the Subjective SCCI

In both SCCIs, objectively and subjectively, cities with an average performance are most sensitive to the modification of modeling choices. Notwithstanding the scenario, cities on top and at the bottom of the SCCIs have a similar smartness. A plausible explanation is that the cities' scores in the SCCIs are quite close to each other around the average ranks and more spread at the top and the bottom (see Tables 6 and 7).

The exact results of the UA and SA are shown in Tables A19 and A20. Solely the ranks are reported there because the scores are not easily comparable because the scenarios which use the simple additive aggregation all have a zero mean, but the scenarios which use the simple geometric aggregation each time have a different mean. The widest range between a minimum and a maximum rank according to both SCCIs is Verona. Verona is ranked 31^{st} in the objective SCCI but would be ranked 22^{nd} in Scenario 7 and 41^{st} in Scenario 8. The only city which is ranked each time the same in the objective SCCI is Palermo on the last rank. The cities which are ranked each time the same in the subjective SCCI are Aalborg on the first, Cardiff on seventh, and again, Palermo on the last rank. While Aalborg substantially outperforms the other cities in the subjective SCCI and Palermo underperforms according to both SCCI, the fact that Cardiff is on the same rank in each scenario is most likely by chance.

Overall, the UA and SA confirm that the outcome of the SCCIs are not too much dependent on the modeling choices because the PCCs between the baseline scenario and the other eleven scenarios are all well above 0.95 (see Tables A21 and A22).[53]

PCCs between the twelve scenarios are calculated with total scores and not with the ranks. However, as already mentioned, these scores are not provided because they are not very indicative as they differ by definition quite substantially between the simple additive aggregation and the simple geometric aggregation.

[53]For illustrative reasons, the entire PCC matrix is displayed in Tables A21 and A22.

4 Discussion of the Smart City Composite Indicators

This section discusses the results of the SCCIs. First, a closer look on connections within and between the SCCIs and their dimensions is provided. Second, the SCCIs are compared descriptively due to population size and capital status of the cities.[54]

4.1 Conjunctions Within and Between the SCCIs

This subsection wants to discuss the relationships within and between the SCCIs, mainly based on a PCC matrix.

Recall that Tables 6 and 7 in Subsection 3.7, as well as Tables A18 and A19, show the results of the objective and subjective SCCI. There are some similarities and dissimilarities which can be observed at first sight. For example, Palermo is on the last rank according to both composite indicators. Prague, on the other hand, is on the first rank with respect to the objective SCCI and on the 34$^{\text{th}}$ rank with respect to the subjective SCCI. Another interesting aspect is that the spread of the scores is lower in the objective SCCI compared to the subjective SCCI. Noteworthy is also that in the subjective SCCI, the score for Aalborg as the first ranked city and the scores for Rome, Naples, Athens, and Palermo as the cities at the bottom differ a lot from the scores of their adjacent ranks.

Due to the sample size, it is difficult to place more precise remarks by solely looking at the results for the SCCI. Therefore, a PCC matrix is calculated. Table 8 shows the PCCs while O stands for the objective SCCI, S stands for the subjective SCCI, and AO, BO, CO, DO, AS, BS, CS as well as DS point on the different dimensions of the SCCIs.

[54] Population size and capital status are depicted here, but several other aspects are open to a closer investigation in future work.

Table 8: Pearson Correlation Coefficients of the SCCIs Performances

	O	S	AO	BO	CO	DO	AS	BS	CS	DS
O	1									
S	0.51	1								
AO	0.62	0.02	1							
BO	0.91	0.43	0.59	1						
CO	0.92	0.55	0.42	0.73	1					
DO	0.24	0.38	–0.29	0.14	0.19	1				
AS	0.53	0.92	0.07	0.45	0.55	0.35	1			
BS	0.34	0.93	–0.05	0.26	0.40	0.31	0.84	1		
CS	0.55	0.96	0.04	0.48	0.60	0.35	0.81	0.84	1	
DS	0.43	0.93	–0.02	0.37	0.44	0.43	0.88	0.85	0.83	1

The relationship between the objective and subjective SCCI is positive with medium strength (0.51). It points out that a city with a good performance in the objective SCCI generally speaking also performs well in the subjective SCCI, and vice versa. Interestingly, dimension DO is solely weakly correlated with the objective SCCI (0.24). This is remarkable because dimension DO itself contributes to the outcome of the objective SCCI and entails that the correlation between DO and the other dimensions of the objective SCCI is weak or negative.

The relationships within each SCCI provide further insights. Despite the correlation between AO and DO (–0.29), the dimensions of the objective SCCI are all positively correlated. The largest correlation in the objective SCCI is between BO and CO (0.73). Both results are not that surprising. If a city performs well in the dimension ’Infrastructure & Mobility’, it is plausible that this could be at the expense of indicators which contribute to the dimension ’Environment & Sustainability’, and vice versa. This interpretation implies a slight trade–off between both dimensions. Just the same idea applies to the high correlation between BO and CO. Note that the PCC does not directly give evidence about causalities between dimensions. However, an intuitive explanation is that a good economy and governance as well as sound living conditions and social cohesion are mutually dependent. Another result of the PCC matrix is that all dimensions of the subjective SCCI are highly correlated with each other (all above 0.8). This is intriguing and necessarily leads to question if surveys are able to capture detailed information about personal impressions or if they, in fact, measure a broader picture. It could be the case that people are not thinking deeply about every question but report more generally about their satisfaction with certain aspects in the city which they live in.

Lastly, the relationships between the dimensions of the objective and the subjective SCCI are of interest. Again, dimension AO is the only one which is negatively correlated with other dimensions (–0.05 with BS, –0.02 with DS). AO is very weakly correlated with every dimension of the subjective SCCI (all below $|0.1|$), including the correlation with AS (0.07). This property is unfavorable in the sense that there is as good as no relationship between the objective measure of 'Infrastructure & Mobility' and the subjective perception of this dimension. A look on the indicators can provide possible explanations for this issue. One challenge is always to find subjective and objective indicators which measure at least broadly the same aspects. This is solely broadly satisfied concerning the two dimensions AO and AS, and an improvement of the indicators could also improve the results. Additionally, subjective perceptions are capable to assess the supply against the requirements for the city whereas the rationale behind the objective SCCI is always that more is better. When looking at the indicator values, especially of those related to transportation, it could simply be that for example in Rostock (40^{th} in AO, sixth in AS), the performance in AO is not that good because not so many means of transport such as trains and planes arrive and depart in and around the city. But people at the same time perceive that they do not live in a large metropolitan region so that they do not require the city to provide a sophisticated public transportation network. A lot of people from Rostock which took part in the survey for the subjective SCCI do maybe seldom use a different way to travel than to go by car. Contrary to that, Paris is ranked first in AO and 41^{th} in AS. Even though people objectively have broad access to many ways of transportation, they are unsatisfied with public transportation. This could be due to overcrowding, frequent delays, and dependency on the use of public transportation.

Furthermore, it is worth noting that despite dimensions BO and BS, every dimension of the objective SCCI has the largest correlation with its correspondent dimension of the subjective SCCI (i.e., AO with AS, CO with CS, and DO with DS). The result that three of four dimensions of the objective SCCI have the strongest positive relationship with their correspondences is highly plausible. Additionally, the correlations of CO with the dimensions of the subjective SCCI are all quite high. Looking at the correlations between the dimensions the other way around is more ambiguous. AS for example is substantially higher correlated with all other dimensions of the objective SCCI than with AO. This is because the outcome of all AS, BS, CS, and DS are similar as the PCCs show (all above 0.8) and as stated before one issue could be that the questions measure a broader picture and not what they in fact ask for.

4.2 Descriptive Comparisons of the SCCIs

The following subsection aims at closer discussing the results of the SCCIs descriptively. An evident way to start a descriptive contemplation is to differentiate between capital and non–capital cities because the sample contains all 28 capital cities of the EU. Thereafter, cities are compared according to their size. This approach is inspired by Montalto et al. (2018a, p. 23).

Table 9 reports the mean ranks and the mean scores for capital as well as for non–capital cities within the objective and the subjective SCCI as measures of central tendency. Table A1 denotes in the column 'Metropolitan Code' which cities are capitals and which are non–capitals. A drawback is that for some countries, there is solely the capital city in the sample and no non–capital city. This could bias the results. Nevertheless, the means indicate that capital cities perform better according to the objective SCCI. But at the same time, non–capital cities perform better according to the subjective SCCI. This is especially surprising when recalling that the correlation between both SCCIs is positive in a medium size (0.51). An explanation for this feature of the SCCIs is not entirely possible here. However, some hypothetical attempts to explain the difference are provided.

Table 9: Mean Ranks and Scores of Capital and Non–Capital Cities

	Objective SCCI		Subjective SCCI	
	Capital	Non–Capital	Capital	Non–Capital
N	28	37	28	37
Ø City Size	1,784,452	538,447	1,784,452	538,447
Ø Rank	25.54	38.65	37.86	29.32
Ø Score	0.14	–0.10	–0.17	0.13

Capitals often get a disproportionate amount of the resources and have agglomeration advantages (J. V. Henderson, 2010, p. 529; Parkinson, Meegan, & Karecha, 2015, p. 1056).[55] Moreover, capitals are substantially larger than non–capitals. Due to that, they have a more skilled mix of employees (Elvery, 2010, p. 377) and it is easier for them to reach a critical mass for new digital solutions which can then contribute to the smartness of the city (Neirotti et al., 2014, p. 29). Therefore, it is evident to suggest that it is easier for capitals to make their city smart from an objective point of view. Contrary to that, some disadvantages of living in the capital city are potentially well captured by perceptional data. Crowds of tourists, overfull streets on permanently occurring events, higher rents and other aspects which include negative externalities could offset the positive effects. Those and other diseconomies of scale could make large cities less smart (Neirotti et al., 2014, p. 29). The residents of capitals also distrust that the local administration has their interests predominantly in mind (de Vries & Sobis, 2018, p. 225). The distrust could influence other indicators and dimensions as well.

Those basic attempts to explain why the ranks and scores of the SCCIs according to the mean give a first idea about possible reasons for their opposite outcome. Furthermore, Table 9 also indicates that the capital cities in the sample on average have substantially more inhabitants than the non–capital cities (1,784,452 inhabitants compared to 538,447 inhabitants). For this reason, the cities are further differentiated descriptively with respect to the number of their inhabitants.

[55] They are typically the political center, contain the best universities and research institutions, can attract human capital more easily and companies headquarters are located there.

Table 10 shows the mean ranks and scores of the cities in the sample for the SCCIs according to their size in terms of the number of inhabitants in 2015. Cities of more than one million inhabitants are classified as XXL, 500,000 to 1,000,000 inhabitants denote XL cities, 250,000 to 500,000 inhabitants are L cities, and 50,000 to 250,000 inhabitants are M cities. This classification follows Montalto et al. (2018a, pp. 21–23).

Table 10: Mean Ranks and Scores According to Population Sizes

	Objective SCCI			
City Size Class	XXL	XL	L	M
N	17	21	16	11
Capitals	14	8	3	3
Non–Capitals	3	13	13	8
% Capitals	82.35	38.10	18.75	27.27
Ø City Size	2,827,967	672,563	342,616	200,553
Ø Rank	25	34.62	34.75	37.3
Ø Score	0.14	–0.04	–0.03	–0.12
	Subjective SCCI			
City Size Class	XXL	XL	L	M
N	17	21	16	11
Capitals	14	8	3	3
Non–Capitals	3	13	13	8
% Capitals	82.35	38.10	18.75	27.27
Ø City Size	2,827,967	672,563	342,616	200,553
Ø Rank	39.24	34.76	28	27.36
Ø Score	–0.25	–0.09	0.21	0.25

Overall, Paris with 9,782,671 inhabitants is the largest city in the sample for the SCCIs and Luxembourg with 111,387 inhabitants is the smallest city (see Table A23 for the concrete classifications of the cities to XXL, XL, L, and M). The results for the mean ranks and scores are similar to the results for the capital and non–capital cities and, again, more or less vice versa. Generally speaking, a city is objectively smarter when it is larger, but people perceive it as smarter when it is smaller. Concerning their geographical position, the cities are quite well distributed across those four size classes so that there should not be a large bias according to their location. Recall again that this is particularly surprising when considering the correlation coefficient between the objective and subjective SCCI (0.51).

More detailed information about the differences in the outcomes of the SCCIs gives another PCC matrix (see Table 11). The terms 'Objective' and 'Subjective' describe the scores of the two SCCIs. Furthermore, the city size in terms of inhabitants is further differentiated in an untreated and a treated version. The untreated version takes the original values for the city sizes as shown in Table A23. The treated version winsorizes outliers as described in Subsection 3.5.

Table 11: PCCs Between the SCCIs and Population Sizes

	Objective	Subjective	Size (Untreat.)	Size (Treat.)
Objective	1			
Subjective	0.51	1		
Size (Untreat.)	0.22	–0.17	1	
Size (Treat.)	0.19	–0.34	0.74	1

The PCCs confirm what is displayed by the mean values for the city ranks and scores. There exist small positive correlation coefficients between the objective SCCI scores and the untreated as well as treated city sizes (0.22 and 0.19) Furthermore, there exist small negative correlation coefficients between the subjective SCCI scores and the untreated as well as treated city sizes (–0.17 and –0.34). Due to the treatment of the six largest cities within the sample, the correlation coefficient between the values of the untreated and the treated city sizes is not that close to one as could have been expected (0.74).

Reasons for the differences in the outcome of the SCCIs according to the city sizes are possibly similar to those mentioned for the differences between capital and non–capital cities. However, further investigations are required to provide more compelling and backed up arguments.

5 Identification of Smart City Drivers

"We live in an era of cheap data but expensive information."

–

Weisberg (2005, p. 211)

This section provides an econometric contemplation to analyze what factors drive a city to be smart. First, some theoretical considerations are made. Second, the independent variables of the models are precisely specified. Lastly, the results of the regression analysis are presented and thereafter discussed.

5.1 Theoretical Considerations

As in other contributions to the literature which study issues closely related to the smart cities notion here (see Subsection 2.3), a multiple linear regression model is used to identify what makes a city smart. A linear regression model is a good starting point and also "the single most useful tool in the econometricians kit" (W. H. Greene, 2012, p. 52). Moreover, a standard multiple linear regression fits the needs of this research project because there are no specific issues for panel data to worry about and appropriate handling of potential clusters is not achievable here.[56]

The multiple linear regression model relies on six assumptions which are presented briefly and in accordance with W. H. Greene (2012, p. 56):

A1. Linearity: There is a linear relationship between the dependent variable and the independent variables.

[56]In principle, it could be the case that there are clustered data. Especially the errors of cities from the same country in the subjective SCCI could be associated because such clusters often occur in survey data. However, clusters can not be analyzed conveniently by a standard cluster regression procedure in this case due to sample size and potential clusters (see Bell, Morgan, Kromrey, and Ferron (2010), McNeish and Harring (2017), and Sarstedt and Mooi (2014)). Therefore, Subsection 5.2 takes this issue into account by introducing suitable independent variables.

A2. Full Rank: There is no exact linear relationship between any of the independent variables in the model.

A3. Exogeneity of the Independent Variables: The expected value of the error term at each observation in the sample is not a function of the independent variables observed at any observation.

A4. Homoscedasticity and Nonautocorrelation: Each disturbance has the same finite variance and is uncorrelated with every other disturbance.

A5. Exogenously Generated Data: The process that generates the sample data is independent of the process that generates the error term.

A6. Normal Distribution: The error term is normally distributed.

Another assumption is added which is not necessarily a property of the standard multiple linear regression model, but its fulfillment can contribute to avoid severely statistical problems (W. H. Greene, 2012, pp. 129–130):

A7. Non–Multicollinearity: The independent variables are not too highly correlated.

The last assumption takes into account that outlying values could have a large influence on the regression results (Casson & Farmer, 2014, p. 595; Weisberg, 2005, p. 194):

A8. No Outliers: No dependent or independent variable has outlying values.

For meaningful results, the multiple regression model ideally fulfills these eight assumptions. It is then possible to use OLS to find the Best Linear Unbiased Estimator (BLUE) because the BLUE satisfies, in any case, the first six assumptions and thus, the Gauss–Markov–Theorem (Weisberg, 2005, p. 27).

A guideline to verify the assumptions of the multiple linear regression comes from Chen, Ender, Mitchell, and Wells (2003). Assumptions A3 and A5 are not explicitly tested because they are mainly an issue for time series or clusters. A closer look at assumption A1 can be done graphically (Casson & Farmer, 2014, p. 592; Osborne & Waters, 2002, pp. 1–2). Assumption A2 is checked automatically by Stata 13 which is the statistical program used here. Statistical tests are implemented for assumptions A4, A6, and A7. The Breusch–Pagan / Cook–Weisberg test (see Breusch and Pagan (1979), and Cook and Weisberg (1983)) is employed for assumption A4 in line with a scatter plot, the Shapiro–Wilk test is used for assumption A6 (see Subsection 3.5), and the Variance Inflation Factor (VIF) checks assumption A7 (see O'Brien (2007)) while a rule of 10 for the VIF is applied because that rule is most common (O'Brien, 2007, pp. 673–674).[57] Outliers (assumption A8) are detected with the help of the Shapiro–Wilk test in line with a graphical assessment (see Subsection 3.5). The assessment of values being outliers is done with the Shapiro–Wilk test for normality because normality of the variables has advantageous features for the regression analysis (Casson & Farmer, 2014, pp. 594–595).

Considering the SCCIs and potential independent variables, it is likely that outliers could be an issue. An option is to employ a logarithmic transformation when the data is skewed (Benoit, 2011, p. 2). Regression results can then be interpreted as elasticities (Benoit, 2011, p. 4). However, it would be necessary to add a constant because some values for the SCCIs are negative. Even though adding a constant changes solely the mean, but not the variance, skewness, or kurtosis, it influences the subsequent transformation (Osborne, 2002, p. 3). Therefore, Osborne (2002, p. 3) suggests that the minimum values of each variable should be moved to one in case that a logarithmic transformation is applied.

In the case that outliers are present in the variables, another possibility is to run a robust regression with an M–estimator in addition to the OLS regression. Verardi and Croux (2009, pp. 441–442) describe a robust regression with an M–estimator which is based on Huber (1964) while observations with a Cook's distance (see Cook (1977)) above one receive a zero weight.[58]

[57] Note that O'Brien (2007) discusses the appropriateness of such a rule of thumb. However, no further investigation is undertaken here as long as there is clearly no violation against the rule of 10.

[58] Verardi and Croux (2009) propose a more robust estimator in their article. However, their empirical investigation is sparse, and pitfalls of this method are hard to ascertain. Therefore, the described M–estimator is used in the presence of outliers because it is nearly as efficient as OLS and also the most common robust estimator (Alma, 2011, p. 413).

Another aspect to bear in mind is that the number of independent variables should not be too large, considering the small sample size for the SCCIs (Brooks & Barcikowski, 2012, p. 2). However, there is no unambiguous rule, and similar studies with a similar amount of observations include a lot of independent variables (see Neirotti et al. (2014)).

Moreover, despite the logarithmic model, all continuous variables are normalized with the help of z–scores before the regression is run. When using z–scores, the beta coefficients show which independent variables are most important in the explanation of the dependent variable. An one standard deviation increase in an independent variable contributes to an expected increase or decrease in the dependent variable by the size of the beta coefficient (Nathans, Oswald, & Nimon, 2012, pp. 2–3).[59] This approach is especially helpful when the measurement units are not readily interpretable (Hoyt, Leierer, & Millington, 2006, p. 227). Dummy variables are not normalized. Even though the normalization would also work in principle, the interpretation of dummy variables is more expedient when assessed against the difference between its counterpart. Furthermore, the dummy variables solely influence the intercept and not the slope of the regression line (Hoyt et al., 2006, p. 229).

When building a regression model, the problem can occur that there are unnecessary variables in the model (overfitting), or that important variables are not in the model (underfitting) (Chatterjee & Simonoff, 2013, pp. 23–24). Common data–driven techniques to overcome these problems are controversial (see Heinze and Dunkler (2017), and Ratner (2010)). Therefore and because the aim is to specify the same models which can explain both SCCIs as dependent variables, it is most appropriate that the selection of variables is theory–driven.

[59] There are disadvantages of the normalization procedure when the independent variables are correlated. However, other approaches do come along with other disadvantages. See Nathans et al. (2012) for a detailed discussion.

5.2 Specification of the Independent Variables

The following specification of the econometric model to further investigate the SCCIs takes into account similar contributions to the literature. A challenge is thereby that it is ambitious to find a model which can fit for both SCCIs as dependent variables. This approach implies the drawback that it is hardly possible to find a model which fits the objective and subjective SCCI to the same extent. However, the approach can offer some interesting insights which other approaches would not be capable of. With this approach, it can be shown if the same variables drive the smartness of a city objectively as well as subjectively and if so, to which amount they do. The dependent variable in the econometric model is the objective SCCI and the subjective SCCI, respectively. Caragliu and Del Bo (2015) employ an akin approach as they calculate a regional smartness indicator which they then use as the dependent variable in their models.

The independent variables are:

- Population:
  'Population' depicts the number of inhabitants in each of the 65 European cities under investigation here. It takes into account that cities have different characteristics with respect to their size. The provision of infrastructure and services differ greatly among various city sizes. 'Population' is also used by Neirotti et al. (2014, p. 33), Oueslati et al. (2015, p. 1607), and Węziak–Białowolska (2016, p. 93). Moreover, Subsection 4.2 shows that larger cities tend to perform better in the objective SCCI and worse in the subjective SCCI. Therefore, it is of interest to investigate the influence of population size more closely.

- GDP per Capita in PPS:
  'GDP per Capita in PPS' considers the economic strength of each city. As a common and convenient practice, GDP per capita is measured in purchasing power standard (PPS) to avoid issues in price differences across countries. GDP per capita is commonly employed in similar studies (see Caragliu and Del Bo (2015, p. 78), Neirotti et al. (2014, p. 33), Oueslati

et al. (2015, p. 1607), and Węziak–Białowolska (2016, p. 93)). A major drawback in the implementation of 'GDP per Capita in PPS' in the multiple linear regression model is that it is also an indicator of the objective SCCI and weights 1.84 % in it. Similar indicators as 'GDP per Capita in PPS' are part of the objective indicator which excludes the option to use a different high–quality variable for economic strength. Nevertheless, 'GDP per Capita in PPS' needs to be part of the model because it assumable has a significant influence on the smartness of the cities and also it is necessary to control for the economic strength of the cities when considering the effects of the other independent variables.

- Population Density:
  'Population Density counts the number of inhabitants per square kilometer. Dense urban areas facilitate social interactions which contribute to the spread of knowledge (Glaeser & Gottlieb, 2006, pp. 1275–1276). On the other hand, density can lead to diseconomies in areas as real estate, security, transportation, and energy (Neirotti et al., 2014, p. 29). 'Population Density' is also used by other authors that investigate issues on a city level directly (see Neirotti et al. (2014, p. 33)) or via a dummy for urbanization (see Caragliu and Del Bo (2015, p. 78)).

- Cooling Degree Days:
  'Cooling Degree Days' is a proxy for the climate in the contemplated cities. Climate can influence several objective factors of the cities such as buildings or environment and it can be highly relevant for subjective perceptions. Węziak–Białowolska (2016, p. 90) follows Knez (2005) and argues that climate shapes the experience of a place. This leads her to introduce a dummy variable for Southern European cities. Furthermore, Oueslati et al. (2015, p. 1607) employ the average temperature for the warmest months of the year in their model.

- Heating Degree Days:
  'Heating Degree Days' is like 'Cooling Degree Days' a proxy for the climate. The rationale behind its usage is the same as for 'Cooling Degree Days'. Including both of these variables in the regression has

the advantage that a more nuanced view about the climatic situation in the cities can be obtained. However, there could arise problems due to multicollinearity (assumption A7). In the case of multicollinearity according to the VIF, one of these two variables is excluded from the set of independent variables.

- Dummy Capital:
  A dummy variable to distinguish between capital and non–capital cities is not used in similar studies. However, following the descriptive statistics in Subsection 4.2 gives reason to include this dummy variable in the regression model because capital cities tend to be smarter objectively and at the same time less smart subjectively. Furthermore, the sample characteristics make it appropriate to use this dummy variable because the 65 European cities in the sample include all 28 capital cities of the EU.

- Dummy New Member State:
  Another dummy variable contrasts new and old member states of the EU. New member states are those that entered the EU in this century. The idea is that there could be differences in the smartness of new and old member states due to historical developments, even if there is a control for economic strength and climate. Since all cities in the sample are in the EU, this differentiation is possible. It goes along with Caragliu and Del Bo (2015, p. 73).

Table A24 shows descriptive statistics for the continuous independent variables. The independent variables are all from 2015.[60] The source is always Eurostat. Despite 'Cooling Degree Days' and 'Heating Degree Days', the independent variables are all on the city level. 'Cooling Degree Days' and 'Heating Degree Days' are measured on a regional level which for these specific variables is a plausible approximation.

[60] Some values for the variables 'Population' and 'Population Density' are from previous years.

Finally, the equation for the multiple regression is given by:

$$SCCI_i = \beta_0 + \boldsymbol{\beta}\boldsymbol{X}_i + \epsilon_i \tag{14}$$

where $SCCI_{\mathrm{i}}$ stands for the i^{th} city while both, the objective and the subjective SCCI can be used as dependent variables. β_0 is the intercept, $\boldsymbol{\beta}$ are the regression coefficients of the independent variables, $\boldsymbol{X}_i$ are the independent variables to estimate the smartness for the i^{th} city, and ϵ_i is the i^{th} error.

5.3 Results of the Econometric Models

The regression model introduced in this section is run once with the objective SCCI as the dependent variable and another time with the subjective SCCI as the dependent variable. In both cases, most assumptions of the multiple linear regression model are satisfied which generally confirms that OLS is an appropriate estimator. Tables A25, A26, and A27 show the test statistics in respect to assumptions A4, A6, and A7. Table A25 indicates that the assumption of homoscedasticity is violated in the case that the subjective SCCI is the dependent variable. A closer look at this violation with the help of a plot of residuals vs. fitted values for the OLS regression when the subjective SCCI is the dependent variable shows that there is not too much change in variation of the residuals (see Figure A1). Even though it seems that heteroscedasticity is not challenging the OLS regression badly, another regression with robust standard errors is run for both SCCIs. Robust standard errors as used here are associated with White (1980) and a common practice to tackle the issue of heteroscedasticity (Imbens & Kolesar, 2012, p. 1).

Furthermore, the Shapiro–Wilk test (see Table A28) in line with a graphical assessment emphasizes that there are variables which contain outliers.[61] The outliers lead to slightly ambiguous results for linearity (assumption A1). For

[61] This is especially prevalent for the variable 'Population'. Paris (9,782,671 inhabitants), and London (8,730,803 inhabitants) are far more populated than the other cities in the sample. Berlin is third with 3,469,849 inhabitants.

the issue of outliers which also slightly influences the linearity assumption, a logarithmic regression and a robust regression with an M–estimator as described in Subsection 5.1 is conducted.[62] This can also improve the issue of heteroscedasticity (Benoit, 2011, p. 5; Wilcox & Keselman, 2004, p. 350).

The coefficients of the multiple regressions are shown in Tables 12 and 13. Model (1) relies on the OLS estimator, model (2) is conducted with robust standard errors, in model (3) a logarithmic transformation is applied to the variables prior to the OLS estimation, and model (4) uses an M–estimator for robust regressions. Note that the reason why the coefficients of model (3) differ substantially is that they are not normalized with the help of z–scores.

The results in Table 12 clearly emphasize that economic strength is the main driver for the objective SCCI. Moreover, population density has a significantly negative impact on the smartness of a city according to all four models while the mere number of inhabitants does not influence the smartness significantly. The results are more ambiguous for the climate variables, but it looks as if colder areas tend to be smarter objectively. Furthermore, three of the four models indicate that a city is objectively speaking smarter when it is capital and less smart when it is a new member state of the EU. Notably, the model fit is quite good, especially when compared to similar studies (see Subsection 2.3). The variables in every model are jointly significant in explaining the objective SCCI and the model fit is quite good.

The results in Table 13 for the subjective SCCI as the dependent variable also point on the importance of the economic state for the smart performance of a city. It is remarkable that population density contributes again negatively to the smartness. Besides, there is some indication that colder areas within the EU contribute in principle to a subjectively perceived smarter performance. If a city is a capital city, its smart performance is significantly worse than the performance of a non–capital city. Population and being a new member state of the EU have no significant impact on the subjective smart performance. The model fit is intermediate, and the independent variables are jointly significant.

[62] The values are not normalized to z–scores in case of the logarithmic regressions because the regression coefficients can then be interpreted as elasticities.

Table 12: Estimations of Smart City Drivers for the Objective SCCI

Independent Variables	Dependent Variable: Objective SCCI			
	(1)	(2)	(3)	(4)
Population	0.099 (0.095)	0.099 (0.089)	0.001 (0.023)	0.096 (0.102)
GDP per Capita in PPS	0.640*** (0.091)	0.640*** (0.093)	0.380*** (0.053)	0.642*** (0.098)
Population Density	–0.256*** (0.094)	–0.256*** (0.093)	–0.056** (0.026)	–0.242** (0.100)
Cooling Degree Days	0.019 (0.131)	0.019 (0.103)	0.003 (0.01)	0.012 (0.140)
Heating Degree Days	0.237* (0.127)	0.237** (0.106)	0.090* (0.051)	0.206 (0.136)
Dummy Capital (1 = Capital)	0.392** (0.194)	0.392** (0.176)	0.051 (0.045)	0.372* (0.207)
Dummy New Member State (1 = New Member State)	–0.459* (0.189)	–0.459** (0.195)	–0.050 (0.045)	–0.463** (0.202)
Constant	–0.006 (0.108)	–0.006 (0.101)	–3.672 (0.534)	–0.004 (0.115)
Obs.	65	65	65	65
R^2	0.7228	0.7228	0.7257	
Adj. R^2	0.6888		0.6920	
$F(7, 57)$	21.23	20.58	21.54	17.99
Prob > F	0.000	0.000	0.000	0.000

Notes: Standard errors in parentheses;
* 0.05 < p–value < 0.1, ** 0.01 < p–value < 0.05, *** p–value < 0.01

Table 13: Estimations of Smart City Drivers for the Subjective SCCI

Independent Variables	Dependent Variable: Subjective SCCI			
	(1)	(2)	(3)	(4)
Population	–0.002 (0.123)	–0.002 (0.097)	–0.048 (0.046)	0.075 (0.116)
GDP per Capita in PPS	0.330*** (0.117)	0.330*** (0.101)	0.338*** (0.107)	0.310*** (0.111)
Population Density	–0.269** (0.094)	–0.269** (0.112)	–0.050 (0.052)	–0.327*** (0.114)
Cooling Degree Days	–0.194 (0.168)	–0.194 (0.165)	–0.061*** (0.020)	–0.343** (0.159)
Heating Degree Days	0.296* (0.163)	0.296** (0.117)	0.120 (0.102)	0.257 (0.155)
Dummy Capital (1 = Capital)	–0.422* (0.250)	–0.422** (0.189)	–0.160* (0.089)	–0.528** (0.236)
Dummy New Member State (1 = New Member State)	–0.166 (0.242)	–0.166 (0.235)	0.116 (0.091)	–0.212 (0.230)
Constant	0.240* (0.138)	0.240* (0.130)	–2.171** (1.071)	0.301** (0.131)
Obs.	65	65	65	65
R^2	0.5420	0.5420	0.5802	
Adj. R^2	0.4857		0.5286	
F(7, 57)	9.64	9.41	11.25	14.68
Prob > F	0.000	0.000	0.000	0.000

Notes: Standard errors in parentheses;
* 0.05 < p–value < 0.1, ** 0.01 < p–value < 0.05, *** p–value < 0.01

Comparing both SCCIs, it is prevalent that they provide similar estimates on 'GDP per Capita in PPS' and on 'Population Density', and to a lesser extent also on the climatic situation. The result for 'GDP per Capita in PPS' is somewhat expectable, but the result for 'Population Density' is remarkable when bearing in mind that the models control for the number of inhabitants. In contrast, capitals can explain an objective smart performance and non–capitals can explain a subjective smart performance. The coefficients of the four models in respect to the two SCCIs are almost always all either negative or positive.[63] This fact demonstrates that the models generally come to similar conclusions and their mutual contemplation raises the reliability. But it is also noteworthy that the significance of the coefficients does sometimes differ.

5.4 Discussion of the Smart City Drivers

Lastly, in this section, the identified smart city drivers are discussed in detail to point on policy implications. These interpretations of the smart city drivers can solely be seen as a first attempt to determine their meaning. Further work on this is required.

The implication of the models that a well–functioning economy is closely related to the objective and subjective smartness of a city is not very much surprising in respect to similar studies, and particularly when taking into consideration that the dimension 'Economy & Governance' is the most important in the SCCIs. The straightforward implication for city officials, policymakers, and important stakeholders to simply improve the state of the economy is probably not so much helpful. However, it is obligatory to illustrate the influence of the economic performance on the smartness performance, and 'GDP per Capita in PPS' also operates as an essential control variable.

Crucial are the results of the econometric models in consideration of the population density. Cities are objectively and subjectively less smart when they

[63] The sole exceptions are the variables 'Population' and the 'Dummy New Member State' when the subjective SCCI is the dependent variable. However, the estimates are all close to zero, and so there is not an inconsistency at hand.

are denser, and that effect is significant. So if a city aims to become smart, it should also try to curtail its density. In light of dense megacities which people would in the first moment usually assign a smartness label to, this finding is not intuitive. But as the smart city notion is defined here, this outcome is plausible. It is way harder to provide the citizens with a well–functioning infrastructure, a sound environment, and welcoming living conditions when there are a lot of people in a limited area. The consequence is not alone that the number of inhabitants should be contained. Population size alone has no significant effect on neither the objective or the subjective smartness. Another way to deal with this issue is to widen the urban areas, and that is actually what is happening within the last decades (Fina & Siedentop, 2008, p. 489; Oueslati et al., 2015, p. 1595). However, there is a trade–off between urban sprawl and population density because urban sprawl comes solely at the sake of other challenges (see European Environment Agency (2006)). For this reason, holistic visions for cities are necessary. Without getting into much detail, Wheeler (2009, pp. 872–872) makes interesting proposals, albeit that he primarily focusses on sustainability. Amongst others, he suggests that balanced local communities can help to tackle mobility challenges in contrast to metropolitan areas with low density and discontiguous communities on the edges (Wheeler, 2009, p. 866).

Furthermore, there is an indication that cities in colder areas within the EU are smarter than cities in warmer areas. The interpretation of this finding is open to many gateways, but an attempt shall be made. Glaeser (2005, pp. 2–4) argues for the USA that warmer cities are growing faster and that aside from climate, education possibly predicts urban growth best while this is especially true for cold–weather cities. Five indicators from the objective SCCI and two indicators from the subjective SCCI which can be quite closely linked to education are all positively associated with the variable 'Heating Degree Days' and negatively with 'Cooling Degree Days' (see Table A29) and thus, support the argument from Glaeser (2005) that colder areas face a need to offer their populations education.[64] On the downside, also in Europe, warmer cities

[64]Even though colder cities perform better in many indicators, it is remarkable that they perform better in *every* indicator related to education.

grow at least slightly faster[65] and this growth leads to many challenges for infrastructure, mobility, cohabitation, and sustainability which colder cities do not face in this markedness. Due to this, it can be argued that it is currently more difficult to make a city smart in a pleasant climate.

Another insight is that the size of the city in terms of inhabitants does not influence the smartness, but it is relevant if the city is a capital or not. There are several reasons why a capital city could be smarter objectively such as agglomeration advantages, (see Subsection 4.2). Subjectively, diverse negative externalities, in contrast, could contribute to a decline in the perceived smartness (see also Subsection 4.2). Both issues can be tackled. Agglomeration advantages could be restricted by means of suitable policies (e.g., the best universities could be located to a place other than the capital), and capitals could try to internalize at least some of the negative externalities they face. Note again that there could be a bias in case that there are country–specific effects because, for some countries, there is solely the capital city part of the sample.

Some evidence exists that a reason for worse performance in the objective SCCI relates to a city being part of a new member state of the EU, but there is no significant effect for the subjective SCCI. The worse objective performance could be attributed to aspects which are broadly related to the economic performance in case that these aspects are not fully accounted for by the inclusion of 'GDP per Capita in PPS' in the models. Moreover, historical reasons and different development paths with respect to the Cold War could still affect every dimension of the objective SCCI negatively.

Overall, the econometric analysis offers some intriguing results. However, this can solely be seen as a first attempt, and more detailed analyses need to be carried out in the future using the SCCIs. An idea is to apply the econometric models on every dimension, following Neirotti et al. (2014). Furthermore, more

[65] The population growth from 2005 to 2015 correlates positively with the variable 'Cooling Degree Days' (0.0549) and negatively with the variable 'Heating Degree Days' (–0.1391). Data to calculate population growth are generally on a metropolitan level because, for the year 2005, they are by far not complete on a city level. However, the metropolitan approximations are appropriate because cities provide a lot of services for their surroundings. Due to data availability on the metropolitan level, city data are used for Leipzig and Rostock, and NUTS 3 data from 2007 instead of 2005 are used for Aalborg and Copenhagen.

evaluation in respect to the choice of the best set of independent variables is a step to take, and logit, as well as probit models, are potential supplements, following Caragliu and Del Bo (2015), and Węziak–Białowolska (2016).

6 Conclusion

Holistic smart city concepts in research are still at their beginning. Therefore, the primary goal of this master thesis is to propose a persuasive conception of a smart city framework to capture the state of 65 European cities concerning their smartness. This is vis–à–vis growing cities a vital research topic. Such a framework which relies on the construction of two Smart City Composite Indicators (SCCIs) as introduced here helps to point on strengths and weaknesses of cities and can be seen as an essential benchmark for city planning and policy actions. Furthermore, a distinguished objective and subjective contemplation of the smart city performances is novel and yields notable findings. This unique approach is of interest when considering that many other researchers merge objective and subjective indicators whereas they strongly rely on objective indicators simply because there are more indicators available to do so. However, the researchers do generally not discuss that this could be an issue and give no reason why it could be appropriate to include a lot more objective indicators. Another important contribution of this thesis is the establishment of econometric models for the identification of smart city drivers. These models are a first attempt to explain the SCCIs, entail implications for the city smart performances and their long–term developments plans, and are open to a broader application on the SCCIs.

The literature review shows that the construction of composite indicators faces several challenges and that it is a specific process which depends on the research topic, sample size, and properties of available indicators. The SCCIs proposed in this work tackle those challenges conveniently. After defining smart cities holistically concerning four dimensions ('Infrastructure & Mobility', 'Living & Social Cohesion', 'Economy & Governance', 'Environment & Sustainability'), a sample of 65 European cities takes into account 73 objective indicators, as well as nineteen subjective indicators due to various criteria. Missing data are imputed by predictive mean matching (PMM) as a tool for multiple imputations (MIs) unless approximations by a previous year or a higher regional level are feasible. As a next step, outliers are detected by the Shapiro–Wilk test in line with a graphical assessment, and then winsorized while the IQR detects outlying values.

The obligatory normalization procedure utilizes z–scores. The last two steps in the construction phase are weighting and aggregation. Weights for the indicators are obtained by a Principal Component Factor Analysis (PCFA) and weights for the dimensions are obtained by an Analytic Hierarchy Process (AHP). The simple additive aggregation implies a full trade–off among indicator values and is therefore employed to finally calculate the two SCCIs. Furthermore, uncertainty and sensitivity analyses show that the results for the SCCIs are not much sensitive to model choices.

Taking a closer look at the results of the SCCIs displays that objective and subjective SCCI are positively correlated with medium strength (0.51). Dimension DO has solely a weak correlation with the objective SCCI (0.24) while the dimensions of the subjective SCCI are all highly correlated with the subjective SCCI. Dimensions AO and DO are negatively correlated (–0.29), suggesting objectively a slight trade–off between the two dimensions 'Infrastructure & Mobility' and 'Environment & Sustainability'. In contrast, dimensions BO and CO go in hand (0.73). The dimensions of the subjective SCCI are all highly correlated with each other (all above 0.8). Between the dimensions of the objective and the subjective SCCI, it is remarkable that dimension AO is very weakly correlated with each dimension of the subjective SCCI (all below $|0.1|$).

Furthermore, capital and non–capital cities, as well as the population size of the cities, are differentiated and delineated concerning the two complete SCCIs. The descriptive contemplation shows that capital and large cities perform better objectively and that non–capital and small cities perform better subjectively. Capitals may perform better objectively because they have agglomeration advantages and get a disproportionate amount of resources. Besides, capital cities in the sample are substantially larger than non–capital cities. Large cities can more easily provide a skilled mix of employees and reach critical masses for new digital solutions. However, residents of capitals distrust their local administration and negative externalities due to diseconomies of scale could influence the subjective performance.

This work also tries to identify variables which drive the cities' smartness. Therefore, multiple linear regression models are proposed in which the objective and the subjective SCCI, respectively, act as dependent variables. The idea is that the same models explain both composite indicators to offer intriguing interpretations. A first model employs an OLS estimator. Due to issues with the OLS assumptions, three more models are submitted. A second model is done with robust standard errors, a third model transforms the dependent and independent variables logarithmically, and a fourth model uses an M–estimator to conduct a robust regression. Despite the case of the logarithmic regression where the coefficients represent elasticities, all other models are normalized with z–scores so that the results indicate the contribution of the variables to the explanation of the model. Independent variables in the models are 'Population', 'GDP per Capita in PPS', 'Population Density', 'Cooling Degree Days', 'Heating Degree Days', 'Dummy Capital', and 'Dummy New Member State'.

The results emphasize that economic strength mainly drives the objective and subjective city smartness. This is not so much surprising and the straightforward policy implications for enhancing the economic state of the city is probably already in the minds of city officials, policymakers, and important stakeholders. Moreover, 'Population Density' is significantly and positively associated with both SCCIs while 'Population' is in no model significant. This result emphasizes the need for city visions. Population density can be reduced by restriction of the population or by urban sprawl. However, more essential than this simple deduction is that cities try to become balanced. Instead of dense city centers, smaller centers all around the cities could help to distribute the population in the city in a more preferable way. There is also a slight indication that colder cities are smarter according to both SCCIs. One reason may be that colder regions put a focus on education as they need to make more efforts to keep their population. Another reason may be that continuing population growth makes it a lot harder for warmer cities to become smart. Capital cities are significantly smarter when the objective SCCI is the dependent variable and non–capitals are significantly smarter when the subjective SCCI is the dependent variable. This could be the case on the one hand due to the mentioned prevalence of agglomeration advantages and on the other hand due to diseconomies of scale. New member states of the EU are less smart objectively while there are no distinct results subjectively. The worse objective performance even when taking into account that the models control for 'GDP per Capita in PPS' could still reflect the different historical development paths of Eastern European cities.

The empirical evidence of this thesis suffers from some limitations. For the objective SCCI, the main issue is that data availability is sometimes not optimal. Many suitable indicators are from a different year than 2015, from a higher regional level than the city level or not at hand at all. The subjective SCCI relies on the European Urban Audit survey from 2015 which is a convenient source but sticking to that source limits the sample size heavily. Besides, more indicators could improve the quality of the subjective SCCI. Another issue is that subjective indicators possibly measure a broader picture and not precisely what the respective question asks for. This is always a concern with respect to survey data and not exclusively a challenge here. Furthermore, it would be an improvement if the indicators of both SCCIs better fit to each other. The econometric analysis of the smartness composite indicators does probably not select the best set of independent variables as the selection is theory–driven and because convenient similar contributions to the literature are still rare.

This work entails a lot of viable research projects. The SCCIs can be used to contemplate various aspects other than population size or capital status as it is done here. Furthermore, the dimensions of the SCCIs can be investigated closer. The construction of the SCCIs in time intervals is especially interesting because it is then possible to monitor the smartness development of the cities over time. Further research questions imply the *if* and the *how* of an appropriate merge of objective and subjective smartness indicators. Undoubtedly, future research must also try to tackle the limitations of this thesis, and interpretation of the SCCIs results and the regression results need to be enhanced.

References

Adefisoye, J. O., Golam Kibria B.M., & George, F. (2016). Performances of Several Univariate Tests of Normality: An Empirical Study. *Journal of Biometrics & Biostatistics*, *7*(4), 1–8. doi:10.4172/2155-6180.1000322

Ahad, N. A., Yin, T. S., Othman, A. R., & Yaacob, C. R. (2011). Sensitivity of Normality Tests to Non–Normal Data. *Sains Malaysiana*, *40*(6), 637–641.

Ahvenniemi, H., Huovila, A., Pinto–Seppä, I., & Airaksinen, M. (2017). What Are the Differences Between Sustainable and Smart Cities? *Cities*, *60*, 234–245. doi:10.1016/j.cities.2016.09.009

Aiginger, K., & Firgo, M. (2015). Regional Competitiveness Under New Perspectives. WWWforEurope Policy Paper No. 26, Wien, Austria. Retrieved from https://core.ac.uk/download/pdf/46117848.pdf

Alawadhi, S., Aldama–Nalda, A., Chourabi, H., Gil–Garcia, J. R., Leung, S., Mellouli, S., … Walker, S. (2012). Building Understanding of Smart City Initiatives. In H. J. Scholl, M. Janssen, M. A. Wimmer, C. E. Moe, & L. S. Flak (Eds.), *Electronic Government* (Vol. 7443, pp. 40–53). doi:10.1007/978-3-642-33489-4_4

Albino, V., Berardi, U., & Dangelico, R. M. (2015). Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology*, *22*(1), 3–21. doi:10.1080/10630732.2014.942092

Alma, Ö. (2011). Comparison of Robust Regression Methods in Linear Regression. *International Journal of Contemporary Mathematical Sciences*, *6*(9), 409–421.

Alonso, J. A., & Lamata, M. T. (2006). Consistency in the Analytical Hierarchy Process: A New Approach. *International Journal of Uncertainty, Fuzziness and Knowledge–Based Systems*, *14*(4), 445–459. doi:10.1142/S0218488506004114

Althouse, L. A., Ware, W. B., & Ferron, J. M. (1998). Detecting Departures from Normality: A Monte Carlo Simulation of a New Omnibus Test Based on Moments. San Diego, CA. Retrieved from https://files.eric.ed.gov/fulltext/ED422385.pdf

Angelidou, M. (2014). Smart City Policies: A Spatial Approach. *Cities*, *41*, 3–11. doi:10.1016/j.cities.2014.06.007

Angelidou, M. (2016). Four European Smart City Strategies. *International Journal of Social Science Studies*, *4*(4), 18–30. doi:10.11114/ijsss.v4i4.1364

Annoni, P., & Dijkstra, L. (2013). *EU Regional Competitiveness Index: RCI 2013*. Luxembourg: Publications Office of the European Union.

Annoni, P., Dijkstra, L., & Gargano, N. (2017). *The EU Regional Competitiveness Index 2016*. Luxembourg: Publications Office of the European Union.

Annoni, P., & Kozovska, K. (2010). *EU Regional Competitiveness Index: RCI 2010*. Luxembourg: Publications Office of the European Union.

Athanasoglou, S., & Dijkstra, L. (2014). *The Europe 2020 Regional Index*. Luxembourg: Publications Office of the European Union.

Bandura, R. (2008). A Survey of Composite Indices Measuring Country Performance: 2008 Update. UNDP/ODS Working Paper, New York, NY. Retrieved from http://www.undp.org/content/dam/undp/library/corporate/Development%20Studies/indices_2008_bandura.pdf

Baraldi, A. N., & Enders, C. K. (2010). An Introduction to Modern Missing Data Analyses. *Journal of School Psychology*, *48*(1), 5–37. doi:10.1016/j.jsp.2009.10.001

Bell, B. A., Morgan, G. B., Kromrey, J. D., & Ferron, J. M. (2010). The Impact of Small Cluster Size on Multilevel Models: A Monte Carlo Examination of Two–Level Models with Binary and Continuous Predictors. *JSM Proceedings, Section on Survey Research Methods*, 4057–4067.

Benoit, K. (2011). Linear Regression Models with Logarithmic Transformations. London, England. Retrieved from https://pdfs.semanticscholar.org/169c/c9bbbd77cb7cec23481f6ecb2ce071e4e94e.pdf

Bifulco, F., Tregua, M., Amitrano, C. C., & D'Auria, A. (2016). ICT and Sustainability in Smart Cities Management. *International Journal of Public Sector Management*, *29*(2), 132–147. doi:10.1108/IJPSM-07-2015-0132

Bosch, P., Jongeneel, S., Rovers, V., Neumann, H.-M., Airaksinen, M., & Huovila, A. (2017). *CITYkeys Indicators for Smart City Projects and Smart Cities*. doi:10.13140/RG.2.2.17148.23686

Box, G. E. P., Hunter, J. S., & Hunter, W. G. (2005). *Statistics for Experimenters: Design, Innovation, and Discovery* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Breusch, T. S., & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, *47*(5), 1287–1294.

Brooks, G. P., & Barcikowski, R. S. (2012). The PEAR Method for Sample Sizes in Multiple Linear Regression. *Multiple Linear Regression Viewpoints*, *38*(2), 1–16.

Calero, C., Moraga, M. Á., & Piattini, M. (2008). *Handbook of Research on Web Information Systems Quality*. doi:10.4018/978-1-59904-847-5

Caragliu, A., & Del Bo, C. F. (2015). Smart Specialization Strategies and Smart Cities: An Evidence–Based Assessment of European Union Policies. In K. Kourtit, P. Nijkamp, & R. Stough (Eds.), *The Rise of the City* (pp. 55–84). Cheltenham, England: Edward Elgar Publishing.

Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart Cities in Europe. *Journal of Urban Technology*, *18*(2), 65–82. doi:10.1080/10630732.2011.601117

Carlin, J. B., Li, N., Greenwood, P., & Coffey, C. (2003). Tools for Analyzing Multiple Imputed Datasets. *The Stata Journal*, *3*(3), 1–20.

Casson, R. J., & Farmer, L. D. M. (2014). Understanding and Checking the Assumptions of Linear Regression: A Primer for Medical Researchers. *Clinical & Experimental Ophthalmology*, *42*(6), 590–596. doi:10.1111/ceo.12358

Castelnovo, W., Misuraca, G., & Savoldelli, A. (2016). Smart Cities Governance. *Social Science Computer Review*, *34*(6), 724–739. doi:10.1177/0894439315611103

Chatterjee, S., & Simonoff, J. S. (2013). *Handbook of Regression Analysis*. Hoboken, NJ: John Wiley & Sons.

Chen, X., Ender, P., Mitchell, M., & Wells, C. (2003). Regression with Stata. Retrieved from https://stats.idre.ucla.edu/stata/webbooks/reg/chapter2/stata-webbooksregressionwith-statachapter-2-regression-diagnostics/

Cherchye, L., Knox Lovell, C. A., Moesen, W., & van Puyenbroeck, T. (2007). One Market, One Number? A Composite Indicator Assessment of EU Internal Market Dynamics. *European Economic Review*, *51*(3), 749–779. doi:10.1016/j.euroecorev.2006.03.011

Cherchye, L., Moesen, W., Rogge, N., & van Puyenbroeck, T. (2007). An Introduction to 'Benefit of the Doubt' Composite Indicators. *Social Indicators Research*, *82*(1), 111–145. doi:10.1007/s11205-006-9029-7

Cook, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics*, *19*(1), 15–18.

Cook, R. D., & Weisberg, S. (1983). Diagnostics for Heteroscedasticity in Regression. *Biometrika*, *70*(1), 1–10.

Cousineau, D., & Chartier, S. (2010). Outliers Detection and Treatment: A Review. *International Journal of Pschological Research*, *3*(1), 58–67.

Dameri, R. P., & Cocchia, A. (2013). Smart City and Digital City: Twenty Years of Terminology Evolution. Genova, Italy. Retrieved from http://www.itais.org/proceedings/itais2013/pdf/119.pdf

de Vries, M. S., & Sobis, I. (2018). Trust in the Local Administration: A Comparative Study Between Capitals and Non–Capital Cities in Europe. *NISPAcee Journal of Public Administration and Policy*, *11*(1), 209–228. doi:10.2478/nispa-2018-0009

Department for Communities and Local Government. (2009). *Multi–Criteria Analysis: A Manual*. Wetherby, England: Communities and Local Government Publications.

Doornik, J. A., & Hansen, H. (1994). An Omnibus Test for Univariate and Multivariate Normality. Oxford, England. Retrieved from http://www.doornik.com/research/normal2.pdf

Dutta, S., Escalona Reynoso, R., Guadagno, F., Lanvin, B., Rivera Leon, L., Saxena, K., & Wunsch–Vincent, S. (2018). *Global Innovation Index 2018: Energizing the World with Innovation*. Geneva, Switzerland: World Intellectual Property Organization.

Elvery, J. A. (2010). City Size and Skill Intensity. *Regional Science and Urban Economics*, *40*(6), 367–379. doi:10.1016/j.regsciurbeco.2010.05.006

Enders, C. K. (2010). *Applied Missing Data Analysis*. New York, NY: Guilford Press.

European Commission. (2010). Europe 2020: A European Strategy for Smart, Sustainable and Inclusive Growth. Brussels, Belgium. Retrieved from http://ec.europa.eu/eu2020/pdf/COMPLET%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf

European Commission. (2016a). 10–Step Guide for the Construction of a Composite Indicator. Ispra, Italy. Retrieved from https://ec.europa.eu/jrc/en/printpdf/68176

European Commission. (2016b). *Quality of Life in European Cities 2015: Flash Eurobarometer 419*. Luxembourg: Publications Office of the European Union.

European Environment Agency. (2006). *Urban Sprawl in Europe: The Ignored Challenge*. Luxembourg: Office for Official Publications of the European Communities.

Eurostat. (2018). Territorial Typologies for European Cities and Metropolitan Regions. Retrieved September 10, 2018, from https://ec.europa.eu/eurostat/statistics-explained/index.php/Territorial_typologies_for_European_cities_and_metropolitan_regions#Functional_urban_areas:_a_city_and_its_commuting_zone

Farrell, P. J., Salibian–Barrera, M., & Naczk, K. (2007). On Tests for Multivariate Normality and Associated Simulation Studies. *Journal of Statistical Computation and Simulation*, *77*(12), 1065–1080. doi:10.1080/10629360600878449

Feroze, S. M., & Chauhan, A. K. (2010). Performance of Dairy Self Help Groups (SHGs) in India: Principal Component Analysis (PCA) Approach. *Indian Journal of Agricultural Economics*, *65*(2), 308–320.

Fina, S., & Siedentop, S. (2008). Urban Sprawl in Europe – Identifying the Challenge. In M. Schrenk, V. V. Popvich, & P. Elisei (Eds.), *REAL CORP 008: Mobility Nodes as Innovation Hubs*. Schwechat–Rannersdorf, Austria: Competence Center of Urban and Regional Planning.

Gan, X., Fernandez, I. C., Guo, J., Wilson, M., Zhao, Y., Zhou, B., & Wu, J. (2017). When to Use What: Methods for Weighting and Aggregating Sustainability Indicators. *Ecological Indicators*, *81*, 491–502. doi:10.1016/j.ecolind.2017.05.068

Gassmann, O., Böhm, J., & Palmié, M. (2018). *Smart City: Innovationen für die Vernetzte Stadt – Geschäftsmodelle und Management [Smart City: Innovations for the Integrated City – Business Models and Management]*. Munich, Germany: Hanser.

Ghasemi, A., & Zahediasl, S. (2012). Normality Tests for Statistical Analysis: A Guide for Non–Statisticians. *International Journal of Endocrinology and Metabolism*, *10*(2), 486–489. doi:10.5812/ijem.3505

Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler–Milanovic, N., & Meijers, E. (2007). Smart Cities: Ranking of European Medium–Sized Cities. Vienna, Austria. Retrieved from http://www.smart-cities.eu/download/smart_cities_final_report.pdf

Glaeser, E. L. (2005). Smart Growth: Education, Skilled Workers, & the Future of Cold–Weather Cities. Policy Briefs No. 2005–1, Cambridge, MA. Retrieved from https://www.hks.harvard.edu/sites/default/files/centers/rappaport/files/skilledcities.pdf

Glaeser, E. L., & Gottlieb, J. D. (2006). Urban Resurgence and the Consumer City. *Urban Studies*, *43*(8), 1275–1299. doi:10.1080/00420980600775683

Glorfeld, L. W. (1995). An Improvement on Horn's Parallel Analysis Methodology for Selecting the Correct Number of Factors to Retain. *Educational and Psychological Measurement*, *55*(3), 377–393.

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2018). On the Methodological Framework of Composite Indices: A Review of the Issues of Weighting, Aggregation, and Robustness. *Social Indicators Research*, *43*(3), 1–34. doi:10.1007/s11205-017-1832-9

Greene, R., Devillers, R., Luther, J. E., & Eddy, B. G. (2011). GIS–Based Multiple–Criteria Decision Analysis. *Geography Compass*, *5*(6), 412–432. doi:10.1111/j.1749-8198.2011.00431.x

Greene, W. H. (2012). *Econometric Analysis* (7th ed.). Harlow, England: Pearson Education.

Grossi, G., & Pianezzi, D. (2017). Smart Cities: Utopia or Neoliberal Ideology? *Cities*, (69), 79–85. doi:10.1016/j.cities.2017.07.012

Hagedoorn, J., & Cloodt, M. (2003). Measuring Innovative Performance: Is There an Advantage in Using Multiple Indicators? *Research Policy*, *32*(8), 1365–1379. doi:10.1016/S0048-7333(02)00137-3

Hain, J. (2010). Comparison of Common Tests for Normality. Würzburg, Germany. Retrieved from http://www.statistik-mathematik.uni-wuerzburg.de/fileadmin/10040800/user_upload/hain/da_hain_final.pdf

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Waltham, MA: Elsevier / Morgan Kaufmann Publishers.

Harzing, A.-W., Baldueza, J., Barner–Rasmussen, W., Barzantny, C., Canabal, A., Davila, A., ... Zander, L. (2009). Rating Versus Ranking: What is the Best Way to Reduce Response and Language Bias in Cross–National Research? *International Business Review*, *18*(4), 417–432. doi:10.1016/j.ibusrev.2009.03.001

Heinze, G., & Dunkler, D. (2017). Five Myths About Variable Selection. *Transplant International*, *30*(1), 6–10. doi:10.1111/tri.12895

Henderson, A. R. (2006). Testing Experimental Data for Univariate Normality. *Clinica Chimica Acta*, *366*(1–2), 112–129. doi:10.1016/j.cca.2005.11.007

Henderson, J. V. (2010). Cities and Development. *Journal of Regional Science*, *50*(1), 515–540. doi:10.1111/j.1467-9787.2009.00636.x

Hermans, E., van den Bossche, F., & Wets, G. (2008). Combining Road Safety Information in a Performance Index. *Accident Analysis and Prevention*, *40*(4), 1337–1344. doi:10.1016/j.aap.2008.02.004

Hermans, E., van den Bossche, F., & Wets, G. (2009). Uncertainty Assessment of the Road Safety Index. *Reliability Engineering & System Safety*, *94*(7), 1220–1228. doi:10.1016/j.ress.2008.09.004

Hollands, R. G. (2008). Will the Real Smart City Please Stand Up? *City*, *12*(3), 303–320. doi:10.1080/13604810802479126

Hoyt, W. T., Leierer, S., & Millington, M. J. (2006). Analysis and Interpretation of Findings Using Multiple Regression Techniques. *Rehabilitation Counseling Bulletin*, *49*(4), 223–233. doi:10.1177/00343552060490040401

Huber, P. J. (1964). Robust Estimation of a Location Parameter. *Annals of Mathematical Statistics*, *35*(1), 73–101.

Hudrliková, L. (2013). Composite Indicators as a Useful Tool for International Comparison: The Europe 2020 Example. *Prague Economic Papers*, *22*(4), 459–473. doi:10.18267/j.pep.462

Huisman, M. (2009). Imputation of Missing Network Data: Some Simple Procedures. *Journal of Social Structure*, *10*(1), 1–29.

Imbens, G. W., & Kolesar, M. (2012). Robust Standard Errors in Small Samples: Some Practical Advice. NBER Working Paper Series No. 18478, Cambridge, MA. Retrieved from https://www.nber.org/papers/w18478.pdf

International Telecommunication Union. (2014). Smart Sustainable Cities: An Analysis of Definitions. Geneva, Switzerland. Retrieved from https://www.itu.int/en/ITU-T/focusgroups/ssc/Documents/Approved_Deliverables/TR-Definitions.docx

IW Consult. (2017). Deutschlands Städte im Vergleich [German Cities in Comparison]. Cologne, Germany. Retrieved November 4, 2018, from https://www.iwconsult.de/leistungen-themen/branchen-und-regionen/staedteranking-2017/

Jacobs, R., & Goddard, M. (2007). How Do Performance Indicators Add Up? An Examination of Composite Indicators in Public Services. *Public Money & Management*, *27*(2), 103–110.

Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Allen Lane.

Kaiser, H. F. (1970). A Second Generation Little Jiffy. *Psychometrika*, *35*(4), 401–415.

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, *34*(1), 111–117.

Kim, H.-Y. (2013). Statistical Notes for Clinical Researchers: Assessing Normal Distribution (2) Using Skewness and Kurtosis. *Restorative Dentistry & Endodontics*, *38*(1), 52–54. doi:10.5395/rde.2013.38.1.52

Kleinke, K. (2017). Multiple Imputation Under Violated Distributional Assumptions: A Systematic Evaluation of the Assumed Robustness of Predictive Mean Matching. *Journal of Educational and Behavioral Statistics*, *42*(4), 371–404. doi:10.3102/1076998616687084

Knez, I. (2005). Attachment and Identity as Related to a Place and its Perceived Climate. *Journal of Environmental Psychology*, *25*(2), 207–218. doi:10.1016/j.jenvp.2005.03.003

Krishnamoorthy, K. (2006). *Handbook of Statistical Distributions With Applications*. Boca Raton, FL: Chapman & Hall.

Lazaroiu, G. C., & Roscia, M. (2012). Definition Methodology for the Smart Cities Model. *Energy*, *47*(1), 326–332. doi:10.1016/j.energy.2012.09.028

Lee, C., Park, S., & Jeong, J. (2016). Comprehensive Comparison of Normality Tests: Empirical Study Using Many Different Types of Data. *Journal* *of the Korean Data and Information Science Society*, *27*(5), 1399–1412. doi:10.7465/jkdi.2016.27.5.1399

Lee, J. H., Hancock, M. G., & Hu, M. C. (2014). Towards an Effective Framework for Building Smart Cities: Lessons from Seoul and San Francisco. *Technological Forecasting and Social Change*, *89*, 80–99. doi:10.1016/j.techfore.2013.08.033

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting Outliers: Do not Use Standard Deviation Around the Mean, Use Absolute Deviation Around the Median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. doi:10.1016/j.jesp.2013.03.013

Li, C. (2013). Little's Test of Missing Completely at Random. *The Stata Journal*, *13*(4), 795–809.

Little, R. J. (1988). A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association*, *83*(404), 1198–1202.

Maletič, D., Maletič, M., Lovrenčić, V., Al–Najjar, B., & Gomišček, B. (2014). An Application of Analytic Hierarchy Process (AHP) and Sensitivity Analysis for Maintenance Policy Selection. *Organizacija*, *47*(3), 177–188. doi:10.2478/orga-2014-0016

McNeese, B. (2016). Are the Skewness and Kurtosis Useful Statistics? Retrieved December 10, 2018, from https://www.spcforexcel.com/knowledge/basic-statistics/are-skewness-and-kurtosis-useful-statistics

McNeish, D. M., & Harring, J. R. (2017). Clustered Data with Small Sample Sizes: Comparing the Performance of Model–Based and Design–Based Approaches. *Communications in Statistics – Simulation and Computation*, *46*(2), 855–869. doi:10.1080/03610918.2014.983648

Meijer, A., & Bolívar, M. P. R. (2016). Governing the Smart City: A Review of the Literature on Smart Urban Governance. *International Review of Administrative Sciences*, *82*(2), 392–408. doi:10.1177/0020852314564308

Mikulić, J., Kožić, I., & Krešić, D. (2015). Weighting Indicators of Tourism Sustainability: A Critical Note. *Ecological Indicators*, *48*, 312–314. doi:10.1016/j.ecolind.2014.08.026

Montalto, V., Jorge Tacao Moura, C., Langedijk, S., & Saisana, M. (2018a). *The Cultural and Creative Cities Monitor: 2017 Edition*. Luxembourg: Publications Office of the European Union.

Montalto, V., Jorge Tacao Moura, C., Langedijk, S., & Saisana, M. (2018b). *The Cultural and Creative Cities Monitor: 2017 Edition – Annex A: The Cultural and Creative Cities Monitor Methodology in Ten Steps*. Luxembourg: Publications Office of the European Union.

Montalto, V., Jorge Tacao Moura, C., Langedijk, S., & Saisana, M. (2018c). *The Cultural and Creative Cities Monitor: 2017 Edition – Annex B: Statistical Assessment of the Cultural and Creative Cities Index 2017*. Luxembourg: Publications Office of the European Union.

Monzon, A. (2015). Smart Cities Concept and Challenges: Bases for the Assessment of Smart City Projects. In M. Helfert, K.-H. Krempels, C. Klein, B. Donellan, & O. Guiskhin (Eds.), *Smart Cities, Green Technologies, and Intelligent Transport Systems* (Vol. 579, pp. 17–31). Cham, Germany: Springer International Publishing.

Morris, T. P., White, I. R., & Royston, P. (2014). Tuning Multiple Imputation by Predictive Mean Matching and Local Residual Draws. *BMC Medical Research Methodology*, *75*(14), 1–13.

Mukaka, M. M. (2012). Statistics Corner: A Guide to Appropriate Use of Correlation Coefficient in Medical Research. *Malawi Medical Journal*, *24*(3), 69–71.

Munda, G., & Nardo Michaela. (2005). Non–Compensatory Composite Indicators for Ranking Countries: A Defensible Setting. Ispra, Italy. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.108.5260&rep=rep1&type=pdf

Nam, T., & Pardo, T. A. (2011). Conceptualizing Smart City with Dimensions of Technology, People, and Institutions. In J. Bertot, K. Nahon, S. A. Chun, L. Luna–Reyes, & V. Atluri (Eds.), *Proceedings of the 12th Annual International Digital Government Research Conference on Digital Government Innovation in Challenging Times* (pp. 282–291). doi:10.1145/2037556.2037602

Nardo, M., Saisana, M., Saltelli, A., & Tarantola, S. (2005). Tools for Composite Indicators Building. Ispra, Italy. Retrieved from http://publications.jrc.ec.europa.eu/repository/bitstream/JRC31473/EUR%2021682%20EN.pdf

Nathans, L. L., Oswald, F. L., & Nimon, K. (2012). Interpreting Multiple Linear Regression: A Guidebook of Variable Importance. *Practical Assessment, Research & Evaluation*, *17*(9), 1–19.

Neirotti, P., de Marco, A., Cagliano, A. C., Mangano, G., & Scorrano, F. (2014). Current Trends in Smart City Initiatives: Some Stylised Facts. *Cities*, *38*, 25–36. doi:10.1016/j.cities.2013.12.010

Nicoletti, G., Scarpetta, S., & Boylaud, O. (2000). Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation. OECD Economics Department Working Papers No. 226, Paris, France. Retrieved from https://www.oecd-ilibrary.org/economics/summary-indicators-of-product-market-regulation-with-an-extension-to-employment-protection-legislation_215182844604;jsessionid=sgyBjmWU0Xgsi1jem9scnACy.ip-10-240-5-154

O'Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity*, *41*(5), 673–690. doi:10.1007/s11135-006-9018-6

OECD. (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Paris, France: OECD Publications.

Opricovic, S., & Tzeng, G.-H. (2004). Compromise Solution by MCDM Methods: A Comparative Analysis of VIKOR and TOPSIS. *European Journal of Operational Research*, *156*(2), 445–455. doi:10.1016/S0377-2217(03)00020-1

Osborne, J. W. (2002). The Effects of Minimum Values on Data Transformations. New Orleans, LA. Retrieved from https://files.eric.ed.gov/fulltext/ED463313.pdf

Osborne, J. W. (2010). Improving Your Data Transformations: Applying the Box–Cox Transformation. *Practical Assessment, Research & Evaluation*, *15*(12), 1–9.

Osborne, J. W., & Waters, E. (2002). Four Assumptions of Multiple Regression that Researchers Should Always Test. *Practical Assessment, Research & Evaluation*, *8*(2), 1–5.

Oueslati, W., Alvanides, S., & Garrod, G. (2015). Determinants of Urban Sprawl in European Cities. *Urban Studies*, *52*(9), 1594–1614. doi:10.1177/0042098015577773

Parkinson, M., Meegan, R., & Karecha, J. (2015). City Size and Economic Performance: Is Bigger Better, Small More Beautiful or Middling Marvellous? *European Planning Studies*, *23*(6), 1054–1068. doi:10.1080/09654313.2014.904998

Paruolo, P., Saisana, M., & Saltelli, A. (2013). Ratings and Rankings: Voodoo or Science? *Journal of the Royal Statistical Society. Series A*, *176*(3), 609–634.

Pearson, K. (1895). Note on Regression and Inheritance in the Case of Two Parents. *Proceedings of the Royal Society of London*, *58*, 240–242.

Pigott, T. D. (2001). A Review of Methods for Missing Data. *Educational Research and Evaluation*, *7*(4), 353–383. doi:10.1076/edre.7.4.353.8937

Raihan, S. (2011). Infrastructure and Growth and Poverty in Bangladesh. MPRA Paper No. 37882, Munich, Germany. Retrieved from https://mpra.ub.uni-muenchen.de/37882/3/MPRA_paper_37882.pdf

Rangone, A. (1996). An Analytical Hierarchy Process Framework for Comparing the Overall Performance of Manufacturing Departments. *International Journal of Operations & Production Management*, *16*(8), 104–119. doi:10.1108/01443579610125804

Ratner, B. (2010). Variable Selection Methods in Regression: Ignorable Problem, Outing Notable Solution. *Journal of Targeting, Measurement and Analysis for Marketing*, *18*(1), 65–75. doi:10.1057/jt.2009.26

Razali, N. M., & Wah, Y. B. (2011). Power Comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling Tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21–33.

Roland Berger. (2017). Smart City, Smart Strategy: Cities Around the World are Embracing the Digital Revolution. But how Well are They Really Doing? Munich, Germany. Retrieved from https://www.rolandberger.com/publications/publication_pdf/ta_17_008_smart_cities_online.pdf

Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association*, *88*(424), 1273–1283.

Rowley, H. V., Peters, G. M., Lundie, S., & Moore, S. J. (2012). Aggregating Sustainability Indicators: Beyond the Weighted Sum. *Journal of Environmental Management*, *111*, 24–33. doi:10.1016/j.jenvman.2012.05.004

Rubin, D. B. (1976). Inference and Missing Data. *Biometrika*, *63*(3), 581–592.

Rubin, D. B. (1996). Multiple Imputation After 18+ Years. *Journal of the American Statistical Association*, *91*(434), 473–489.

Saaty, T. L. (1977). A Scaling Method for Priorities in Hierarchial Structures. *Journal of Mathematical Psychology*, *15*(3), 234–281.

Saaty, T. L. (1990). How to Make a Decision: The Analytical Hierarchy Process. *European Journal of Operational Research*, *48*(1), 9–26.

Saaty, T. L. (2008a). Decision Making with the Analytic Hierarchy Process. *International Journal of Services Sciences*, *1*(1), 83–98. doi:10.1504/IJSSCI.2008.017590

Saaty, T. L. (2008b). The Analytical Network Process. *Iranian Journal of Operations Research*, *1*(1), 1–27.

Saisana, M. (2012). A Do–It–Yourself Guide in Excel for Composite Indicator Development. Ispra, Italy. Retrieved from https://ec.europa.eu/jrc/sites/jrcsh/files/20140922_JRC_COIN_COIN2014_ExcelguideforCIs.pdf

Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and Sensitivity Analysis Techniques as Tools for the Quality Assessment of Composite Indicators. *Journal of the Royal Statistical Society. Series A*, *168*(2), 307–323.

Salkind, N. J. (2010). *Encyclopedia of Research Design*. Thousand Oaks, CA: Sage Publications.

Saltelli, A. (2007). Composite Indicators Between Analysis and Advocacy. *Social Indicators Research*, *81*(1), 65–77. doi:10.1007/s11205-006-0024-9

Saltelli, A., & Annoni, P. (2010). How to Avoid a Perfunctory Sensitivity Analysis. *Environmental Modelling & Software*, *25*(12), 1508–1517. doi:10.1016/j.envsoft.2010.04.012

Saltelli, A., Ratto M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., . . . Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer*. Chichester, England: John Wiley & Sons.

Salvati, L., & Carlucci, M. (2014). A Composite Index of Sustainable Development at the Local Scale: Italy as a Case Study. *Ecological Indicators*, *43*, 162–171. doi:10.1016/j.ecolind.2014.02.021

Sands, G. R., & Podmore, T. H. (2000). A Generalized Environmental Sustainability Index for Agricultural Systems. *Agriculture, Ecosystems and Environment*, *1*(79), 29–41.

Sarstedt, M., & Mooi, E. (2014). Cluster Analysis. In M. Sarstedt & E. Mooi (Eds.), *A Concise Guide to Market Research* (pp. 273–324). doi:10.1007/978-3-642-53965-7_9

Saunders, J. A., Morrow–Howell, N., Spitznagel, E., Doré, P., Proctor, E. K., & Pescarino, R. (2006). Imputing Missing Data: A Comparison of Methods for Social Work Researchers. *Social Work Reserach*, *30*(1), 19–31.

Schafer, J. L., & Graham, J. W. (2002). Missing Data: Our View of the State of the Art. *Psychological Methods*, *7*(2), 147–177. doi:10.1037//1082-989X.7.2.147

Senlier, N., Yildiz, R., & Aktaş, E. D. (2009). A Perception Survey for the Evaluation of Urban Quality of Life in Kocaeli and a Comparison of the Life Satisfaction with the European Cities. *Social Indicators Research*, *94*(2), 213–226. doi:10.1007/s11205-008-9361-1

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, *52*(3–4), 591–611.

Sharpe, A., & Andrews, B. (2012). An Assessment of Weighting Methodologies for Composite Indicators: The Case of the Index of Economic Well–Being. CSLS Research Report No. 2012–10, Ottawa, Canada. Retrieved from http://www.csls.ca/reports/csls2012-10.pdf

Sinharay, S., Stern, H. S., & Russell, D. (2001). The Use of Multiple Imputation for the Analysis of Missing Data. *Psychological Methods*, *6*(4), 317–329. doi:10.1037/1082-989X.6.4.317

Song, Q., & Shepperd, M. (2007). Missing Data Imputation Techniques. *International Journal of Business Intelligence and Data Mining*, *2*(3), 261–290. doi:10.1504/IJBIDM.2007.015485

Stanković, J., Džunić, M., Džunić, Ž., & Marinković, S. (2017). A Multi–Criteria Evaluation of the European Cities' Smart Performance: Economic, Social and Environmental Aspects. *Proceedings of Rijeka Faculty of Economics:*

Journal of Economics and Business, *35*(2), 519–550. doi:10.18045/zbefri.2017.2.519

Stiglitz, J. E., Fitoussi, J.-P., & Durand, M. (2018). *Beyond GDP: Measuring What Counts for Economic and Social Performance*. Paris, France: OECD Publishing.

Taleb, N. N. (2007). *The Black Swan: The Impact of the Highly Improbable*. New York, NY: Random House.

Talukder, B., Hipel, K. W., & vanLoon, G. W. (2017). Developing Composite Indicators for Agricultural Sustainability Assessment: Effect of Normalization and Aggregation Techniques. *Resources*, *66*(6), 1–27. doi:10.3390/resources6040066

Tate, E. (2012). Social Vulnerability Indices: A Comparative Assessment Using Uncertainty and Sensitivity Analysis. *Natural Hazards*, *63*(2), 325–347. doi:10.1007/s11069-012-0152-2

United Nations. (2018). World Urbanization Prospects 2018. Retrieved November 27, 2018, from https://esa.un.org/unpd/wup/Download/

van de Vijver, F., & Tanzer, N. K. (2004). Bias and Equivalence in Cross–Cultural Assessment: An Overview. *European Review of Applied Psychology*, *54*(2), 119–135. doi:10.1016/j.erap.2003.12.004

Verardi, V., & Croux, C. (2009). Robust Regression in Stata. *The Stata Journal*, *9*(3), 439–453.

Vink, G., Frank, L. E., Pannekoek, J., & van Buuren, S. (2014). Predictive Mean Matching Imputation of Semicontinuous Variables. *Statistica Neerlandica*, *68*(1), 61–90. doi:10.1111/stan.12023

Wang, M. C., & Bushman, B. J. (1998). Using the Normal Quantile Plot to Explore Meta–Analytic Data Sets. *Psychological Methods*, *3*(1), 46–54.

Weisberg, S. (2005). *Applied Linear Regression* (3rd ed.). doi:10.1002/0471704091

Westfall, P. H. (2014). Kurtosis as Peakedness, 1905 – 2014. R.I.P. *The American Statistician*, *68*(3), 191–195. doi:10.1080/00031305.2014.917055

Węziak–Białowolska, D. (2016). Quality of Life in Cities – Empirical Evidence in Comparative European Perspective. *Cities*, *58*, 87–96. doi:10.1016/j.cities.2016.05.016

Wheeler, S. (2009). Regions, Megaregions, and Sustainability. *Regional Studies*, *43*(6), 863–876. doi:10.1080/00343400701861344

White, H. (1980). A Heteroskedasticity–Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, *48*(4), 817–838.

Wilcox, R. R., & Keselman, H. J. (2004). Robust Regression Methods: Achieving Small Standard Errors when there Is Heteroscedasticity. *Understanding Statistics*, *3*(4), 349–364. doi:10.1207/s15328031us0304_8

Wulff, J. N., & Ejlskov, L. (2017). Multiple Imputation by Chained Equations in Praxis: Guidelines and Review. *The Electronic Journal of Business Research Methods*, *15*(1), 41–56.

Yuan, Lim, Lan, Yuen, B., & Low, C. (1999). *Urban Quality of Life: Critical Issues and Options*. Singapore: School of Building and Real Estate, National University of Singapore.

Zhou, P., & Ang, B. W. (2009). Comparing MCDA Aggregation Methods in Constructing Composite Indicators Using the Shannon–Spearman Measure. *Social Indicators Research*, *94*(1), 83–96. doi:10.1007/s11205-008-9338-0

Zhou, P., Ang, B. W., & Poh, K. L. (2006). Comparing Aggregating Methods for Constructing the Composite Environmental Index: An Objective Measure. *Ecological Economics*, *59*(3), 305–311. doi:10.1016/j.ecolecon.2005.10.018

Zhou, P., Fan, L.-W., & Zhou, D.-Q. (2010). Data Aggregation in Constructing Composite Indicators: A Perspective of Information Loss. *Expert Systems with Applications*, *37*(1), 360–365. doi:10.1016/j.eswa.2009.05.039

Appendix

Table A1: Sample of European Cities and Their Classification Codes

#	City	Country	NUTS 0	NUTS 2	Metropolitan Code	City Code
1	Graz	Austria	AT	AT22	AT002M	AT002C1
2	Vienna	Austria	AT	AT13	AT001MC	AT001C1
3	Antwerp	Belgium	BE	BE10	BE002M	BE002C1
4	Brussels	Belgium	BE	BE20	BE001MC	BE001C1
5	Liège	Belgium	BE	BE33	BE005M	BE005C1
6	Burgas	Bulgaria	BG	BG34	BG004M	BG004C1
7	Sofia	Bulgaria	BG	BG41	BG001MC	BG001C1
8	Nicosia	Cyprus	CY	CY00	CY001MC	CY001C1
9	Ostrava	Czech Republic	CZ	CZ08	CZ003M	CZ003C1
10	Prague	Czech Republic	CZ	CZ01	CZ001MC	CZ001C1
11	Berlin	Germany	DE	DE30	DE001MC	DE001C1
12	Hamburg	Germany	DE	DE60	DE002M	DE002C1
13	Leipzig	Germany	DE	DED5	DE008M	DE008C1
14	Munich	Germany	DE	DE21	DE003M	DE003C1
15	Rostock	Germany	DE	DE80	DE043M	DE043C1
16	Aalborg	Denmark	DK	DK05	DK004M	DK004C2
17	Copenhagen	Denmark	DK	DK01	DK001MC	DK001C1
18	Tallinn	Estionia	EE	EE00	EE001MC	EE001C1
19	Athens	Greece	EL	EL30	EL001MC	EL001C1
20	Barcelona	Spain	ES	ES51	ES002M	ES002C1
21	Madrid	Spain	ES	ES30	ES001MC	ES001C1
22	Málaga	Spain	ES	ES61	ES006M	ES006C1
23	Oviedo	Spain	ES	ES12	ES013M	ES013C1
24	Helsinki	Finland	FI	FI1B	FI001MC	FI001C2
25	Bordeaux	France	FR	FR61	FR007M	FR007C1
26	Lille	France	FR	FR30	FR009M	FR009C1
27	Marseille	France	FR	FR82	FR203M	FR203C1
28	Paris	France	FR	FR10	FR001MC	FR001C1
29	Rennes	France	FR	FR52	FR013M	FR013C2
30	Strasbourg	France	FR	FR42	FR006M	FR006C2
31	Zagreb	Croatia	HR	HR04	HR001MC	HR001C1
32	Budapest	Hungary	HU	HU10	HU001MC	HU001C1
33	Miskolc	Hungary	HU	HU31	HU002M	HU002C1
34	Dublin	Ireland	IE	IE02	IE001MC	IE001C1
35	Bologna	Italy	IT	ITH5	IT009M	IT009C1
36	Naples	Italy	IT	ITF3	IT003M	IT003C1
37	Palermo	Italy	IT	ITG1	IT005M	IT005C1
38	Rome	Italy	IT	ITI4	IT001MC	IT001C1
39	Turin	Italy	IT	ITC1	IT004M	IT004C1
40	Verona	Italy	IT	ITH3	IT012M	IT012C1

41	Vilnius	Lithuania	LT	LT00	LT001MC	LT001C1
42	Luxembourg	Luxembourg	LU	LU00	LU001MC	LU001C1
43	Riga	Latvia	LV	LV00	LV001MC	LV001C1
44	Valletta	Malta	MT	MT00	MT001MC	MT001C1
45	Amsterdam	Netherlands	NL	NL32	NL002MC	NL002C1
46	Groningen	Netherlands	NL	NL11	NL007M	NL007C1
47	Rotterdam	Netherlands	NL	NL33	NL003M	NL003C1
48	Białystok	Poland	PL	PL34	PL011M	PL011C1
49	Gdańsk	Poland	PL	PL63	PL006M	PL006C1
50	Kraków	Poland	PL	PL21	PL003M	PL003C1
51	Warsaw	Poland	PL	PL12	PL001MC	PL001C1
52	Lisbon	Portugal	PT	PT17	PT001MC	PT001C1
53	Bucharest	Romania	RO	RO32	RO001MC	RO001C1
54	Cluj-Napoca	Romania	RO	RO11	RO002M	RO002C1
55	Malmö	Sweden	SE	SE22	SE003M	SE003C1
56	Stockholm	Sweden	SE	SE11	SE001MC	SE001C1
57	Ljubljana	Slovenia	SI	SI0	SI001MC	SI001C1
58	Bratislava	Slovakia	SK	SK01	SK001MC	SK001C1
59	Košice	Slovakia	SK	SK04	SK002M	SK002C1
60	Belfast	United Kingdom	UK	UKN0	UK012M	UK012C1
61	Cardiff	United Kingdom	UK	UKL2	UK009M	UK009C1
62	Glasgow	United Kingdom	UK	UKM3	UK004M	UK004C1
63	London	United Kingdom	UK	UKI1	UK001MC	UK001K2
				UKI2		
64	Manchester	United Kingdom	UK	UK3	UK008M	UK008C1
65	Newcastle	United Kingdom	UK	UKC2	UK013M	UK013C1

NUTS 2 Codes are from 2010.; Metropolitan codes for capitals end with a 'C'.

Table A2: Indicators of the Objective SCCI

#	Dimensions & Indicators	Unit	Year	Level
A	**Infrastructure & Mobility**			
A1O	Access to high-level passenger transport infrastructure	Minute-equivalent	2012	City
A2O	Accessibility to population by rail	Number of persons	2011	City
A3O	Accessibility of public transportation	Score	2015	City
A4O	Departures of public transportation	Number of departures	2015	City
A5O	Accessibility of motorways	Index	2014	Regional
A6O	Accessibility to passenger flights	Number of flight access	2014	City
A7O	Potential accessibility by roads	Number of road access	2015	City
A8O	Length of local roads	Meter per inhabitant	2014	City
A9O	Death by transport accident	Rate	2015	Regional
A10O	Network efficiency	Score	2015	City
A11O	Transportation and storage	Local units per inhabitant	2015	Regional
A12O	Construction	Local units per inhabitant	2015	Regional
A13O	Establishments	Local units per inhabitant	2015	Regional
A14O	Built-up areas	Square meter per inhabitant	2011	City
A15O	Old buildings	Percentage	2015	City
A16O	Next generation internet access	Score	2016	City
A17O	Sights & landmarks	Number per 100,000 inhabitants	2016	City
A18O	Green infrastructure	Percentage	2010 - 2020	City
A19O	Nature based recreation opportunities	Score	2010 - 2020	City
B	**Living & Social Cohesion**			
B1O	Infant mortality	Percentage	2015	Regional
B2O	Life expectancy	Years	2015	Regional
B3O	Medical doctors	Number per 100,000 inhabitants	2015	Regional
B4O	Hospital beds	Number per 100,000 inhabitants	2015	Regional
B5O	Educational attainment	Percentage	2015	Regional
B6O	Early leavers from education and training	Percentage	2015	Regional
B7O	Pre-primary education	Percentage	2015	Regional
B8O	Museums	Number per 100,000 inhabitants	2016	City
B9O	Rooms	Number per inhabitant	Various	Regional
B10O	Rent	Index	2015	City
B11O	Tourism	Number per 1,000 inhabitants	2015	Regional
B12O	Tourism intensity	Ratio	2016	City
B13O	Net migration	Percentage	2005 - 2015	City
B14O	Foreign-born population	Percentage	2011	City
B15O	People at risk of poverty or social exclusion	Percentage	2015	Regional
B16O	Difference female vs. male employment	Percentage	2016	Regional
B17O	Gini coefficient	Scale	2015	National
B18O	Total population change	Percentage	2015	City
C	**Economy & Governance**			
C1O	GDP in PPS	Euro per inhabitant	2015	City
C2O	Young people neither in employment nor education/training	Percentage	2015	Regional
C3O	Employment	Percentage	2015	City
C4O	Long-term unemployment	Percentage	2015	Regional
C5O	Patents	Number per inhabitant	2012	City
C6O	Research & development	Percentage	2015	Regional
C7O	Employment in high-technology sectors	Percentage	2015	Regional
C8O	Potential market size	Index	2013	Regional
C9O	Risk factors linked to globalisation	Score	2016	Regional
C10O	Regional resilience	Score	2015	Regional
C11O	Vulnerability to tourism	Score	2016	Regional
C12O	Ease of doing business	Rank	2015	National

C13O	Government bond yields	Percentage	2015	National
C14O	Efficiency of the legal system - criminal cases	Percentage	2015	National
C15O	Efficiency of the legal system - civil and/or commercial cases	Percentage	2015	National
C16O	Interaction with public authorities via the internet	Percentage	2015	Regional
C17O	Public procurement - single bidder	Percentage	2015	Regional
C18O	Public procurement - open call for tender	Percentage	2015	Regional
C19O	Voter turnout at national elections	Percentage	2015	Regional
C20O	Information and communication	Local units per inhabitant	2015	Regional
C21O	Theft	Per 100,000 inhabitants	2015	National
C22O	Intentional homicide	Per 100,000 inhabitants	2015	National
D	**Environment & Sustainability**			
D1O	Green urban areas	Percentage	2012	City
D2O	Green urban areas in neighbourhood	Percentage	2012	City
D3O	Green metropolitan area	Square meters per inhabitant	2014	City
D4O	Artificial areas	Square meters per inhabitant	2010 - 2020	City
D5O	PM10 concentration	Microgram per cubic meter	2010 - 2020	City
D6O	NOx emissions	100 kg per year and inhabitant	2010 - 2020	Regional
D7O	NO2 concentration	Microgram per cubic meter	2010 - 2020	Regional
D8O	Removal capacity of NO2 by vegetation	Kilogram per hectare and year	2015	City
D9O	Noise from roads	Percentage	2012 - 2017	City
D10O	Waste & water	Local units per inhabitant	2015	Regional
D11O	Waste generated	1,000 tonnes per inhabitant	2013	Regional
D12O	Recycling of material	1,000 tonnes per inhabitant	2013	Regional
D13O	Transport fuel taxes	Percentage	2015	National
D14O	Renewable sources	Percentage	2015	Regional

Table A3: Indicators of the Subjective SCCI

#	Dimensions & Indicators	Questions and Statements
A	**Infrastructure & Mobility**	
A1S	Public transportation	How satisfied are you with public transport, for example the bus, tram or metro in [CITY NAME]?
A2S	Streets and buildings	How satisfied are you with the state of the streets and buildings in your neighborhood in [CITY NAME]?
A3S	Public spaces	How satisfied are you with public spaces such as markets, squares, pedestrian areas in [CITY NAME]?
B	**Living & Social Cohesion**	
B1S	Health care	How satisfied are you with health care services, doctors and hospitals in [CITY NAME]?
B2S	Educational facilities	How satisfied are you with schools and other educational facilities in [CITY NAME]?
B3S	Cultural facilities	How satisfied are you with cultural facilities such as concert halls, theatres, museums and libraries in [CITY NAME]?
B4S	Affordable housing	It is easy to find good housing at a reasonable price in [CITY NAME].
B5S	Acceptance of foreigners	The presence of foreigners is good for [CITY NAME].
B6S	Satisfaction to live in the city	I am satisfied to live in [CITY NAME].
C	**Economy & Governance**	
C1S	Job situation	It is easy to find a job in [CITY NAME].
C2S	Job opportunities	How satisfied are you with your personal job situation?
C3S	Administrative services	The administrative services of [CITY NAME] help people efficiently.
C4S	Trust in public administration	Generally speaking, the public administration of [CITY NAME] can be trusted.
C5S	Safety in city	I feel safe in [CITY NAME].
D	**Environment & Sustainability**	
D1S	Green spaces	How satisfied are you with green spaces such as parks and gardens in [CITY NAME]?
D2S	Air quality	How satisfied are you with the quality of the air in [CITY NAME]?
D3S	Noise	How satisfied are you with the noise level in [CITY NAME]?
D4S	Cleanliness	How satisfied are you with the cleanliness in [CITY NAME]?
D5S	Climate change	[CITY NAME] is committed to fight against climate change (e.g.: energy efficiency, green transport).

Some questions and statements were shortened for better clarity.

Table A4: Little's MCAR Test

Dimensions	A	B	C	D
χ^2 distance	271.7813	292.5054	236.4106	151.7572
Df	218	212	208	134
Prob > χ^2	0.0077	0.0002	0.0667	0.1400

Table A5: Doornik–Hansen Multivariate Normality Test

Dimensions	A	B	C	D
χ^2 distance	146.289	52.998	340.588	70.627
Df	38	36	44	28
Prob > χ^2	0.0000	0.0337	0.0000	0.0000

Table A6: Shapiro–Wilk Normality Test Statistics of the Objective SCCI

Indicator	Obs.	W	V	z	Prob > z	Indicator	Obs.	W	V	z	Prob > z
A1O	65	0.90817	5.323	3.621	0.00015	**C1O**	65	0.95056	2.866	2.28	0.01130
A2O	65	0.83645	9.481	4.871	0.00000	C2O	65	0.89052	6.346	4.001	0.00003
A3O	65	0.94666	3.092	2.445	0.00725	**C3O**	65	0.96949	1.769	1.235	0.10846
A4O	65	0.6943	17.721	6.225	0.00000	C4O	65	0.80913	11.065	5.205	0.00000
A5O	65	0.87652	7.158	4.262	0.00001	C5O	65	0.74885	14.559	5.799	0.00000
A6O	65	0.77452	13.071	5.566	0.00000	C6O	65	0.93032	4.039	3.023	0.00125
A7O	65	0.93413	3.818	2.901	0.00186	C7O	65	0.93193	3.946	2.973	0.00148
A8O	65	0.68524	18.246	6.288	0.00000	C8O	65	0.69491	17.686	6.221	0.00000
A9O	65	0.9586	2.4	1.896	0.02899	**C9O**	65	0.99050	0.551	-1.292	0.90174
A10O	65	0.20248	46.232	8.302	0.00000	**C10O**	65	0.99364	0.369	-2.161	0.98466
A11O	65	0.82391	10.208	5.031	0.00000	C11O	65	0.90903	5.273	3.600	0.00016
A12O	65	0.92133	4.56	3.286	0.00051	**C12O**	65	0.96117	2.251	1.757	0.03947
A13O	65	0.49961	29.007	7.292	0.00000	C13O	65	0.72204	16.113	6.019	0.00000
A14O	65	0.94599	3.131	2.472	0.00673	**C14O**	65	0.95160	2.806	2.234	0.01274
A15O	65	0.95861	2.399	1.895	0.02905	C15O	65	0.89642	6.004	3.881	0.00005
A16O	65	0.96122	2.248	1.754	0.03969	**C16O**	65	0.98413	0.92	-0.18	0.57154
A17O	65	0.86202	7.999	4.503	0.00000	C17O	65	0.91438	4.963	3.469	0.00026
A18O	65	0.92051	4.608	3.308	0.00047	C18O	65	0.92657	4.257	3.137	0.00085
A19O	65	0.85035	8.675	4.678	0.00000	**C19O**	65	0.97681	1.344	0.641	0.26083
						C20O	65	0.83676	9.463	4.867	0.00000
B1O	65	0.82632	10.068	5.001	0.00000	C21O	65	0.89897	5.857	3.828	0.00006
B2O	65	0.91603	4.868	3.427	0.00031	C22O	65	0.64194	20.757	6.567	0.00000
B3O	65	0.88822	6.480	4.046	0.00003						
B4O	65	0.93016	4.049	3.028	0.00123	D1O	65	0.90386	5.573	3.72	0.00010
B5O	65	0.92115	4.571	3.291	0.00050	D2O	65	0.81815	10.542	5.1	0.00000
B6O	65	0.92370	4.423	3.220	0.00064	D3O	65	0.91873	4.711	3.356	0.00039
B7O	65	0.98352	0.955	-0.099	0.53942	D4O	65	0.87418	7.294	4.303	0.00001
B8O	65	0.83410	9.617	4.902	0.00000	**D5O**	65	0.96704	1.911	1.402	0.08048
B9O	65	0.95322	2.712	2.160	0.01538	D6O	65	0.74092	15.019	5.867	0.00000
B10O	65	0.96028	2.303	1.806	0.03546	**D7O**	65	0.98445	0.901	-0.225	0.58903
B11O	65	0.87448	7.276	4.298	0.00001	D8O	65	0.90862	5.297	3.61	0.00015
B12O	65	0.70527	17.085	6.146	0.00000	**D9O**	65	0.97541	1.425	0.767	0.22142
B13O	65	0.99195	0.467	-1.651	0.95059	D10O	65	0.65666	19.903	6.477	0.00000
B14O	65	0.93941	3.512	2.720	0.00326	**D11O**	65	0.98768	0.714	-0.729	0.76688
B15O	65	0.85418	8.453	4.622	0.00000	D12O	65	0.92335	4.444	3.230	0.00062
B16O	65	0.93154	3.968	2.985	0.00142	D13O	65	0.93862	3.558	2.748	0.00299
B17O	65	0.98423	0.914	-0.195	0.57718	D14O	65	0.92582	4.300	3.159	0.00079
B18O	65	0.97707	1.329	0.616	0.26888						

Indicators assumable from a normal distribution at a significance level of 0.01 are bold.

Table A7: Skewness and Kurtosis Values of the Objective SCCI

Indicator	Obs.	Skewness	Kurtosis	Indicator	Obs.	Skewness	Kurtosis
A1O	65	-0.9930759	0.38277766	C1O	65	0.91616664	1.36175752
A2O	65	1.39595303	1.50992062	C2O	65	-1.49344183	3.30340235
A3O	65	0.59548528	0.79431928	C3O	65	-0.108660734	0.549785766
A4O	65	2.62327497	7.53012523	C4O	65	-1.76999028	3.2134624
A5O	65	1.06388345	0.23201139	C5O	65	1.965330417	3.716351538
A6O	65	2.00036961	4.61913705	C6O	65	0.96116201	0.622739695
A7O	65	0.97060469	0.91294555	C7O	65	0.68708401	-0.50243705
A8O	65	3.48383607	17.2398988	C8O	65	2.840917882	10.23533884
A9O	65	-0.48453284	-0.48872024	C9O	65	-0.051539848	-0.622693073
A10O	65	-7.60657502	59.8939238	C10O	65	0.00997134	-0.1413438
A11O	65	2.25224694	8.89587112	C11O	65	-1.146571	0.95942534
A12O	65	1.19782126	2.76336464	C12O	65	-0.52200038	0.75789138
A13O	65	4.9699287	30.6647432	C13O	65	-2.945457471	13.36290938
A14O	65	0.97345545	1.49828294	C14O	65	0.026903137	-0.351678157
A15O	65	0.82563855	1.8395405	C15O	65	-0.93798135	1.294091062
A16O	65	0.5918007	-0.25040012	C16O	65	0.01360512	-0.55647759
A17O	65	1.60242314	3.11962994	C17O	65	-0.91269015	0.1164746
A18O	65	1.07311891	1.22738089	C18O	65	-0.82137009	-0.00277348
A19O	65	1.33264547	1.08102651	C19O	65	-0.206551568	-0.642339373
				C20O	65	1.6241959	2.75842338
B1O	65	-1.94666838	5.03534779	C21O	65	-0.86246104	0.66300092
B2O	65	-0.86114613	-0.06265156	C22O	65	-3.300443078	13.79399311
B3O	65	1.581470786	3.741482936				
B4O	65	0.696767855	0.986042559	D1O	65	1.15361874	1.10848122
B5O	65	-1.037173768	0.796089071	D2O	65	-1.99887174	5.34313052
B6O	65	-0.993335592	0.752131347	D3O	65	1.023819744	1.01200204
B7O	65	-0.238867518	0.282685371	D4O	65	1.66928364	7.80429915
B8O	65	1.593666995	2.355733575	D5O	65	-0.0832106	1.22817205
B9O	65	-0.414531864	-0.866129599	D6O	65	-1.86439391	2.81868244
B10O	65	-0.803255747	0.828653001	D7O	65	-0.10193783	0.10325148
B11O	65	1.466833895	2.47602276	D8O	65	0.72923508	-0.37719255
B12O	65	2.645396949	8.647907606	D9O	65	-0.326868092	-0.331579229
B13O	65	-0.005917916	-0.050462618	D10O	65	3.413924198	14.67297776
B14O	65	0.870212719	1.647183712	D11O	65	0.097486355	-0.472531263
B15O	65	-1.806575618	4.675814218	D12O	65	1.177514319	2.819834993
B16O	65	-1.021409936	1.431526607	D13O	65	0.84117417	0.20460954
B17O	65	-0.178592429	-0.684022811	D14O	65	0.806798569	-0.292017544
B18O	65	0.372240084	0.181954691				

Table A8: Shapiro–Wilk Normality Test Statistics of the Subjective SCCI

Indicator	Obs.	W	V	z	Prob > z	Indicator	Obs.	W	V	z	Prob > z
A1S	65	0.92633	4.271	3.144	0.00083	**C1S**	65	0.96539	2.007	1.508	0.06577
A2S	65	0.95379	2.679	2.134	0.01643	**C2S**	65	0.96586	1.979	1.478	0.06967
A3S	65	0.899	5.855	3.827	0.00006	**C3S**	65	0.96025	2.304	1.808	0.03532
B1S	65	0.92914	4.108	3.06	0.00111	**C4S**	65	0.97718	1.323	0.606	0.27224
B2S	65	0.97752	1.303	0.574	0.28315	**C5S**	65	0.97387	1.515	0.899	0.18432
B3S	65	0.95482	2.619	2.085	0.01855	D1S	65	0.89992	5.802	3.807	0.00007
B4S	65	0.97731	1.315	0.594	0.27634	**D2S**	65	0.97735	1.313	0.59	0.27759
B5S	65	0.96556	1.996	1.497	0.06720	**D3S**	65	0.96457	2.054	1.558	0.05957
B6S	65	0.9315	3.971	2.986	0.00141	**D4S**	65	0.97338	1.543	0.939	0.17384
						D5S	65	0.96013	2.311	1.814	0.03483

Indicators assumable from a normal distribution at a significance level of 0.01 are bold.

Table A9: Skewness and Kurtosis Values of the Subjective SCCI

Indicator	Obs.	Skewness	Kurtosis	Indicator	Obs.	Skewness	Kurtosis
A1S	65	-1.183034948	2.13728946	C1S	65	-0.56362709	-0.13567343
A2S	65	-0.705252292	0.41789435	C2S	65	-0.41799865	-0.53705193
A3S	65	-1.205859154	1.22396156	C3S	65	-0.70642086	0.37674407
B1S	65	-0.37176093	-1.17336032	C4S	65	-0.47468362	-0.1765631
B2S	65	-0.32299069	-0.53693188	C5S	65	-0.43683256	-0.41998561
B3S	65	-0.75457753	0.59502432	D1S	65	1.03848526	-1.14391365
B4S	65	-0.26273541	-0.49823392	D2S	65	-0.76997387	-0.19416396
B5S	65	-0.45397299	-0.51565642	D3S	65	-0.80718333	-0.33387777
B6S	65	-0.87722398	0.34785626	D4S	65	0.2266839	-0.50720111
				D5S	65	-0.37471363	-0.5499912

Table A10: KMO of the Objective SCCI

Indicator	KMO	Indicator	KMO	Indicator	KMO	Indicator	KMO
A1O	0.7450	B1O	0.4827	C1O	0.8171	D1O	0.5222
A2O	0.7622	B2O	0.6590	C2O	0.7308	D2O	0.3919
A3O	0.5822	B3O	0.4321	C3O	0.8997	D3O	0.7236
A4O	0.5633	B4O	0.5753	C4O	0.6785	D4O	0.4980
A5O	0.7311	B5O	0.4804	C5O	0.7100	D5O	0.5699
A6O	0.8496	B6O	0.4650	C6O	0.6822	D6O	0.3757
A7O	0.8888	B7O	0.2806	C7O	0.6729	D7O	0.5974
A8O	0.6629	B8O	0.5435	C8O	0.3666	D8O	0.6920
A9O	0.6276	B9O	0.6116	C9O	0.5388	D9O	0.3885
A10O	0.7252	B10O	0.5870	C10O	0.6878	D10O	0.5445
A11O	0.5707	B11O	0.7121	C11O	0.3626	D11O	0.4481
A12O	0.2718	B12O	0.6772	C12O	0.4711	D12O	0.5572
A13O	0.3866	B13O	0.6731	C13O	0.6280	D13O	0.7238
A14O	0.6060	B14O	0.7200	C14O	0.3133	D14O	0.3952
A15O	0.6658	B15O	0.7136	C15O	0.3421		
A16O	0.7254	B16O	0.3474	C16O	0.6632		
A17O	0.5008	B17O	0.6351	C17O	0.7249		
A18O	0.6402	B18O	0.6548	C18O	0.5889		
A19O	0.3524			C19O	0.2466		
				C20O	0.6952		
				C21O	0.6302		
				C22O	0.3263		
Overall:	0.6771	**Overall:**	0.5824	**Overall:**	0.6271	**Overall:**	0.5425

Table A11: KMO of the Subjective SCCI

Indicator	KMO	Indicator	KMO	Indicator	KMO	Indicator	KMO
A1S	0.8495	B1S	0.6117	C1S	0.7430	D1S	0.8574
A2S	0.6695	B2S	0.7217	C2S	0.6756	D2S	0.7837
A3S	0.6699	B3S	0.8580	C3S	0.7120	D3S	0.7691
		B4S	0.3754	C4S	0.7266	D4S	0.8959
		B5S	0.4914	C5S	0.8747	D5S	0.9212
		B6S	0.6064				
Overall:	0.7119	**Overall:**	0.6254	**Overall:**	0.7388	**Overall:**	0.8369

Table A12: Eigenvalues and Variance of the PCFA of the Objective SCCI

Dimension A:

Factor	Eigenvalue	Difference	Variance (%)	Cumulative Variance (%)
1	5.58719	2.46551	0.2941	0.2941
2	3.12167	1.13403	0.1643	0.4584
3	1.98764	0.43584	0.1046	0.563
4	1.5518	0.18161	0.0817	0.6446
5	1.37018	0.32069	0.0721	0.7168
6	1.0495	0.18375	0.0552	0.772
7	0.86575	0.15529	0.0456	0.8176
8	0.71046	0.09317	0.0374	0.855
9	0.61728	0.07245	0.0325	0.8874
10	0.54483	0.19576	0.0287	0.9161
11	0.34907	0.02517	0.0184	0.9345
12	0.3239	0.06098	0.017	0.9515
13	0.26293	0.02922	0.0138	0.9654
14	0.23371	0.07814	0.0123	0.9777
15	0.15557	0.05335	0.0082	0.9859
16	0.10222	0.02159	0.0054	0.9912
17	0.08063	0.01109	0.0042	0.9955
18	0.06954	0.0534	0.0037	0.9992
19	0.01614		0.0008	1

Retained factors are bold.

Dimension B:

Factor	Eigenvalue	Difference	Variance (%)	Cumulative Variance (%)
1	4.41778	1.25234	0.2454	0.2454
2	3.16543	1.47190	0.1759	0.4213
3	1.69353	0.34654	0.0941	0.5154
4	1.34699	0.04997	0.0748	0.5902
5	1.29703	0.20706	0.0721	0.6623
6	1.08997	0.10200	0.0606	0.7228
7	0.98796	0.17573	0.0549	0.7777
8	0.81223	0.15089	0.0451	0.8228
9	0.66135	0.09213	0.0367	0.8596
10	0.56922	0.08236	0.0316	0.8912
11	0.48686	0.07069	0.0270	0.9182
12	0.41617	0.11206	0.0231	0.9414
13	0.30411	0.09365	0.0169	0.9583
14	0.21045	0.01995	0.0117	0.9699
15	0.19050	0.02822	0.0106	0.9805
16	0.16228	0.03765	0.0090	0.9895
17	0.12463	0.06113	0.0069	0.9965
18	0.06350		0.0035	1

Retained factors are bold.

Dimension C:

Factor	Eigenvalue	Difference	Variance (%)	Cumulative Variance (%)
1	6.93415	4.57745	0.3152	0.3152
2	2.35670	0.27626	0.1071	0.4223
3	2.08044	0.26042	0.0946	0.5169
4	1.82003	0.24519	0.0827	0.5996
5	1.57484	0.30329	0.0716	0.6712
6	1.27155	0.30151	0.0578	0.7290
7	0.97004	0.07263	0.0441	0.7731
8	0.89741	0.22766	0.0408	0.8139
9	0.66975	0.12637	0.0304	0.8443
10	0.54338	0.00683	0.0247	0.8690
11	0.53655	0.08283	0.0244	0.8934
12	0.45372	0.02862	0.0206	0.9140
13	0.42510	0.10538	0.0193	0.9333
14	0.31972	0.02605	0.0145	0.9479
15	0.29367	0.08831	0.0133	0.9612
16	0.20535	0.01892	0.0093	0.9706
17	0.18643	0.02491	0.0085	0.9790
18	0.16152	0.04908	0.0073	0.9864
19	0.11245	0.03891	0.0051	0.9915
20	0.07353	0.00307	0.0033	0.9948
21	0.07046	0.02722	0.0032	0.998
22	0.04323		0.0020	1

Retained factors are bold.

Dimension D:

Factor	Eigenvalue	Difference	Variance (%)	Cumulative Variance (%)
1	2.88793	0.63591	0.2063	0.2063
2	2.25202	0.55820	0.1609	0.3671
3	1.69382	0.15909	0.1210	0.4881
4	1.53473	0.29446	0.1096	0.5977
5	1.24027	0.23141	0.0886	0.6863
6	1.00886	0.17064	0.0721	0.7584
7	0.83822	0.25732	0.0599	0.8183
8	0.58091	0.04325	0.0415	0.8598
9	0.53765	0.13857	0.0384	0.8982
10	0.39908	0.06436	0.0285	0.9267
11	0.33472	0.02744	0.0239	0.9506
12	0.30728	0.08879	0.0219	0.9725
13	0.21849	0.05246	0.0156	0.9881
14	0.16603		0.0119	1

Retained factors are bold.

Table A13: Eigenvalues and Variance of the PCFA of the Subjective SCCI

Dimension A

Factor	Eigenvalue	Difference	Variance (%)	Cumulative Variance (%)
1	2.41233	2.01032	0.8041	0.8041
2	0.40201	0.21635	0.134	0.9381
3	0.18566		0.0619	1

Retained factors are bold.

Dimension B:

Factor	Eigenvalue	Difference	Variance (%)	Cumulative Variance (%)
1	2.88479	1.69409	0.4808	0.4808
2	1.1907	0.15461	0.1985	0.6792
3	1.03609	0.58627	0.1727	0.8519
4	0.44982	0.15267	0.075	0.9269
5	0.29715	0.15571	0.0495	0.9764
6	0.14144		0.0236	1

Retained factors are bold.

Dimension C:

Factor	Eigenvalue	Difference	Variance (%)	Cumulative Variance (%)
1	3.48375	2.70294	0.6968	0.6968
2	0.78081	0.27782	0.1562	0.8529
3	0.50299	0.35543	0.1006	0.9535
4	0.14756	0.06268	0.0295	0.983
5	0.08488		0.017	1

Retained factors are bold.

Dimension D:

Factor	Eigenvalue	Difference	Variance (%)	Cumulative Variance (%)
1	3.85574	3.35668	0.7711	0.7711
2	0.49906	0.17946	0.0998	0.871
3	0.3196	0.09414	0.0639	0.9349
4	0.22545	0.1253	0.0451	0.98
5	0.10015		0.02	1

Retained factors are bold.

Table A14: Factor Loadings and Explained Variance of the Objective SCCI

Dimension A:

Indicator	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
A1O	0.459	0.3529	0.5341	0.0798	0.2358	0.1292
A2O	0.9388	0.0326	-0.0018	-0.032	0.171	-0.0515
A3O	0.0373	0.5906	0.3485	0.1436	0.5209	0.023
A4O	-0.0295	0.7987	0.1736	0.0941	0.2343	-0.0793
A5O	0.929	0.1018	0.0833	0.0521	0.1632	-0.059
A6O	0.7161	0.2388	0.3231	0.209	0.1628	-0.1077
A7O	0.9415	0.1359	0.0407	0.0439	0.1047	0.0244
A8O	-0.2449	-0.8557	0.1451	0.1291	0.0273	0.1694
A9O	0.3255	0.0285	0.7098	-0.1152	0.3459	-0.0787
A10O	0.7867	-0.0811	0.0125	-0.066	-0.1377	0.2924
A11O	-0.2596	0.3298	-0.0915	0.6895	-0.2022	0.169
A12O	0.0888	-0.2292	0.0052	0.8328	-0.0447	-0.0878
A13O	0.0459	-0.1476	-0.1866	-0.1031	-0.1201	0.8493
A14O	0.0249	-0.8251	-0.0596	0.0016	0.2578	0.0967
A15O	-0.6447	0.1299	-0.1636	0.294	-0.0385	0.262
A16O	0.1659	0.4963	0.0187	0.5331	0.1654	0.0522
A17O	-0.1214	-0.1675	0.2173	0.352	0.3877	0.6895
A18O	-0.2878	-0.0044	0.0797	0.1484	-0.7891	0.0014
A19O	-0.0194	0.0276	0.7958	-0.0144	-0.3721	-0.0874
Explained Variance	4.7605	3.0937	1.8356	1.8158	1.6786	1.4836
% of Total Variance	32.46	21.09	12.51	12.38	11.44	10.11

Dimension B:

Indicator	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
B1O	-0.0351	0.5052	0.0113	0.2096	0.6049	0.1541
B2O	-0.5271	0.571	0.3103	0.1435	0.2732	-0.0853
B3O	0.213	0.0131	0.2324	0.4318	0.0621	-0.4933
B4O	0.7109	0.0242	-0.0226	0.0819	-0.404	-0.1154
B5O	0.8471	0.1379	0.0214	0.1476	-0.1449	0.1385
B6O	0.6479	0.3631	0.0147	0.1113	0.2489	0.0002
B7O	0.1157	0.1233	0.0246	0.1007	-0.0199	0.8918
B8O	-0.027	-0.0138	0.3879	-0.0315	0.7947	-0.174
B9O	-0.6612	0.3683	0.1577	0.3976	0.0175	-0.2182
B10O	0.7431	-0.0084	-0.3279	0.0969	0.2275	-0.134
B11O	-0.188	0.222	0.7885	-0.0543	0.0919	-0.0727
B12O	-0.0294	-0.024	0.878	-0.086	0.2133	-0.098
B13O	-0.218	0.4799	0.3229	0.1902	0.0232	-0.0432
B14O	-0.1334	0.0442	0.6461	0.4537	0.1862	0.2069
B15O	0.1183	0.8341	0.0286	0.0449	0.1419	0.0594
B16O	0.1447	0.0627	-0.0482	0.8889	0.008	0.0373
B17O	0.2095	0.7732	0.0396	-0.111	-0.0658	0.1121
B18O	0.0793	0.5106	0.5816	0.242	-0.2956	-0.0101
Explained Variance	3.1573	2.7243	2.6924	1.5859	1.5756	1.2752
% of Total Variance	24.27	20.94	20.69	12.19	12.11	9.80

Dimension C:

Indicator	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
C1O	0.6849	0.4518	0.2199	0.1351	0.137	0.0619
C2O	0.3179	0.3796	0.7368	0.0586	-0.113	0.1091
C3O	0.5447	0.0574	0.6061	0.0879	0.0474	0.0407
C4O	0.1459	0.1178	0.8264	-0.1242	0.0731	-0.0917
C5O	0.2286	0.724	0.1779	-0.049	0.152	0.0516
C6O	0.7245	0.4157	0.2002	0.2624	0.1229	0.0578
C7O	0.8591	0.0778	0.1576	-0.0145	0.0683	-0.0311
C8O	-0.1925	0.0836	0.4343	0.2205	0.6612	-0.1693
C9O	0.1935	0.3147	0.5086	0.1679	-0.1511	0.3287
C10O	0.5511	0.0121	0.6597	-0.162	0.2	-0.1164
C11O	0.2212	-0.1285	0.0389	-0.189	0.6118	-0.0614
C12O	-0.0467	0.2895	0.5796	0.1302	-0.058	-0.3445
C13O	0.1409	0.8243	0.158	-0.0852	0.111	0.0873
C14O	-0.0434	-0.332	0.0277	0.8333	0.0809	-0.2593
C15O	-0.073	-0.0668	-0.0275	0.0541	0.0777	0.8902
C16O	0.318	0.7383	0.2038	-0.0097	-0.0943	-0.1285
C17O	-0.0086	0.5582	0.1616	0.0577	0.1287	-0.5392
C18O	-0.7107	-0.0398	0.1084	0.3704	-0.1295	-0.0333
C19O	0.082	0.1861	-0.2376	-0.0539	0.7796	0.2161
C20O	0.7874	0.1677	0.3642	-0.0248	-0.1977	-0.1224
C21O	0.0059	-0.8308	-0.0566	0.0398	0.1437	0.1885
C22O	0.0195	0.1309	-0.0268	0.8165	-0.2017	0.2956
Explained Variance	3.8932	3.6975	3.2569	1.789	1.735	1.6657
% of Total Variance	24.28	23.06	20.31	11.16	10.82	10.39

Dimension D:

Indicator	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
D1O	-0.148	0.2426	0.259	0.778	0.0843	0.1339
D2O	0.0901	0.0841	-0.5016	0.6948	-0.0316	0.0001
D3O	-0.342	0.7411	0.1605	0.1969	-0.0553	-0.0429
D4O	0.3696	0.6763	0.2475	0.1872	-0.385	-0.1294
D5O	0.2784	-0.3525	0.2733	0.6854	-0.0591	0.0738
D6O	-0.3515	0.0285	-0.0802	0.0806	0.8148	-0.2049
D7O	-0.0533	0.2429	0.8522	0.0821	-0.1077	-0.039
D8O	0.3507	-0.0571	-0.7108	-0.1182	-0.0553	0.175
D9O	0.4531	0.0644	0.0449	-0.0513	0.7315	0.2429
D10O	0.0821	0.7576	0.0547	-0.1397	0.2378	0.0782
D11O	-0.4761	0.201	0.0441	0.0928	0.1217	0.7163
D12O	0.8032	0.2104	-0.1695	0.2153	0.0212	-0.1889
D13O	-0.7767	0.1702	0.1968	0.0431	0.0658	-0.0236
D14O	-0.0857	0.11	0.1743	-0.0799	0.0673	-0.7708
Explained Variance	2.3049	1.9637	1.8231	1.7442	1.4572	1.3247
% of Total Variance	21.71	18.49	17.17	16.43	13.72	12.48

Table A15: Factor Loadings and Explained Variance of the Subjective SCCI

Dimension A:

Indicator	Factor 1	Factor 2
A1S	0.3667	0.9303
A2S	0.8857	0.35
A3S	0.8865	0.3485
Explained Variance	1.7048	1.1094
% of Total Variance	60.58	39.42

Dimension B:

Indicator	Factor 1	Factor 2	Factor 3
B1S	0.9009	0.0002	-0.1787
B2S	0.8681	0.2493	0.1171
B3S	0.5384	0.488	-0.37
B4S	-0.0507	-0.0698	0.9694
B5S	-0.0399	0.9551	-0.1274
B6S	0.4157	0.8496	0.0414
Explained Variance	2.0321	1.9392	1.1402
% of Total Variance	39.75	37.94	0.2231

Dimension C:

Indicator	Factor 1	Factor 2	Factor 3
C1S	0.4496	0.7478	0.3764
C2S	0.1765	0.9559	0.1183
C3S	0.9269	0.2119	0.2378
C4S	0.881	0.2987	0.2928
C5S	0.3201	0.2198	0.9183
Explained Variance	1.9711	1.6554	1.1412
% of Total Variance	41.34	34.72	23.94

Dimension D:

Indicator	Factor 1
C1S	0.8986
C2S	0.8535
C3S	0.9366
C4S	0.8844
C5S	0.8126
Explained Variance	3.8556
% of Total Variance	100

Table A16: Intensities of Importance and CR of the SCCIs Dimensions

	Criteria			
Criteria	A	B	C	D
A	1.00	0.50	0.50	2.00
B	2.00	1.00	0.50	2.00
C	2.00	2.00	1.00	2.00
D	0.50	0.50	0.50	1.00
CR	0.045			

Table A17: Results of the Objective SCCI – Dimensions

Dimension A:

#	City	Score	#	City	Score
1	Paris	0.8457	34	Warsaw	0.0244
2	Amsterdam	0.6777	35	Nicosia	0.0206
3	Rotterdam	0.5985	36	Helsinki	0.0154
4	Bratislava	0.5586	37	Newcastle	-0.0596
5	Brussels	0.5198	38	Turin	-0.0742
6	Verona	0.5121	39	Oviedo	-0.1333
7	Antwerp	0.4642	40	Rostock	-0.134
8	Lisbon	0.4049	41	Bordeaux	-0.1434
9	Dublin	0.3684	42	Rennes	-0.1617
10	London	0.367	43	Kraków	-0.1634
11	Barcelona	0.3495	44	Belfast	-0.172
12	Leipzig	0.325	45	Aalborg	-0.176
13	Madrid	0.2745	46	Gdańsk	-0.1872
14	Ljubljana	0.273	47	Glasgow	-0.2291
15	Rome	0.2521	48	Málaga	-0.2375
16	Liège	0.2387	49	Groningen	-0.2452
17	Marseille	0.2231	50	Budapest	-0.2562
18	Bologna	0.2167	51	Zagreb	-0.2691
19	Luxembourg	0.2073	52	Ostrava	-0.271
20	Naples	0.1919	53	Malmö	-0.2925
21	Prague	0.1881	54	Valletta	-0.3404
22	Munich	0.1624	55	Vilnius	-0.3464
23	Hamburg	0.1582	56	Tallinn	-0.3848
24	Berlin	0.1548	57	Miskolc	-0.4021
25	Manchester	0.1246	58	Košice	-0.4436
26	Copenhagen	0.1219	59	Sofia	-0.4786
27	Strasbourg	0.1212	60	Riga	-0.5225
28	Vienna	0.1172	61	Białystok	-0.6294
29	Stockholm	0.1127	62	Palermo	-0.6514
30	Athens	0.1106	63	Burgas	-0.6921
31	Lille	0.105	64	Bucharest	-0.7128
32	Cardiff	0.0796	65	Cluj-Napoca	-0.7404
33	Graz	0.0646			

Dimension B:

#	City	Score	#	City	Score
1	Prague	0.8584	34	Barcelona	-0.0367
2	Stockholm	0.7427	35	Groningen	-0.0398
3	Luxembourg	0.6442	36	Madrid	-0.043
4	Munich	0.6121	37	Budapest	-0.0551
5	Vienna	0.5382	38	Kraków	-0.107
6	Rostock	0.5251	39	Valletta	-0.1412
7	Bratislava	0.5231	40	Athens	-0.1813
8	Graz	0.5075	41	Lille	-0.2186
9	Amsterdam	0.5057	42	Oviedo	-0.2289
10	Ljubljana	0.4929	43	Warsaw	-0.2295
11	Helsinki	0.4629	44	Newcastle	-0.2315
12	Bordeaux	0.45	45	Vilnius	-0.2324
13	Nicosia	0.4105	46	Málaga	-0.2432
14	Copenhagen	0.3815	47	Ostrava	-0.2773
15	Malmö	0.3798	48	Riga	-0.2876
16	Hamburg	0.3776	49	Turin	-0.291
17	Brussels	0.3531	50	Zagreb	-0.2957
18	Paris	0.3489	51	Cardiff	-0.3044
19	Antwerp	0.3278	52	Rome	-0.318
20	Berlin	0.3219	53	Sofia	-0.3382
21	Dublin	0.2339	54	Manchester	-0.3669
22	Rennes	0.2314	55	Glasgow	-0.3807
23	Verona	0.2084	56	Gdańsk	-0.4206
24	Aalborg	0.2036	57	Belfast	-0.447
25	Marseille	0.1794	58	Bucharest	-0.4879
26	London	0.1767	59	Košice	-0.5499
27	Leipzig	0.1637	60	Burgas	-0.5671
28	Strasbourg	0.1492	61	Białystok	-0.6072
29	Tallinn	0.1158	62	Miskolc	-0.8946
30	Bologna	0.0889	63	Naples	-0.9209
31	Lisbon	0.0855	64	Cluj-Napoca	-0.9221
32	Rotterdam	0.0112	65	Palermo	-0.9261
33	Liège	-0.0198			

Dimension C:

#	City	Score	#	City	Score
1	Copenhagen	0.7962	34	Tallinn	-0.0559
2	Munich	0.7956	35	Manchester	-0.0585
3	Graz	0.7852	36	Sofia	-0.0588
4	Vienna	0.7788	37	Ljubljana	-0.0616
5	Helsinki	0.7561	38	Bologna	-0.1116
6	Luxembourg	0.703	39	Strasbourg	-0.1353
7	Stockholm	0.6234	40	Brussels	-0.1399
8	Prague	0.6202	41	Barcelona	-0.1527
9	Bratislava	0.5676	42	Valletta	-0.1661
10	Hamburg	0.5127	43	Belfast	-0.177
11	Amsterdam	0.5119	44	Białystok	-0.1788
12	Paris	0.4391	45	Rennes	-0.1798
13	Warsaw	0.4192	46	Bordeaux	-0.1825
14	Berlin	0.3726	47	Gdańsk	-0.1853
15	Groningen	0.3586	48	Bucharest	-0.2045
16	Dublin	0.3529	49	Rome	-0.252
17	London	0.3035	50	Marseille	-0.2792
18	Leipzig	0.2442	51	Athens	-0.3601
19	Antwerp	0.1861	52	Lille	-0.3609
20	Malmö	0.119	53	Turin	-0.3819
21	Glasgow	0.0942	54	Oviedo	-0.3939
22	Cardiff	0.0847	55	Verona	-0.4114
23	Vilnius	0.0841	56	Zagreb	-0.4303
24	Madrid	0.0799	57	Košice	-0.4317
25	Ostrava	0.0781	58	Málaga	-0.4557
26	Rostock	0.0393	59	Cluj-Napoca	-0.4902
27	Riga	0.0377	60	Liège	-0.5014
28	Aalborg	0.0319	61	Nicosia	-0.5224
29	Rotterdam	0.0297	62	Miskolc	-0.7187
30	Kraków	0.0216	63	Naples	-0.8583
31	Lisbon	0.0092	64	Burgas	-0.8951
32	Budapest	-0.0324	65	Palermo	-0.9567
33	Newcastle	-0.0554			

Dimension D:

#	City	Score	#	City	Score
1	Miskolc	0.7461	34	Rennes	-0.0134
2	Ljubljana	0.6819	35	Oviedo	-0.0153
3	Rostock	0.6078	36	Verona	-0.0396
4	Vilnius	0.5292	37	Groningen	-0.0723
5	Leipzig	0.511	38	Paris	-0.0762
6	Bordeaux	0.5024	39	Malmö	-0.1238
7	Stockholm	0.5006	40	Burgas	-0.1265
8	Tallinn	0.4402	41	Vienna	-0.1364
9	Zagreb	0.3582	42	Copenhagen	-0.1408
10	Graz	0.3134	43	Palermo	-0.1569
11	Riga	0.3115	44	Nicosia	-0.1784
12	Košice	0.3085	45	Newcastle	-0.1826
13	Berlin	0.3027	46	London	-0.1916
14	Hamburg	0.2855	47	Liège	-0.2111
15	Gdańsk	0.2637	48	Madrid	-0.2608
16	Prague	0.2511	49	Valletta	-0.27
17	Helsinki	0.2396	50	Lille	-0.2714
18	Sofia	0.2131	51	Brussels	-0.2804
19	Belfast	0.2014	52	Rome	-0.2809
20	Białystok	0.1947	53	Antwerp	-0.2933
21	Cluj-Napoca	0.1723	54	Málaga	-0.3171
22	Warsaw	0.1667	55	Amsterdam	-0.3233
23	Bratislava	0.1323	56	Dublin	-0.3286
24	Cardiff	0.1305	57	Manchester	-0.331
25	Aalborg	0.1282	58	Rotterdam	-0.399
26	Budapest	0.1228	59	Athens	-0.4165
27	Munich	0.1175	60	Bologna	-0.4756
28	Luxembourg	0.1094	61	Lisbon	-0.4975
29	Glasgow	0.0727	62	Turin	-0.5123
30	Ostrava	0.0493	63	Bucharest	-0.5226
31	Strasbourg	0.011	64	Naples	-0.5552
32	Marseille	0.0053	65	Barcelona	-0.9811
33	Kraków	0.0012			

Table A18: Results of the Subjective SCCI – Dimensions

Dimension A:

#	City	Score	#	City	Score
1	Vienna	1.6461	34	Gdańsk	0.1131
2	Munich	1.151	35	Zagreb	0.1118
3	Rotterdam	1.0464	36	Dublin	0.0594
4	Aalborg	1.036	37	Warsaw	-0.0081
5	Graz	0.9928	38	Tallinn	-0.0421
6	Rostock	0.9902	39	Barcelona	-0.0706
7	Helsinki	0.988	40	Lille	-0.0728
8	Luxembourg	0.975	41	Paris	-0.1811
9	Groningen	0.9661	42	Budapest	-0.205
10	Białystok	0.9528	43	Vilnius	-0.264
11	Malmö	0.8768	44	Cluj-Napoca	-0.2787
12	Rennes	0.864	45	Málaga	-0.2871
13	Stockholm	0.8154	46	Miskolc	-0.2983
14	Strasbourg	0.7523	47	Brussels	-0.3188
15	Bordeaux	0.7328	48	Bologna	-0.4411
16	Ljubljana	0.6845	49	Turin	-0.4506
17	Burgas	0.6103	50	Riga	-0.4842
18	Cardiff	0.5888	51	Košice	-0.6548
19	Hamburg	0.5796	52	Madrid	-0.7206
20	Copenhagen	0.5777	53	Verona	-0.7512
21	Amsterdam	0.5669	54	Liège	-0.8597
22	London	0.5606	55	Marseille	-1.0991
23	Leipzig	0.5522	56	Bratislava	-1.129
24	Manchester	0.4851	57	Lisbon	-1.2636
25	Oviedo	0.4424	58	Sofia	-1.3173
26	Prague	0.4395	59	Valletta	-1.3374
27	Belfast	0.4336	60	Athens	-1.3672
28	Newcastle	0.4296	61	Nicosia	-1.4047
29	Glasgow	0.3994	62	Bucharest	-1.6995
30	Ostrava	0.3478	63	Naples	-2.1396
31	Antwerp	0.3105	64	Palermo	-2.1486
32	Kraków	0.2617	65	Rome	-2.1983
33	Berlin	0.1527			

Dimension B:

#	City	Score	#	City	Score
1	Aalborg	1.3086	34	Dublin	0.1141
2	Groningen	1.1687	35	Prague	0.1138
3	Belfast	0.9254	36	Malmö	0.1123
4	Glasgow	0.8223	37	Strasbourg	0.107
5	Graz	0.768	38	Tallinn	0.0684
6	Cardiff	0.7452	39	Berlin	-0.069
7	Newcastle	0.7092	40	London	-0.094
8	Leipzig	0.6471	41	Verona	-0.1471
9	Vienna	0.5505	42	Košice	-0.1491
10	Rotterdam	0.4599	43	Barcelona	-0.2185
11	Cluj-Napoca	0.4446	44	Lille	-0.2959
12	Copenhagen	0.4391	45	Valletta	-0.3019
13	Oviedo	0.4359	46	Budapest	-0.3319
14	Málaga	0.4301	47	Paris	-0.3692
15	Zagreb	0.3922	48	Liège	-0.4423
16	Munich	0.382	49	Madrid	-0.4553
17	Manchester	0.3747	50	Nicosia	-0.457
18	Amsterdam	0.3718	51	Warsaw	-0.4807
19	Luxembourg	0.3593	52	Brussels	-0.4844
20	Antwerp	0.3392	53	Bologna	-0.492
21	Hamburg	0.324	54	Turin	-0.5404
22	Rennes	0.3138	55	Miskolc	-0.5523
23	Ljubljana	0.3077	56	Lisbon	-0.5922
24	Stockholm	0.3049	57	Bucharest	-0.6291
25	Burgas	0.2918	58	Riga	-0.6503
26	Vilnius	0.2704	59	Marseille	-0.772
27	Gdańsk	0.2274	60	Bratislava	-0.773
28	Helsinki	0.2187	61	Sofia	-0.8757
29	Rostock	0.1882	62	Naples	-1.1635
30	Ostrava	0.1876	63	Rome	-1.42
31	Bordeaux	0.1795	64	Palermo	-1.4628
32	Białystok	0.1719	65	Athens	-1.4714
33	Kraków	0.1158			

Dimension C:

#	City	Score	#	City	Score
1	Aalborg	1.591	34	Paris	-0.0058
2	Graz	1.3264	35	Oviedo	-0.0235
3	Munich	1.2721	36	Riga	-0.0332
4	Luxembourg	1.2146	37	Brussels	-0.1167
5	Copenhagen	1.1841	38	Białystok	-0.16
6	Stockholm	1.0357	39	Kraków	-0.1772
7	Helsinki	0.9386	40	Lille	-0.1968
8	Cardiff	0.926	41	Berlin	-0.2024
9	Vienna	0.8784	42	Dublin	-0.2062
10	Cluj-Napoca	0.8151	43	Warsaw	-0.269
11	Hamburg	0.7724	44	Liège	-0.2952
12	Antwerp	0.7579	45	Verona	-0.3356
13	Belfast	0.721	46	Ostrava	-0.3711
14	Newcastle	0.7097	47	Budapest	-0.3713
15	Valletta	0.696	48	Bratislava	-0.5199
16	Glasgow	0.6527	49	Barcelona	-0.5545
17	Malmö	0.6505	50	Málaga	-0.5662
18	Groningen	0.6395	51	Košice	-0.6181
19	Rennes	0.6026	52	Nicosia	-0.6193
20	Manchester	0.5817	53	Bologna	-0.6757
21	London	0.5763	54	Zagreb	-0.6981
22	Rostock	0.5092	55	Bucharest	-0.7406
23	Bordeaux	0.4313	56	Marseille	-0.7525
24	Leipzig	0.3927	57	Sofia	-0.8559
25	Tallinn	0.3329	58	Miskolc	-0.9094
26	Amsterdam	0.3044	59	Lisbon	-1.0399
27	Rotterdam	0.2986	60	Madrid	-1.146
28	Strasbourg	0.2928	61	Turin	-1.3559
29	Ljubljana	0.1464	62	Rome	-1.8125
30	Burgas	0.0951	63	Naples	-1.9346
31	Gdańsk	0.0875	64	Athens	-1.9375
32	Vilnius	0.0742	65	Palermo	-2.0122
33	Prague	0.0055			

Dimension D:

#	City	Score	#	City	Score
1	Aalborg	1.1933	34	Riga	0.0899
2	Vienna	1.181	35	Lille	-0.0618
3	Malmö	1.1556	36	Berlin	-0.1018
4	Białystok	1.1277	37	Burgas	-0.1127
5	Luxembourg	1.1059	38	Verona	-0.156
6	Groningen	1.0688	39	Ostrava	-0.2231
7	Newcastle	1.0652	40	Miskolc	-0.231
8	Munich	1.0531	41	Cluj-Napoca	-0.242
9	Cardiff	1.0435	42	Košice	-0.2707
10	Rostock	1.0393	43	Turin	-0.2991
11	Belfast	0.9403	44	Bologna	-0.3137
12	Ljubljana	0.8934	45	Prague	-0.3864
13	Rennes	0.8553	46	Warsaw	-0.3936
14	Glasgow	0.8424	47	Nicosia	-0.4343
15	Stockholm	0.8278	48	Liège	-0.4768
16	Helsinki	0.7552	49	Brussels	-0.5749
17	Oviedo	0.7335	50	Málaga	-0.6043
18	Leipzig	0.6796	51	Budapest	-0.6566
19	Bordeaux	0.6584	52	Kraków	-0.7184
20	Manchester	0.6411	53	Barcelona	-0.7891
21	Hamburg	0.6409	54	Paris	-0.8403
22	Vilnius	0.5719	55	Lisbon	-0.9332
23	Copenhagen	0.538	56	Bratislava	-1.0458
24	Tallinn	0.5205	57	Valletta	-1.0615
25	Dublin	0.4305	58	Marseille	-1.0814
26	Graz	0.4258	59	Madrid	-1.3443
27	London	0.3765	60	Sofia	-1.4174
28	Strasbourg	0.3579	61	Bucharest	-1.5014
29	Antwerp	0.2851	62	Rome	-1.5246
30	Amsterdam	0.2341	63	Naples	-1.7743
31	Rotterdam	0.2189	64	Palermo	-2.1195
32	Gdańsk	0.1872	65	Athens	-2.1527
33	Zagreb	0.1048			

Table A19: Results of the UA and SA of the Objective SCCI

	Scenario														
City	1	2	3	4	5	6	7	8	9	10	11	12	Min	Max	Ø
Prague	1	2	2	2	1	4	1	2	3	4	2	4	1	4	2.33
Stockholm	2	3	1	1	5	1	2	1	1	2	1	2	1	5	1.83
Munich	3	1	4	4	4	2	6	3	2	1	5	1	1	6	3
Luxembourg	4	5	6	6	6	11	5	11	5	5	6	5	4	11	6.25
Graz	5	4	5	5	2	5	4	4	4	3	4	3	2	5	4
Bratislava	6	6	3	3	3	7	3	5	6	6	3	6	3	7	4.75
Helsinki	7	8	8	9	8	3	9	7	7	8	8	8	3	9	7.5
Vienna	8	7	12	11	7	6	12	12	8	7	12	7	6	12	9.08
Amsterdam	9	9	10	8	10	12	11	13	10	9	10	9	8	13	10
Paris	10	12	7	12	9	10	7	9	11	12	7	12	7	12	9.83
Copenhagen	11	11	14	15	11	9	16	15	9	11	14	11	9	16	12.25
Hamburg	12	10	11	7	12	8	10	6	12	10	11	10	6	12	9.92
Berlin	13	13	15	14	13	13	13	10	13	13	15	13	10	15	13.17
Leipzig	14	14	13	13	15	15	14	14	14	14	13	14	13	15	13.92
Ljubljana	15	15	9	10	14	14	8	8	15	15	9	15	8	15	12.25
Dublin	16	18	20	19	19	17	23	22	16	18	18	18	16	23	18.67
Rostock	17	16	16	16	16	18	15	16	17	16	16	16	15	18	16.25
Antwerp	18	17	17	18	17	16	18	18	18	17	17	17	16	18	17.33
London	19	19	18	20	18	20	19	21	19	19	20	19	18	21	19.25
Warsaw	20	20	22	21	20	19	20	20	20	20	21	20	19	22	20.25
Brussels	21	25	21	23	22	24	21	24	21	25	22	25	21	25	22.83
Bordeaux	22	21	19	17	21	21	17	17	22	21	19	21	17	22	19.83
Rotterdam	23	22	24	25	24	26	27	25	24	24	23	24	22	27	24.25
Malmö	24	24	29	27	27	23	31	27	23	22	26	22	22	31	25.42
Groningen	25	23	33	29	25	25	32	32	25	23	33	23	23	33	27.33
Aalborg	26	29	25	31	23	27	26	26	28	30	27	30	23	31	27.33
Lisbon	27	28	32	33	30	30	36	37	26	27	30	27	26	37	30.25
Madrid	28	27	30	32	31	31	33	36	27	28	31	28	27	36	30.17
Strasbourg	29	26	26	22	26	22	25	19	29	26	28	26	19	29	25.33
Tallinn	30	31	28	28	33	33	29	30	31	32	29	32	28	33	30.5
Verona	31	33	23	24	29	41	22	33	35	35	25	35	22	41	30.5
Marseille	32	30	27	26	28	28	24	23	30	29	24	29	23	32	27.5
Cardiff	33	32	34	30	37	32	34	29	32	31	34	31	29	37	32.42
Vilnius	34	38	31	34	34	35	28	31	33	38	32	38	28	38	33.83
Rennes	35	35	35	35	32	29	30	28	34	33	35	33	28	35	32.83
Bologna	36	34	39	37	36	38	38	38	36	34	39	34	34	39	36.58
Kraków	37	36	37	38	35	34	35	35	37	36	37	36	34	38	36.08
Budapest	38	37	36	36	39	36	37	34	38	37	36	37	34	39	36.75
Ostrava	39	39	40	40	38	37	40	39	40	40	41	40	37	41	39.42
Glasgow	40	40	41	39	43	42	47	42	39	39	40	39	39	47	40.92
Nicosia	41	46	38	47	40	52	39	51	41	48	38	48	38	52	44.08
Newcastle	42	41	44	41	45	40	49	41	42	41	44	41	40	49	42.58
Riga	43	43	42	44	41	39	41	40	43	43	42	43	39	44	42
Barcelona	44	44	53	53	44	46	54	53	44	42	51	42	42	54	47.5
Manchester	45	42	48	48	47	43	50	49	45	44	50	44	42	50	46.25
Rome	46	45	47	45	50	48	48	47	47	45	47	45	45	50	46.67
Sofia	47	49	50	50	46	45	46	48	46	46	46	46	45	50	47.08
Liège	48	47	43	43	48	47	43	43	48	47	45	47	43	48	45.75
Gdańsk	49	50	45	42	42	49	42	44	49	50	43	50	42	50	46.25
Belfast	50	48	46	46	54	50	52	50	50	49	49	49	46	54	49.42
Valletta	51	54	55	55	52	58	55	60	51	54	55	54	51	60	54.5
Lille	52	51	51	52	49	44	45	46	52	51	52	51	44	52	49.67
Athens	53	55	54	54	55	57	56	58	53	55	53	55	53	58	54.83
Oviedo	54	52	52	51	53	53	51	52	55	53	54	53	51	55	52.75
Zagreb	55	53	49	49	51	51	44	45	54	52	48	52	44	55	50.25
Turin	56	56	59	57	57	55	59	57	56	56	58	56	55	59	56.83
Białystok	57	57	57	58	58	54	58	55	57	57	57	57	54	58	56.83
Málaga	58	58	58	59	59	59	60	59	58	58	59	58	58	60	58.58
Košice	59	59	56	56	56	56	53	54	59	59	56	59	53	59	56.83
Bucharest	60	60	61	61	60	61	62	61	60	60	61	60	60	62	60.58
Miskolc	61	61	60	60	61	60	57	56	61	61	60	61	56	61	59.92
Cluj-Napoca	62	62	62	62	62	62	61	62	62	62	62	62	61	62	61.92
Naples	63	63	63	63	64	64	64	64	63	63	63	63	63	64	63.33
Burgas	64	64	64	64	63	63	63	63	64	64	64	64	63	64	63.67
Palermo	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65

Table A20: Results of the UA and SA of the Subjective SCCI

	Scenario														
City	1	2	3	4	5	6	7	8	9	10	11	12	Min	Max	Ø
Aalborg	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Graz	2	2	6	5	2	2	6	5	2	2	6	5	2	6	3.75
Munich	3	4	3	4	3	4	4	4	3	4	3	4	3	4	3.58
Vienna	4	5	2	3	5	5	2	3	4	5	2	3	2	5	3.58
Luxembourg	5	6	5	6	4	6	5	6	5	6	5	6	4	6	5.42
Groningen	6	3	4	2	6	3	3	2	6	3	4	2	2	6	3.67
Cardiff	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
Copenhagen	8	11	13	14	9	11	16	14	8	10	13	14	8	16	11.75
Belfast	9	10	8	10	8	9	8	10	10	11	9	10	8	11	9.33
Stockholm	10	13	9	13	10	15	10	13	9	12	8	13	8	15	11.25
Newcastle	11	8	10	8	11	8	9	8	12	8	11	8	8	12	9.33
Helsinki	12	18	11	17	13	18	12	18	11	17	10	16	10	18	14.42
Glasgow	13	9	15	9	12	10	13	11	13	9	14	9	9	15	11.42
Malmö	14	24	12	23	17	24	15	24	14	24	12	23	12	24	18.83
Rennes	15	15	16	12	14	12	14	12	16	15	16	11	11	16	14
Hamburg	16	14	17	16	16	16	17	16	15	14	17	15	14	17	15.75
Rostock	17	16	14	11	15	14	11	9	17	18	15	12	9	18	14.08
Leipzig	18	17	18	15	19	17	18	15	18	16	18	17	15	19	17.17
Manchester	19	21	20	21	18	21	19	21	19	22	20	21	18	22	20.17
Antwerp	20	12	24	20	20	13	25	20	20	13	24	20	12	25	19.25
Rotterdam	21	19	22	18	22	20	22	19	21	19	22	18	18	22	20.25
Bordeaux	22	20	23	19	21	19	21	17	22	20	23	19	17	23	20.5
Ljubljana	23	25	21	25	24	25	23	25	23	25	19	25	19	25	23.58
Cluj-Napoca	24	29	31	37	23	29	31	36	24	30	31	36	23	37	30.08
Amsterdam	25	22	27	22	27	22	27	22	25	21	26	22	21	27	24
London	26	26	28	26	25	26	28	27	26	26	28	26	25	28	26.5
Strasbourg	27	23	26	24	28	23	26	23	27	23	25	24	23	28	24.92
Białystok	28	35	19	28	29	35	20	28	28	35	21	28	19	35	27.83
Oviedo	29	27	25	27	26	27	24	26	29	27	27	27	24	29	26.75
Tallinn	30	33	30	32	31	32	30	31	30	33	29	33	29	33	31.17
Burgas	31	43	29	39	30	43	29	39	31	42	30	40	29	43	35.5
Vilnius	32	34	32	34	33	34	32	34	32	34	32	34	32	34	33.08
Gdańsk	33	37	33	38	32	36	33	38	33	39	33	38	32	39	35.25
Prague	34	28	35	29	34	28	35	29	34	28	35	29	28	35	31.5
Dublin	35	38	34	35	35	38	34	37	35	38	34	35	34	38	35.67
Ostrava	36	30	36	30	36	30	37	30	36	29	36	30	29	37	33
Berlin	37	31	38	31	40	33	38	32	37	31	38	31	31	40	34.75
Kraków	38	46	39	46	37	46	39	46	38	45	39	45	37	46	42
Zagreb	39	40	37	36	38	40	36	35	39	41	37	37	35	41	37.92
Valletta	40	42	52	50	39	39	50	48	41	43	53	50	39	53	45.58
Lille	41	32	40	33	41	31	40	33	40	32	40	32	31	41	36.25
Málaga	42	45	41	44	42	45	41	43	44	47	41	46	41	47	43.42
Paris	43	39	45	42	44	41	45	42	42	37	44	41	37	45	42.08
Riga	44	49	42	48	43	49	42	49	43	49	42	49	42	49	45.75
Warsaw	45	52	43	52	46	53	43	52	45	52	43	52	43	53	48.17
Verona	46	41	44	41	45	42	44	41	47	40	45	42	40	47	43.17
Brussels	47	36	46	40	47	37	46	40	46	36	46	39	36	47	42.17
Budapest	48	54	47	53	48	54	47	53	48	53	47	53	47	54	50.42
Barcelona	49	48	48	47	49	48	48	47	49	48	49	47	47	49	48.08
Košice	50	50	49	49	51	50	49	50	50	50	48	48	48	51	49.5
Liège	51	44	53	43	50	44	53	44	51	44	52	43	43	53	47.67
Bologna	52	47	50	45	52	47	52	45	52	46	50	44	44	52	48.5
Miskolc	53	56	51	56	53	57	51	56	53	56	51	55	51	57	54
Nicosia	54	53	55	55	54	52	55	55	54	54	55	56	52	56	54.33
Bratislava	55	59	56	59	55	59	56	59	55	58	56	59	55	59	57.17
Turin	56	55	54	51	56	55	54	51	56	55	54	51	51	56	54
Marseille	57	51	58	54	57	51	58	54	57	51	57	54	51	58	54.92
Madrid	58	57	57	57	58	56	57	57	59	57	58	57	56	59	57.33
Lisbon	59	58	59	58	59	58	59	58	58	59	59	58	58	59	58.5
Bucharest	60	61	61	61	60	61	61	61	60	61	61	61	60	61	60.75
Sofia	61	60	60	60	61	60	60	60	61	60	60	60	60	61	60.25
Rome	62	62	63	62	64	62	64	62	62	62	62	62	62	64	62.42
Naples	63	63	64	63	62	63	63	63	63	63	64	63	62	64	63.08
Athens	64	64	62	64	63	64	62	64	64	64	63	64	62	64	63.5
Palermo	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65

Table A21: PCCs Between the UA and SA Scenarios of the Objective SCCI

	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.9978	1										
3	0.9879	0.9864	1									
4	0.9845	0.9883	0.9965	1								
5	0.9959	0.9942	0.9905	0.9872	1							
6	0.9871	0.9912	0.9785	0.9818	0.9888	1						
7	0.9735	0.9728	0.9943	0.9911	0.9855	0.9714	1					
8	0.9726	0.9781	0.9882	0.9927	0.982	0.9869	0.9911	1				
9	0.9993	0.9978	0.9866	0.9841	0.9948	0.9889	0.9719	0.9735	1			
10	0.9963	0.9993	0.9845	0.9875	0.9923	0.9925	0.9706	0.9786	0.9977	1		
11	0.9887	0.9875	0.9992	0.9961	0.9905	0.9801	0.9925	0.988	0.9888	0.9869	1	
12	0.9848	0.9889	0.9953	0.9992	0.9867	0.9829	0.9889	0.9922	0.9857	0.9893	0.9964	1

Table A22: PCCs Between the UA and SA Scenarios of the Subjective SCCI

	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.9775	1										
3	0.9939	0.9707	1									
4	0.9786	0.9955	0.9819	1								
5	0.9997	0.9764	0.9931	0.9772	1							
6	0.978	0.9997	0.9707	0.9949	0.9775	1						
7	0.9939	0.9703	0.9998	0.9813	0.9937	0.9708	1					
8	0.9792	0.9955	0.9823	0.9998	0.9784	0.9953	0.9821	1				
9	0.9999	0.9788	0.9938	0.9797	0.9992	0.979	0.9935	0.9801	1			
10	0.9766	0.9999	0.9699	0.9954	0.9752	0.9993	0.9692	0.9952	0.9781	1		
11	0.9938	0.9717	0.9999	0.9827	0.9928	0.9715	0.9995	0.9829	0.9939	0.9711	1	
12	0.9779	0.9954	0.9812	0.9999	0.9763	0.9946	0.9805	0.9995	0.9792	0.9955	0.9823	1

Table A23: City Population Size Classification

XXL		XL		L		M	
City	Inhabitants	City	Inhabitants	City	Inhabitants	City	Inhabitants
Paris	9,782,671	Naples	978,399	Gdańsk	461,489	Nicosia	241,000
London	8,730,803	Lille	902,970	Bratislava	419,678	Košice	239,464
Berlin	3,469,849	Turin	896,773	Tallinn	413,782	Oviedo	221,870
Madrid	3,141,991	Marseille	893,431	Strasbourg	401,308	Rennes	215,366
Rome	2,872,021	Amsterdam	810,938	Bologna	386,181	Valletta	212,885
Athens	2,641,511	Zagreb	799,999	Liège	382,852	Rostock	204,167
Bucharest	2,107,399	Kraków	761,873	Cardiff	355,727	Aalborg	203,448
Vienna	1,766,746	Palermo	678,492	Belfast	339,243	Burgas	198,725
Hamburg	1,762,791	Riga	641,007	Cluj-Napoca	321,916	Groningen	198,317
Budapest	1,757,618	Bordeaux	635,780	Malmö	302,835	Miskolc	159,554
Warsaw	1,735,442	Helsinki	620,715	Białystok	295,459	Luxembourg	111,287
Barcelona	1,604,555	Rotterdam	618,357	Ostrava	292,681		
Stockholm	1,579,896	Glasgow	602,990	Newcastle	291,359		
Munich	1,429,584	Málaga	569,130	Ljubljana	287,218		
Prague	1,267,449	Copenhagen	559,440	Graz	269,997		
Sofia	1,228,282	Leipzig	544,479	Verona	260,125		
Brussels	1,196,831	Vilnius	542,626				
		Vilnius	542,626				
		Manchester	525,254				
		Dublin	516,255				
		Antwerp	515,593				
		Lisbon	509,312				

Table A24: Summary Statistics of the Continuous Independent Variables

Independent Variable	Unit	Mean	Std. Dev.	Min	Max
Population	# inhabitants	1,075,187	1,652,190	111,287	9,782,671
GDP per Capita in PPS	PPS/inhabitant	33,000	11,837.44	11,000	75,000
Population Density	# inhabitants/km	2736.82	1929.10	161.57	8947.69
Cooling Degree Days	Weather-based technical index	128	164.52	0	685.57
Heating Degree Days	Weather-based technical index	2,546.66	776.86	541.82	4051.73

Table A25: Breusch–Pagan / Cook–Weisberg Test of Heteroscedasticity

	Dependent Variable	
	Objective SCCI	Subjective SCCI
H_0	Constant Variance	Constant Variance
$\chi^2(1)$	0.04	11.81
Prob > χ^2	0.8320	0.0006

Table A26: Shapiro–Wilk Test Statistics of Normality in Residuals

Dependent Variable	Obs.	W	V	z	Prob > z
Objective SCCI	65	0.98898	0.639	-0.97	0.83403
Subjective SCCI	65	0.96519	2.018	1.52	0.06421

Table A27: VIFs of the Independent Variables

Independent Variable	VIF
Cooling Degree Days	3.52
Heating Degree Days	3.32
Dummy Capital	1.92
Population	1.87
Population Density	1.80
GDP per Capita	1.72
Dummy New Member State	1.70

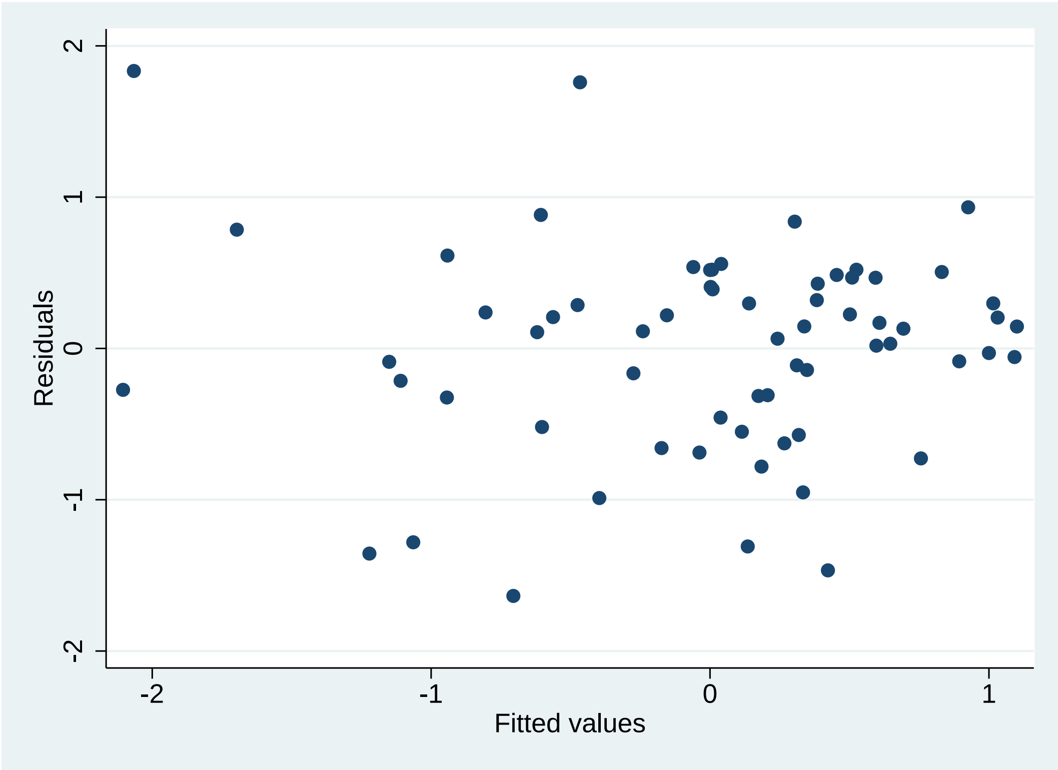

Figure A1: Residuals vs. Fitted Values of the Subjective SCCI

Table A28: Shapiro–Wilk Test Statistics of the Regression Variables

Variable	Obs.	W	V	z	Prob > z
Objective SCCI	65	0.97295	1.568	0.974	0.16508
Subjective SCCI	65	0.95274	2.74	2.182	0.01454
Population	65	0.5107	28.364	7.244	0.00000
GDP per Capita	65	0.95056	2.866	2.28	0.01130
Population Density	65	0.82316	10.251	5.04	0.00000
Cooling Degree Days	65	0.77206	13.213	5.589	0.00000
Heating Degree Days	65	0.94781	3.025	2.397	0.00826

Table A29: PCCs Between Climatic Variables and Educational Indicators

	Cooling Degree Days	Heating Degree Days
Cooling Degree Days	1	
Heating Degree Days	-0.7639	1
B5O	-0.5317	0.7007
B6O	-0.4195	0.4364
B7O	-0.1532	0.1592
C5O	-0.3211	0.331
C6O	-0.1594	0.1632
B2S	-0.5071	0.3953
B3S	-0.6355	0.6999