Conditional Regional Differences in Subjective Well-Being in Central and Eastern Europe and the Role of Social Capital

Thesis

to obtain the degree of Master of Science

in Economics and Business

from Erasmus University of Rotterdam

under the supervision of Dr. Martijn Burger

by

Igor Georgiev[a]

[a] SID: 353842; email: igeorgiev@student.eur.nl

Abstract: This paper tests empirically the conditional difference in subjective well-being between and within regions in 5 countries from Central and Eastern Europe after controlling for a set of widely-accepted socio-demographic and regional characteristics from 2002 to 2014. Further, the paper tests if the unexplained conditional difference in subjective well-being might be partially explained by social capital. Using a 3-level hierarchical model, mediation analysis and an ordered probit model, this paper concludes that on average individuals living in the big cities of the capital city regions are happier and more life satisfied as compared to individuals living in other regions and communities. This is largely explained by the macroeconomic and environmental characteristics. Although it seems that social capital mediates the effect of the place of living on SWB in the case of rural communities versus big city communities, the total mediating effect of social capital for all regions and communities is insignificant. Further, results indicate that the biggest contributors to the mediation effect are the norms and sanctions. Thus, it seems that social capital at least partly might be accounted for the conditional gap in SWB between the rural community and the big city community within CEE countries. Finally, results suggest that happiness and life satisfaction should not be used interchangeably as this might lead to false generalized conclusions. However, due to various limitations the obtained results should not be interpreted as a causal claims, but just as mere associations.

I would like to thank Dr. Burger and Dr. Stavropoulos for their patience and helpful comments throughout the writing process. Further, I would like to thank the teaching teams of seminars Multinationals and Business Networks and Topics in Organisation and Strategy, namely Dr. Burger, Dr. Stavropoulos, Caroline Witte, Efstratia Arampatzi and Dr. Bhaskarabhatla for their contribution to my empirical techniques knowledge. I would also like to thank my former classmates Albert van den Pol, Sam Hullegien and Jan Quist for the enjoyable gatherings and healthy competition throughout the academic year. Last, but not least, I would like to express my gratitude and appreciation to my family, especially my mother Irina Georgieva, my father Stanislav Georgiev, my grandmother Nadiya Druzhinina and my grandfather Prof. dr. Anatoliy Druzhinin for their endless love and support.

TABLE OF CONTENTS

1. Introduction

In recent years more and more attention is being paid to the topic of subjective well-being (SWB. By recognizing the need for *"a more inclusive, equitable and balanced approach to economic growth that promotes sustainable development, poverty eradication, happiness and the well-being of all peoples"*,[1] the United Nations General Assembly established in 2012 an international day of happiness. Hence, people are celebrating the "Happiness Day" each year on 20th March. The importance of SWB is also highlighted by the World Happiness Report. It is being published since 2012 and its aim is to study the global state of happiness and its causes (Helliwell, Layard and Sachs, 2016). As discussed in the report, politicians from leading world economies such as the United Kingdom, South Korea, the United Arab Emirates and Dubai all acknowledge the importance of SWB from a policy perspective (Helliwell, Layard and Sachs, 2015). Another colorful example of the importance of the SWB topic in public policy might be tracked to Bhutan. In 1970s, the "Gross National Happiness" (GNH) index was established, representing the government's goal to develop a sustainable economy. This commitment was reassured by the creation of the Gross National Happiness Commission, which aims to *"ensure that GNH is mainstreamed into the planning, policy making and implementation process"* (Gross National Happiness Commission, 2007 pp.1). Although the topic of SWB is becoming more and more popular in the recent decades, it is not a new phenomenon. Thus, individual's life satisfaction as the ultimate goal can be tracked back to Aristotle's *eudaimonia*. From ancient Greek it translates as "happiness", "welfare", "human flourishing" (Helliwell et al., 2015).

Meanwhile, increasing number of economists is opposing the neo-classical belief that individuals are purely self-interested actors (Coleman, 1988). For example, separate stream of economy, established in Italy and supported by the Italian economist and philosopher Antonio Genovesi, promotes efficiency, equity and public happiness. Known as Civil Economy, it challenges the general economic belief that individuals should be treated as *homines economici*, or in other words - purely self-interested. Its paradigm states that social capital, represented by trust and relationships among the individuals, is one of the key aspects of a successful economy (Becchetti et al., 2015).

[1] United Nations General Assembly Resolution 66/281, (2012, pp. 1).

SWB and social capital topics are even more important in the case of Central and Eastern Europe (CEE). First of all, CEE countries are consistently less happy as compared to their Western counterparts. This becomes evident from the results presented in the World Happiness Report 2016. On a possible scale from 0 to 10, the average mean value of happiness in CEE countries is 5.554 as compared to a score of 6.575 in Western Europe (Helliwell et al., 2016 pp. 15). Further, some countries in CEE report notoriously low levels of happiness, which cannot be explained by their economic development. For example, Bulgaria is ranked 129th in the World Happiness Report 2016, which puts it behind much weaker economies such as Ethiopia, Nepal and Sierra Leone.

On the other hand, the history of CEE has been marked by the transition from communism and planned economy to democracy and market economy (Nikolova, 2015). It has been argued that dictatorship in its various forms (totalitarianism, ordinary dictatorship and absolutism) can destroy social capital. It achieves this through attempting to establish a full control of the relationships among individuals. This in turn erodes trust and individuals' own initiatives and enhances fear and distrust (Paldam and Svendsen, 2000). Moreover, the strong effect of social capital on SWB is well-documented. Thus, very large income-equivalent values are needed to compensate for low social capital (Helliwell et al., 2009). This should be especially important concern in the case of CEE. The reasoning behind this is threefold. First, in transition countries the magnitude of the effect of social capital on SWB is comparable to the effect of GDP (Bartolini et al., 2015). This finding highlights the importance of social capital since the link between economic growth and happiness in transition economies is especially strong (for example, Nikolova, 2015; Djankov et al., 2016). Second, one might expect low levels of social capital in countries from CEE since social capital stock could have been eroded during the years under communism and planned economy. Third, when formulating new policies for SWB, the young market economies might struggle to compensate for the low social capital with income-equivalent policy tools. Thus, a particular attention should be paid to social capital when studying the SWB topic since social capital potentially might be able to explain the differences in SWB between and within countries (Helliwell et al., 2009).

Furthermore, one might expect to observe a variation in SWB and social capital between and within regions. First of all, previous academic findings conclude that there is a trend of economic

divergence between capital city regions and other regions within CEE (for example, Dogaru et al., 2015). However, there is a strong link between economic growth and happiness in transition economies (for example, Nikolova, 2015; Djankov et al., 2016). Thus, one might naturally hypothesize that SWB should be higher in the capital city regions before controlling for various macroeconomic and socio-demographic characteristics. Furthermore, Hayo (2007) finds that individuals living in rural communities have higher SWB, *ceteris paribus*. On the other hand, there is some evidence that the levels of social capital stock differ between the different types of communities and regions (for example, Onyx and Bullen, 2000; Beugelsdijk and Van Schaik, 2005; Rodríguez-Pose and von Berlepsch, 2014). First, social capital varies across Europe. Empirical evidence suggests that its level is particularly low in eastern regions of Europe (Rodríguez-Pose and von Berlepsch, 2014). Second, while in some European countries the level of social capital might be relatively homogenous, in others it substantially differs (Beugelsdijk and Van Schaik, 2005). Finally, the evidence suggests that in general social capital tends to be higher in rural areas as compared to urban areas (Onyx and Bullen, 2000). These findings justify the need to study the conditional regional differences in SWB in CEE and the role of social capital.

Hence, combining data from European Social Survey, EUROSTAT, The World Bank, AirBase and ESPON, this study aims to test empirically if there is a conditional difference in SWB between and within regions in CEE after controlling for a set of widely-accepted socio-demographic and regional characteristics. Moreover, this paper aims to test if the unexplained conditional difference in SWB between and within regions in CEE might be partially explained by social capital.

Analyses are conducted on a data sample consisting of 5 CEE countries and 41 NUTS2 regions from 2002 to 2014. These countries are Bulgaria, Czech Republic, Hungary, Poland and Slovakia. Further, NUTS2 regions are classified either as capital city regions or other regions under the classification available in Dogaru et al. (2014, pp. 211). Moreover, within each of the regions, 3 different community types are distinguished. These are: 1) a big city or the outskirts of a big city; 2) towns or small cities; 3) country villages or a farm/home in the countryside. Using a 3-level hierarchical model, mediation analysis and an ordered probit model, this paper concludes that on average individuals living in the big cities of the capital city regions are

happier and more life satisfied as compared to individuals living in other regions and communities. However, this finding is largely explained by the macroeconomic and environmental characteristics. Although it seems that social capital mediates the effect of the place of living on SWB in the case of rural communities versus big city communities, the total mediating effect of social capital for all regions and communities is insignificant. Further, results indicate that the biggest contributors to the mediation effect are the norms and sanctions. Thus, it seems that social capital at least partly might be accounted for the conditional gap in SWB between the rural community and the big city community within CEE countries. Further, according to the results from previous academic literature, all three components of social capital as suggested in Coleman (1988) have positive effect on SWB. However, while the magnitude of the effect of trust and norms and sanctions on SWB is relatively similar, the effect of social interactions on SWB is much smaller. Finally, results suggest that happiness and life satisfaction should not be used interchangeably as this might lead to false generalized conclusions. However, due to various limitations the obtained results should not be interpreted as a causal claims, but just as mere associations.

The main academic contributions of this study are as follows. First, this paper contributes to the academic literature stream, which studies SWB in CEE (for example, Easterlin, 2009; Guriev and Zhuravskaya, 2009; Djankov et al., 2016). Further, this work touches a topic, which was hardly studied before. More precisely, it studies empirically the conditional difference in subjective well-being between and within regions in CEE after controlling for a set of widely-accepted socio-demographic and regional characteristics. Moreover, this work contributes to the academic literature stream, which studies the effect of social capital on SWB in Europe (for example, Rodríguez-Pose and von Berlepsch, 2014; Bartolini et al., 2015). More precisely, this paper tests if the unexplained conditional difference in SWB between and within regions in CEE might be partially explained by social capital, which was hardly researched before. By doing so, this paper contributes to the understanding of the relationship between social capital and SWB in CEE.

The social relevance relates to the fact that SWB and social capital are crucial factors for healthy and active society. People are facing various aspects of social capital on a daily basis. Thus, feeling safety after dark, having trust in national institutions, participating in various non-profit

organizations, and, in general, being socially active would increase people's self-esteem and quality of life. Moreover, understanding what affects SWB will enable making the society in CEE happier and more life satisfied. Not only that, but a better understanding of the causes for a difference in SWB between and within regions in CEE will also make it possible to achieve SWB cohesion across regions.

The rest of the paper is organized as follows. Section II presents an overview of the previous literature and builds the hypotheses. Section III describes the data and methodology. Section IV presents the results. Robustness checks are performed in section V. Final discussions and conclusion are presented in section VI.

2. Literature Review

2.1. Subjective Well-Being and Its Part in Applied Economics

Subjective well-being (SWB) data is based on questionnaires regarding such perceptual life aspects as life purpose, various types of emotions, overall life evaluation, *et cetera* (Nikolova, 2015). As further discussed by the author, these aspects have individual, economic and institutional determinants.

Further, a distinction between two facets of SWB should be made. As discussed in Wiest et al. (2011), SWB has emotional and cognitive aspects. While the former is more correlated with positive and negative experiences and might fluctuate on a daily basis, the latter is assumed to be more stable. Further, the emotional aspect of SWB, also known as affective, can be traced back to the hedonic theories of happiness. On the other hand, the cognitive aspect of SWB might be linked to the subjective evaluation theories of well-being (Schimmack et al. 2008). This paper focuses on the more stable in time cognitive aspect of SWB.

Although some might question the feasibility of using subjective indicators as a prime source of empirical data, there is a wide range of academic evidence proving the opposite. Individuals' rating of their own happiness has been statistically linked to various personal characteristics, the surrounding environment as well as objective indicators. Responses collected from asking SWB questions could predict future behavior. Moreover, these responses have been linked to individuals' neurological functioning, health outcomes and the way individuals spend their time (Kahneman and Krueger, 2006; Nikolova, 2015). Furthermore, as pointed out in Alesina, Di

Tella and MacCulloch (2004), referring to Sutton and Davidson (1997), the happiness data is correlated with the *"electroencephalogram measures of prefrontal brain activity"* which is in fact "responsible" for happiness.

As discussed in Veenhoven (2000), a potential issue, which might occur when dealing with SWB literature, is the confusion caused by the meanings of terms "quality of life", "wellbeing" and "happiness". In many fields the term "wellbeing" is used interchangeably with the term "quality of life" and it measures the overall living conditions. It is used to distinguish between the pure economic "welfare", and incorporates in itself also subjective measures. Nevertheless, the summing of all wellbeing factors might cause unreliable results (Veenhoven, 2000).[2] As further discussed in Veenhoven (2000), the term "happiness" is most widely used in social indicators research. It defines a more broad evaluation of life: the general subjective evaluation of life. An individual might consider himself happy even if the objective characteristics of his life are low. On the other hand, the opposite scenario might be also true - an individual might consider himself unhappy even if the objective characteristics of his life are high. Finally, the author concludes that happiness is the best available summary measure of SWB.

Two suggestive definitions of happiness are presented in Veenhoven (2000) and Veenhoven (2004):

> *"…happiness is both a merit in itself, and indicative of good life-chances. Subjective happiness implies two things: Firstly that the minimal conditions for humans thriving are apparently met, secondly that the fit between opportunities and capacities must be sufficient. Hence happiness says more about the quality of life-chances that the sum-scores do." (Veenhoven, 2000 pp. 25)*
>
> *"Happiness can be defined as the 'overall enjoyment of your life as-a-whole." (Veenhoven, 2004 pp.6)*

However, one major criticism of using happiness as a measurement of the overall SWB is the potential inaccurate translation of the word "happiness" across countries. Further, the understanding of "happiness" by individuals might differ across cultures and religions (Veenhoven, 2000). Thus, in some surveys the word "satisfied" is used instead of the word "happy" due to raised concerns regarding the translation of the meaning of "happy" (Alesina et

[2] Further elaboration of this potential issue might be found in Veenhoven (2000).

al., 2004). However, Veenhoven (2004) discusses these potential pitfalls and concludes that using happiness is a valid measure of overall quality of life. The author concludes that happiness levels can be compared across individuals and countries. Furthermore, the answers recorded from the questions asking about happiness and life satisfaction are highly correlated (Alesina et al., 2004).

SWB data has also been challenged in terms of its applicability in social policies. While most politicians favor objective indicators for tailoring social policies, they are harder to aggregate accurately. Assigning weights to the various objective indicators, while aggregating them, might lead to false conclusions since different individuals might praise to a different extent certain objective measures such as income or the availability of transport links between cities. On the other hand, the overall assessment of life is much less problematic when using subjective indicators. The reasoning behind this is as follows. Each individual will assess and assign weights for each of the indicators affecting his happiness while evaluating his own happiness level (Veenhoven, 2002). Thus, understanding which indicators increase individuals' SWB will not only increase traditional welfare, but will also increase the amount of happy years lived by individuals. Furthermore, several papers argue that SWB might influence the voting behavior of individuals. Dolan, Metcalfe and Powdthavee (2008) suggest that voting intentions might be affected by SWB. Further, Liberini, Redoano and Proto (2014), argue that there is a link between SWB and individuals voting behavior even after controlling for traditional welfare indicators. Finally, Ward (2015) concludes that the variation in countrywide voting behavior is better explained with SWB as compared to the standard macroeconomic indicators. Not only this, but many academic papers suggest that SWB influences objective outcomes which are of prime interest for policymakers and the society – health, income levels, productivity, unemployment levels and social behavior (De Neve, Diener and Xuereb, 2013). These findings highlight the importance of SWB as a tool for policy formulation.

Dolan, Peasgood and White (2008) provide an extensive literature review on the determinants of SWB. Some of the main determinants of SWB are income, personal characteristics, socially developed characteristics, the work and activities which an individual undertakes, attitudes and believes, relationships, and finally, wider political, economic and social environment.

2.2.Subjective Well-Being in Central and Eastern Europe

The gradual economic convergence between Central and Eastern Europe (CEE) and Western Europe has been well documented (for example, Fischer and Stirbock, 2004; Paas and Schlitte, 2006; Ezcurra, Pascual and Rapun, 2007; Dogaru, Burger, Karreman and van Oort, 2015). Further, there is a strong link between economic growth and happiness in transition economies (Easterlin, 2009; Nikolova, 2015; Djankov, Nikolova and Zilinsky, 2016). However, findings presented in the World Happiness Report 2016 show that the average mean value of happiness in CEE countries is 5.554, on a possible scale from 0 to 10. On the other hand, the average mean value of happiness in Western Europe is 6.575, which is more than one point higher as compared to CEE (Helliwell et al., 2016 pp. 15). Further, some of the CEE countries report notoriously low levels of happiness. For example, out of 157 countries, Bulgaria is ranked 129th. As reported in UNCTAD (United Nations Conference on Trade and Development), the GDP per capita in current US$ for 2014 in Bulgaria was 7,876.41$. For comparison, Ethiopia, with GDP per capita of 553.20$, Nepal, with GDP per capita of 691.71$, and Sierra Leone, with GDP per capita of 774.70$ all report higher happiness levels than Bulgaria (Helliwell et al., 2016). Although ranked relatively higher, the other CEE countries are not performing much better in the happiness ranking. For example, Hungary is ranked 91st, Romania is ranked 71st, while Poland is ranked 57th. On the other hand, Czech Republic is ranked 27th, making it the most happy country in CEE (Helliwell et al., 2016). Observing the very low rankings of some CEE countries and the large differences in happiness between CEE countries, one might reasonably expect that there is something else than the level of economic development, which affects SWB in CEE.

Furthermore, the conditional difference in country-level subjective well-being between CEE and the West has been acknowledged in the academic literature (for example, Guriev and Zhuravskaya, 2009; Nikolova, 2015; Djankov et al., 2016). Although this conditional difference in SWB has been documented, few papers study the possible causes of this phenomenon (Nikolova, 2015). The macroeconomic volatility, raising inequality, the deterioration in public goods provision and mismatch of human capital are all potential candidates in explaining the SWB gap (Guriev and Zhuravskaya, 2009). Moreover, Nikolova (2015) proposes the rule of law as a potential explanation for the conditional difference in SWB between developed and countries in transition. Bartolini et al. (2015) find that social trust is a powerful explanatory tool

for SWB in transition countries. Finally, Djankov, et al. (2016), using SWB data from Pew, Eurobarometer, European Values Study and Life in Transition Survey, find no evidence that the happiness gap between transition countries and developed countries is closing. As it is discussed, although the happiness level has been rising after the 1990s, it stagnated after early 2000s. Authors suggest that high corruption levels and poor government quality might explain the happiness gap not only at the country-level but also at the individual-level.

On the other hand, even less has been done in order to explore if there is a persistent conditional difference in SWB levels between and within the various types of regions in CEE. Hayo (2007) analyzes life satisfaction determinant in seven CEE countries. The author finds significant difference in inter-country life satisfaction levels, which cannot be fully explained by the widely-used socio-demographic characteristics. Further, the results from the ordered logit model suggest that increasing community size has statistically significant negative effect on life satisfaction, *ceteris paribus*. Rodriguez-Pose and Maslauskaite (2012) study life satisfaction determinants in 10 CEE countries. The authors discuss that the pace of life satisfaction convergence in CEE does not follow the same speed as the economic convergence between those countries. As further argued, the main factors which cause this are the macroeconomic and institutional differences between the countries in the data sample. The paper concludes that the institutional factors with the highest influence on life satisfaction are corruption, government spending and decentralization. The regression results for 2008 show that living in a big city has a negative impact on happiness, *ceteris paribus*. This effect remains statistically significant also after the introduction of country fixed effects into the model. On the other hand, living in a remote area has a positive effect on happiness, *ceteris paribus*. However, this effect becomes statistically insignificant once the country fixed effects are introduced into the model. However, all of the previous literature touching the topic of SWB in CEE, at best has concentrated on studying SWB on an inter-country level. Thus, the topic of conditional difference in SWB levels between and within the various types of regions in CEE countries remains virtually unexplored. Potentially the only two notable exceptions are Andrén and Martinsson (2006) and Lelkes (2006). The first article studies life satisfaction determinants in Romania for 2001. The authors conclude that life satisfaction determinants are similar to those observed in Western countries. Further, the regression results suggest that different residence areas have different negative effect in life satisfaction, *ceteris paribus*. This finding remains statistically significant after controlling for

various socio-demographic characteristics. Regarding Lelkes (2006), the article studies life satisfaction determinants in Hungary. The regression results for 1998 suggest that living in Budapest has a negative and statistically significant effect on life satisfaction, *ceteris paribus*.

Nevertheless, there is some worldwide evidence that there might be a conditional difference in SWB between and within the various types of regions within a country. However, while there had been conducted extensive research on objective measurements of quality of life from a regional perspective, the evidence of the relationship between happiness and specific region types is rather scarce (Ballas, 2013).[3] Among the few, who have been researching this topic, Glaeser, Gottlieb and Ziv (2014) find that happiness levels differ across U.S. metropolitan areas. Individuals living in declining cities tend to be unhappier, *ceteris paribus*. However, historical data suggest that this lower happiness levels were persistent even when these areas were prosperous. Furthermore, the happiness levels are particularly low in areas with low income growth rates, *ceteris paribus*. These findings are persistent after using various controls, including income and individual fixed effects. Moreover, authors find that inequality does not affect happiness levels in these areas. Two possible explanations are offered: the first one suggests that these cities have been always unhappy, and this caused their decline. The second possible explanation proposes that individuals willingly sacrifice their happiness in order to receive some material compensation in return, such as higher income or lower housing rents. Further, Ballas and Tranmer (2011), applying multilevel modeling techniques on geographical data from the UK Census of Population and BHPS data, find that happiness levels differ across regions. However, this finding loses its statistical significance after including controls for individual, household and area characteristics.

Since empirical evidence from previous literature on conditional regional difference in SWB in CEE countries is scarce, the hypotheses building is mostly open to logical reasoning. First, although the gradual economic convergence between CEE countries and Western European countries has been well documented, the opposite trend is observed between the regions within CEE. Thus, one can observe the trend of economic divergence between capital city regions and other regions within CEE (for example, Rodriguez-Pose and Ezcurra, 2010; Chapman and Meliciani, 2012; Dogaru et al., 2015). Further, as already has been discussed, there is a strong

[3] For an extensive literature review of objective determinants of quality of life read Ballas (2013, pp. 40-42).

link between economic growth and happiness in transition economies (Nikolova, 2015; Djankov et al., 2016). On the other hand, one might reasonably argue that the economic development will not only differ between regions but also within regions. For example, it is reasonable to expect that, for example, the capital city will differ in its economic development to a rural village, located somewhere at the outskirts of the region. Intuitively, the same difference in economic development between the big cities and rural areas in other regions might be persistent. Hence, the following hypothesis is proposed:

Hypothesis 1a: *After controlling for country fixed effects and time fixed effects, SWB is higher in the capital city regions as compared to the other regions within CEE countries.*

Hypothesis 1b: *After controlling for country fixed effects and time fixed effects, SWB is higher in the big cities as compared to rural areas within the regions within CEE countries.*

Furthermore, it will be interesting to test for conditional regional difference in SWB even after controlling for a set of the widely-used socio-demographic and regional characteristics. As it has been already discussed, Glaeser et al. (2014) finds persistent difference in happiness levels across U.S. metropolitan areas, which is persistent after using various controls, including income and individual fixed effects. Further, results obtained in Andrén and Martinsson (2006) and Lelkes (2006) indicate that residence area has statistically significant effect on life satisfaction even after using various controls. Moreover, Gerdtham and Johannesson (2001), using a data sample of 5 106 individuals from the Level of Living Survey, find that urbanization negatively effects happiness, *ceteris paribus*. On the other hand, the authors do not find any statistically significant difference in the effect of urbanization on happiness for rural areas or areas with small cities. Hayo (2007), studying a data sample from seven CEE countries, finds that individuals living in rural areas are on average happier compared to individuals living in large cities, *ceteris paribus*. Hence, a second hypothesis is proposed:

Hypothesis 2a: *After controlling for a set of widely-accepted socio-demographic, macroeconomic and environment characteristics, there is an unexplained conditional gap in SWB between the capital city regions and other regions within CEE countries.*

Hypothesis 2b: *After controlling for a set of widely-accepted socio-demographic, macroeconomic and environment characteristics, there is an unexplained gap in SWB within regions of CEE countries.*

2.3. Social Capital and Its Link with SWB

What becomes evident from the literature review in section 2.2 is that a large number of academic articles have tried, intentionally or not, to relate the differences in SWB between CEE countries to social capital. For example, Bartolini et al. (2015) study the effect of social trust on SWB; Nikolova (2015) studies the effect of the rule of law on SWB; Rodriguez-Pose and Maslauskaite (2012) and Djankov et al. (2016) partly study the effect of corruption and institutions quality on SWB. The next paragraphs will try shed life on the questions regarding what is exactly social capital and does it affect in any way SWB.

Throughout the years, academics tried to introduce a precise and monogamous definition of social capital. However, this led to even broader variety of potential candidates for the term (Woolcock, 1998; van Oorschot et al., 2006). Below are presented four available definitions of social capital based on Bourdieu (1986), Coleman (1988), Woolcock (1998) and OECD (2001)[4]:

> *"the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition – or in other words, to membership in a group – which provides each of its members with the backing of the collectivity-owned capital, a "credential" which entitles them to credit, in the various senses of the word." (Bourdieu, 1986 pp. 248)*

> *"[social] capital is defined by its functions. It is not a single entity, but a variety of different entities having two characteristics in common: they all consist of some aspect of social structures, and they facilitate certain actions of actors – whether persons or corporate actors – within the structure." (Coleman, 1988 pp. 98)*[5]

> *"[social] capital, a broad term encompassing the norms and networks facilitating collective action for mutual benefit." (Woolcock, 1998 pp. 155)*

> *"networks together with shared norms, values and understandings that facilitate co-operation within or among groups." (OECD, 2001 pp. 41)*

[4] For an overview of previous academic literature, theoretical and empirical efforts on the topic of social capital read, for example, Woolcock (1998); Woolcock and Narayan (2000); and Adam and Roncevic (2003).
[5] Variety of illustrative examples of social capital can be found in Coleman (1988, pp. 98-100).

Idealistic situation, known as the "social capital dream", would be that all established proxies of social capital will ultimately measure the same thing. Nevertheless, as it will become evident, this assumption might be overoptimistic.

Coleman (1988) distinguishes between two intellectual streams, which have been dealing with social capital – the first one being mostly attributed to the work of sociologists while the second one is related to the work done by economists. The former stream argues that individual's actions are guided by the environment. More precisely, individual's actions are guided by established social norms, rules and obligations. On the other hand, the latter intellectual stream believes that the individual is purely self-interested actor, who always aims to maximize its own utility (Coleman, 1988). Nevertheless, Civil Economy challenges the general economic belief that individuals should be treated as *homines economici*, or in other words purely self-interested (Becchetti et al., 2015). Supported by the Italian economist and philosopher Antonio Genovesi, the Civil Economy paradigm states that trust and relationships among the individuals are the key aspects of a successful economy. The interdependence between a successful economy and high levels of social capital has been also proven empirically. For example, as acknowledged in Glaeser et al. (2002), referring to Knack and Keefer (1997) and LaPorta et al. (1997), increase of trust increases economic growth and judicial efficiency. Knack and Keefer (1997) find that the relationship between trust and economic growth is especially strong in less developed countries. Countries blessed with high levels of social capital should expect higher economic growth (Woolcock, 1998). Trust, civic participation, high density of interpersonal networks should all facilitate higher efficiency. As argued in Paldam (2000), social capital levels can be influenced by the quality of the government and institutions.

Coleman (1988) tries to combine elements of the two intellectual streams – sociological and economical. The author proposes to think of social capital as a kind of resource, available to the individual. As further discussed, social capital is a resource, which might be accumulated not only by individuals, but also by organizations. The author further argues that social capital consists of three distinctive components: 1) trust and obligations, 2) information channels, and 3) norms and sanctions. The first component relies on mutual trust between the economic agents during the execution of a deal. Further, information channels might not only enhance knowledge diffusion but also serve as a knowledge resource. Lastly, norms and sanctions reduce criminal

actions and develop stronger community ties (Rodriguez-Pose and von Berlepsch, 2014). The multifaceted character of social capital is acknowledged in Van Oorschot et al. (2006). As further discussed, social capital has three dimensions: networks, trust, and civism. However, each dimension has a two-level perspective – individual level and aggregated level.

Rodriguez-Pose and von Berlepsch (2014), referring to Putnam (2000) and Olson (1982), distinguish between informal and formal social interactions. While the former largely consist of informal interactions with friends and family, the latter are formed from participation in various groups. For example, political parties, trade unions, professional organizations, *et cetera*. As further discussed, referring to Olson (1982), interpersonal networks might not necessarily lead to socially-beneficial outcomes. Thus, for example Olson-type formal social interaction might trigger resource redistribution from non-members to members of the various organizations and groups, which might not be optimal for the society as a whole.

Some of the accepted proxies for social capital are social trust, corruption, density of civic participation, density of interpersonal networks, *et cetera* (Paldam and Svendsen, 2000; Paldam, 2000).

The effect of social capital on SWB has been well-documented by previous empirical literature. For example, Helliwell and Putnam (2004) find that social capital indirectly through health and directly influence SWB. These findings are consistent for various proxies of social capital such as family strength, community and religious ties, civic engagement, neighborhood ties and trust. Helliwell et al. (2009) argue that together with income, social capital is an important determinant of SWB, which could explain between and within countries differences in SWB. Furthermore, the authors argue that very large income-equivalent values should be used in order to compensate for low social capital stock. Further, Helliwell and Barrington-Leigh (2010) find that trust and social networks affect SWB. Moreover, Leung et al. (2013), using data sample from Canadian population, and after controlling for a matrix of individual and demographic characteristics, find that social capital affect happiness. This finding was consistent for all three types of social capital as suggested in Coleman (1988).

Further, Onyx and Bullen (2000) find that levels of social capital substantially differ across various types of communities in Australia. According to the obtained results, in general social capital tends to be higher in rural areas as compared to urban areas. Further, Beugelsdijk and Van

Schaik (2005) conclude that levels of trust differ between European regions. For example, while the level of trust is rather homogeneous in Netherlands, it substantially differs in Italy. These findings are in line with the results obtained in this work. Table 1 in the appendix presents the mean values of social capital for the three possible communities, separately for capital city regions and other regions. As it could be seen in the table, on average all three components of social capital tend to be lower in the big cities or the outskirts of the big cities as compared to the other possible community types.

Further, the positive effect of social capital on SWB in Europe has been also documented. For example, Rodríguez-Pose and von Berlepsch (2014) study the effect of social capital on happiness in Europe. Consistent with Leung et al. (2013), the findings support the hypothesis that social capital is correlated with happiness. These results are consistent for all three dimensions of social capital as suggested in Coleman (1988). Furthermore, the authors find that social capital levels vary across European regions, with the East scoring particularly low. Furthermore, as discussed in the beginning of this subsection, various academic papers has found that various variables, which might be related to social capital, has a strong effect on SWB in CEE (for example, Rodriguez-Pose and Maslauskaite, 2012; Bartolini et al., 2015; Nikolova, 2015; and Djankov et al., 2016). Further, Bartolini et al. (2015) find that the magnitude if the effect of social capital on SWB in transition countries might be compared to that of GDP in the medium-term. Hence, the last hypothesis is proposed:

Hypothesis 3a: *The unexplained conditional gap in SWB between the capital city regions and other regions within CEE countries is partly explained by social capital.*

Hypothesis 3b: *The unexplained conditional gap in SWB within regions of CEE countries is partly explained by social capital*

3. Data & Methodology

3.1. Data

In order to test the proposed hypotheses this study combines data from several sources. First, European Social Survey (ESS) is used as a source for the individual-level data. ESS is suitable for the purposes of this research as it provides information regarding SWB and all components of social capital. Survey waves are released each two years, with first one dating back to 2002, while the last one is from 2014. Data for 5 CEE countries is available throughout the survey

waves, but with gaps. These are Bulgaria (BG), Czech Republic (CZ), Hungary (HU), Poland (PL) and Slovakia (SK). Initially, Romania (RO) was also considered as a potential candidate for analysis. However, data for this country is available only for one wave. Thus, it was dropped out of the final data set. Information regarding the availability of yearly data for each of the sampled countries might be found in Table 2 in the appendix.

Overall, the data set consists of 41 NUTS 2 regions, located in the above mentioned 5 countries. Each of the regions is classified either as a capital city region, a second-tier city regions or other region, accordingly to the classification available at Dogaru et al. (2014, pp. 211). Table 3 in the appendix presents NUTS2 codes, their matching region names, and the region type according to the classification.[6] Cartographical illustration is available in figure 1 in the appendix. Furthermore, the created data set further distinguishes between three possible living locations within each region. These are: 1) a big city or the outskirts of a big city; 2) towns or small cities; 3) country villages or a farm/home in the countryside.

Table 4 presents description of level 1 sampling units. Column 1 presents the list of sampled countries; column 2 presents the number of NUTS2 regions in each of the sampled countries. Column 3 presents the number of observation available for each of the countries. Finally, columns 4-6 present, respectively, the minimum, the mean, and the maximum individual observations per region within a country.

Table 4: Description of level 1 sampling units

			#Obs per Unit		
Stratum	#Units	#Obs	min	mean	max
BG	6	8324	1061	1387.3	2023
CZ	8	12943	1316	1617.9	2138
HU	7	11518	1197	1645.4	3012
PL	16	12430	307	776.9	1766
SK	4	8764	1030	2191.0	2903
Total: 5	41	53979	307	1316.6	3012

Note: This table presents the description of level 1 sampling units. The data is obtained from European Social Survey. The survey data is divided into five stratums. These are Bulgaria (BG), Czech Republic (CZ), Hungary (HU), Poland (PL) and Slovakia (SK). Column 2 presents the number of NUTS2 regions in each of the countries.

[6] Under the classification available in Dogaru et al. (2014, pp. 211), the capital city regions are 6 instead of 5. This is the case because there are two regions from the Czech Republic which represents this group - CZ01 and CZ02.

Common issues when dealing with survey data, as it becomes evident from tables 2 and 4, are the inclusion probabilities of an individual and the ratio of observations to the total population size of a region/country. For example, as one could observe in table 2 and 4, regardless to the different population of the five countries in the sample, the number of yearly observations per country is very similar. Thus, weights are applied in order to tackle this issue. Accordingly to the survey guidelines, two types of weights are used when dealing with ESS data – post-stratification weights (*SPSWGHT*), which take into account the differences of the inclusion probabilities, and population size weights (*PWEIGHT*), which are the same for all individuals within a country, but differ between countries. They account for the fact that different countries have different populations but similar sample size. According to the survey guidelines the final weight is created by multiplying both weights: *WEIGHT= SPSWGHT*PWEIGHT*.

After all of the cleaning and recoding procedures are completed, the individual level data set is merged with regional and country level data sets, which incorporate regional controls. These controls are retrieved from EUROSTAT, The World Bank, AirBase and ESPON. Data matching is done using NUTS2 region codes and years. However, ESS waves 1-4 use NUTS2 labels instead of NUTS2 codes. Thus, corresponding NUTS2 codes are created from a matching procedure of the region names and their corresponding codes, using information available in EUROSTAT. Data from the World Bank database is matched using country codes and years.

3.1.1. Dependent variable

In general, the two most widely accepted proxies of SWB are estimated by asking some variation of the following two question: (1) *"Taking all things together, how would you say things are these days – would you say you're very happy, fairly happy, or not too happy these days?"* or (2) *"On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with the life you lead?"* (Alesina et al., 2004). As discussed in Alesina et al. (2004), the answers recorded from these two questions are highly correlated. However, there has been an ongoing debate in the academic literature regarding which of the two variables is most reliable. For example, Alesina et al. (2004) prefer the question with the word "satisfied", since some concerns are raised regarding the translation of the meaning of "happy" across different nations and cultures. On the other hand, Veenhoven (2000; 2004) discusses this potential pitfall and

concludes that using happiness is a valid measure of overall quality of life. The author concludes that happiness levels can be compared across individuals and countries.

For robustness purposes this paper uses as a proxy for SWB two questions available throughout all of the ESS waves:

The first dependent variable, named ***HAPPY***, is constructed using the answers to the question regarding how happy is the individual: *"Taking all together, how happy would you say you are?"*, with answers ranging on a 0-10 scale from *"Extremely unhappy*" to *"Extremely happy"*.

The second dependent variable, named ***LIFESAT***, is constructed using the answers to the question regarding the general life satisfaction: *"All things considered, how satisfied are you with your life as a whole these days?"*, with answers ranging on a 0-10 scale from *"Extremely dissatisfied*" to *"Extremely satisfied"*.

Other possible answers for both questions such as *"Refusal"*, *"Don't know"* or *"No answer"* are treated as missing. Both variables are treated as continuous since previous findings in Ferrer-i-Carbonell and Frijters (2004) suggest that this assumption yields reliable and consistent results.

The correlation between ***LIFESAT*** and ***HAPPY*** is 0.6845, which gives some space for potential variation in results, depending on which of the variables is chosen as proxy for SWB.

3.1.2. Independent Variable

The creation of the independent variable, which will test the three hypotheses, is a two-step process. First, two initial variables are created. The first one is a dummy variable, ***CAPITAL***, taking value 1 if a NUTS2 region is classified as a capital city region and 0 otherwise. The second variable, ***AREA***, indicates the place of living within the region. It is a categorical variable with category 1 representing a big city or the outskirts of a big city. Category 2 represents towns or small cities, and category 3 represents a country village or a farm/home in the countryside. The second step involves the creation of a single categorical variable, ***LOCATION***, which has 6 categories. Thus, a specific region belongs to category 1 if the dummy *CAPITAL* equals 1 and the categorical variable *AREA* equals 1. Further, the region is assigned to category 2 if the dummy *CAPITAL* equals 1 and the categorical variable *AREA* equals 2. To category 3 are assigned regions for which the dummy *CAPITAL* equals 1 and the categorical variable *AREA*

equals 3. For regions assigned to category 4, the dummy *CAPITAL* equals 0 and the categorical variable *AREA* equals 1. Further, category 5 consists of regions for which the dummy *CAPITAL* equals 0 and the categorical variable *AREA* equals 2. Finally, category 6 consists of regions for which the dummy *CAPITAL* equals 0 and the categorical variable *AREA* equals 3.

3.1.3. Social Capital Controls

The multifaceted concept of social capital has been well-acknowledged in previous literature (for further discussion read, for example, Bjornskov, 2006; Van Oorschot et al., 2006; Rodriguez-Poze & von Berlepsch, 2014). Idealistic situation, known as the "social capital dream", would be that all established measurements will ultimately measure the same thing – social capital. Nevertheless, this assumption might be overoptimistic. The idea of creating a single social capital index variable, which would incorporate all aspects of social capital in it, was eventually overthrown. In line with Rodriguez-Poze & von Berlepsch (2014), social capital takes the form of:

$$SCi\ [TRUSTi\ (STRUSTi, ITRUSTi), SINTi\ (INSINTi, FSINTi), NORMi\ ] \quad (1)$$

where SC_i is the level of social capital of the individual i. Social capital, SC, is the combination of trust ($TRUST_i$), which by itself is formed from social trust ($STRUST_i$) and institutional trust ($ITRUST_i$); social interactions, formed from informal social interaction ($INSINT_i$) and formal social interaction ($FSINT_i$). The final component, forming social capital, is the norms and sanction ($NORM_i$).

This paper uses three different indexes in order to measure each of the social capital dimensions. ***TRUST*** is an index variable, incorporating in itself proxies for social and institutional trust. Further, ***SOCIALINTERACTION*** is an index variable, incorporating in itself proxies for informal and formal social interactions. Finally, ***NORMS*** is the third and last index variable, which incorporates in itself the proxies for norms and sanctions. These index variables are created after applying factor analysis (command *factor* in STATA), using the *pcf* option for principal-components factoring. All of the proxy variables, which are used for the creation of the three index variables, have been established in previous literature as valid measures of social capital's components (for example read, Rodriguez-Poze & von Berlepsch, 2014; Bartolini et al., 2015).

Table 5 presents a summary of the variables which are used for the creation of each of the three indexes, the number of non-missing observation, as well as the survey questions and the scale range for the answers.

Table 5: Summary Statistics of Social Capital Variables I

Variable	Survey Question	Range	#Obs
TRUST			
PPLTRST	*"Most people can be trusted in dealing with people"*	0-10	53664
PPLHLP	*"Would you say that most of the time people try to be helpful or that they are mostly looking out for themselves?"*	0-10	53538
TRUSTPRL	*"Please tell me on a score of 0-10 how much you personally trust country's parliament?"*	0-10	52382
TRUSTLGL	*"Please tell me on a score of 0-10 how much you personally trust the legal system?"*	0-10	51863
TRUSTPLC	*"Please tell me on a score of 0-10 how much you personally trust the police?"*	0-10	52734
TRUSTPLT	*"Please tell me on a score of 0-10 how much you personally trust politicians?"*	0-10	52533
STATEHLTH	*"Please say what you think overall about the state of health services in [country] nowadays?"*	0-10	52437
SOCIALINTERACTION			
SOCIAL	*"How often do you meet socially with friends, relatives or work colleagues?"*	0-6	53663
ACHURCH	*"Apart from special occasions such as weddings and funerals, about how often do you attend religious services nowadays?"*	0-6	53488
WORKEDPRT	*There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you worked in a political party or action group?*	0-1	53655
WORKEDORG	*There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you worked in orked in another organisation or association?*	0-1	53606
CONTACTEDPLT	*There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you contacted a politician, government or local government official?*	0-1	53642
BADGE	*There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you worn or displayed a campaign badge/sticker?*	0-1	53584
PETITION	*There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you signed a petition?*	0-1	53533

TRADEUNION	*Are you or have you ever been a member of a trade union or similar organisation? If yes, is that currently or previously?*	0-1	53507
NORMS			
PPLFAIR	*"Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair?"*	0-10	53131
SAFEDARK	*"How safe do you – or would you - feel walking alone in this area after dark?"*	0-3	52779

Note: Other possible answers for all questions such as *"Refusal"*, *"Don't know"* or *"No answer"* are treated as missing. It should be noted that some answers to different questions use, for example, value 1 as the highest and 4 as the lowest or vice versa. This was accounted for by checking manually survey's answer cards, for each of the questions used in this paper, and then recoding the variables' values to 0 being always the lowest possible answer.

Furthermore, in this study each of the index variables for social capital is measured simultaneously in two levels – individual and regional, according to previous findings in Van Oorschot et al. (2006) and Rodriguez-Poze & von Berlepsch (2014). The authors find evidence indicating that values of some of the social capital's components vary between regions in Europe. Thus, regional values of the index variables ***TRUST***, ***SOCIALINTERACTION*** and ***NORMS*** are computed by taking the mean value of all individual-level observations, for each of the regions in the sample. Table 6 presents summary statistics for the three index variables, both on individual and regional levels.

Table 6: Summary Statistics of Social Capital Variables II

Variable	Definition	#Obs	Mean	Std. Dev.	Min	Max
TRUST	Index variable measuring trust, which is formed from social trust and institutional trust	49079	-2.74e-10	1,00	-2,48	3,83
SOCIALINTERACTION	Index variable measuring social interactions, which is formed from informal social interactions and formal social interactions	52142	2.38e-09	1,00	-0,59	7,42
NORMS	Index variable measuring norms and sanctions	52024	0,0012	1,00	-2,91	2,61
REG_TRUST	Regional mean value of TRUST	53979	-0,0046	0,33	-0,80	1,05
REG_SOCIALINTERACTION	Regional mean value of SOCIALINTERACTION	53979	0,0012	0,18	-0,45	0,60
REG_NORMS	Regional mean value of NORMS	53979	-0,0005	0,23	-0,76	0,88

Note: This table reports the summary statistics of social capital variables. The data is obtained from European Social Survey. *TRUST*, *SOCIALINTERACTION* and *NORMS* are created after applying factor analysis (command factor in STATA), using the pcf option for principal-components factoring. *REG_TRUST*, *REG_SOCIALINTERACTION* and *REG_NORMS* are the regional means of *TRUST*, *SOCIALINTERACTION* and *NORMS,* respectively.

3.1.4. Socio-Demographic Controls

In line with previous findings, socio-demographic control variables include individual's age, the squared value of the age (age^2), a dummy indicating if the individual is unemployed and actively searching for a job, years of completed education, number of people living in the household, categorical variable for marital status, dummy indicating the gender, and subjective health. All of these variables have been found to affect SWB (for summary, Dolan et al., 2008).

Further, specific attention should be paid to personal income since data regarding this factor is problematic. The effect of personal income on SWB has been proven highly significant in previous empirical literature (for example, Alesina et al. 2004; Luttmer 2004; Rodriguez-Poze & von Berlepsch, 2014). Thus, omitting personal income out of the model might violate the zero conditional mean assumption that the explanatory variables are uncorrelated with the error term. In turn this will cause biased results. However, retrieving individual income data appears to be a problematic issue for a sample consisting of relatively small number of countries and all survey waves. Variables *netpay* and *grspay*, which incorporate information regarding the individual net and gross wages, respectively, are not available systematically across all of the survey waves.

Instead, relative household income variable is constructed after combining two identical variables – *hinctnt* and *hinctnta*. The question which is asked during the survey interview is *"If you add up the income from all sources, which letter describes your household's total net income? If you don't know the exact figure, please give an estimate."* Original answer values are coded with letters which are then recoded into numbers after checking the possible answers (Survey Card 53).

Unfortunately, for a very high percentage of surveyed individuals this question is marked either as *"Refusal"* or *"Don't know"*. Thus, setting these two answers as missing would yield only 34903 valid observations out of 53979. In order to preserve the observations, the newly created variable ***INCOME*** is treated as categorical. Categories from 0 to 10 represent the different income levels, with 10 being the highest income. Category 11 incorporates answers *"Refusal"* and *"No answer"*, while category 12 represents answers *"Don't know"*.

Another worth discussion control variable is *lrscale*, which is based on a survey question *"In politics people sometimes talk of "left" and "right". Using this card, where would you place*

yourself on this scale, where 0 means the left and 10 means the right?". The variable is recoded into a new variable named ***LEFTRIGHT***, which has 4 categories. Category 1 represents the "left" political spectrum, category 2 stays for "in the middle", while category 3 is the "right" political spectrum. The assigning of the initial 10 categories into the newly created "left", "in the middle", and "right" followed the methodology used in Alesina et al. (2004). Thus, first 4 categories are recoded into category 1, the middle three categories are recoded into category 2, and the last four categories are recoded into category 3. Finally, category 4 incorporates the answers marked as "Don't know". The last category is created assuming that some individuals might be indifferent to politics and do not take any particular side. Other possible answers such as *"Refusal"* or *"No answer"* are treated as missing. The inclusion of this variable is justified by previous findings, which reveal difference in reaction to different socio-demographic and macroeconomic factors conditional on the political ideology (for example, Alesina et al., 2004; Van Oorschot et al., 2006).

3.1.5. Regional Controls

In line with previous findings, the set of macroeconomic controls is as follows (for summary, Dolan et al., 2008).

Natural logarithm of the GDP per capita at PPP measures the regional economic situation. Another regional economic variable is the percentage of regional unemployment for the population aged from 20 to 64 years. Data for both variables is retrieved from EUROSTAT and is available in NUTS2 level.

The effect of income inequality on happiness has been well researched in the past. Among others, Alesina, Di Tella and MacCulloch (2004), find that relative inequality on a global scale has a negative effect on individuals' happiness. This finding is well supported by other empirical papers (for example, Easterlin, 1974; 1995; 2003; Luttmer, 2005). However, relative inequality's effect is stronger in Europe. This finding is persistent after controlling for various individual characteristics, including income, as well as year and country fixed effects (Alesina, Di Tella and MacCulloch, 2004). Unfortunately, data for the Gini coefficient of income inequality is not available disaggregated to NUTS2 level. As discussed in Djankov et al. (2016), acquiring information regarding this coefficient is very problematic. Indeed, together with individual income, most problematic was the collection of information regarding Gini coefficients. Most

data sources have missing yearly observations for particular countries or such information does not exist for a particular country at all. One might reasonably expect that income inequality might differ between country regions. For example, intuitively one might expect that income inequality might be higher in large urbanized regions. As argued in Glaeser (2011), prosperous places attract the poor, who hope to find a better future there. On the other hand, agglomeration economies and knowledge accumulation, which are common for highly developed regions, might create favorable condition for the rich to reinvest their money and get even richer. On the contrary, one might oppose such claim with another argument – that the economy of rural places is mainly driven by agriculture. This might also favor inequality if small percentage of individuals, represented by big corporations, owns the land, while most of the population works for them. Whichever case might be true, not accounting for regional income inequality might yield false conclusions. After extensive search for a potential data source, which might provide the needed statistical information, only one source appeared to have this information – OECD.Stat. Unfortunately, the database maintained by OECD provides information regarding Gini coefficient on NUTS2 level only for 2014. Furthermore, from the countries of interest, there is data available only for Czech Republic, Hungary and Poland. After taking into account the missing values from the other variables in the model, this leads to approximately 3000 observations. Considering the relatively large number of variables, it was decided not to use this data as results would not be reliable. Thus, the construction of the variable controlling for income inequality is based on the available country-level data from the World Bank.[7] This might be considered as one of this study's limitations.

While the effect of income inequality on happiness has been well researched in the past, very little attention has been paid to happiness inequality as a potential explanatory tool for happiness (Helliwell et al. 2016). Goff, Helliwell and Mayraz (2016) find that SWB inequality has negative effect on SWB. This result remained statistically significant after introducing to the model

[7] It should be mentioned that several observations had missing values for the Gini coefficient. In order to preserve these observations, the following commands were executed in STATA:
bys cntry: g gr = D.gini/L.gini
xtreg gr year, fe vce(robust)
predict growth
replace gini=gini[_n-1]+(gini[_n-1]*growth[_n]) if gini==. & gini[_n-1]!=.
where gr is the annual growth rate of the Gini coefficient and growth is the predicted values of the annual growth rate of the Gini coefficient.

controls for income inequality, region dummies and various personal controls. Moreover, the negative coefficient of SWB inequality consistently remained higher as compared to the coefficient of income inequality. Similar findings are presented in Bolle, Okhrin and Vogel (2009). In line with these findings, an additional control variable for regional SWB inequality is introduced. The standard deviation of SWB levels in a region is used as a measure of SWB inequality. This is in line with the method used in Goff, Helliwell and Mayraz (2016). Two variables are created, ***HAPPYINEQ*** and ***LIFESATINEQ***. While the first is used as a control in the models where the dependent variable is ***HAPPY***, the latter is used in the models where the dependent variable is ***LIFESAT***.

Further, the level of regional urbanization is controlled with ***POPDENSITY***, which is the regional population density. Data regarding population density is obtained from EUROSTAT. Population density is calculated by taking the total regional population and dividing this number by the total area in km^2 of the region. Summary of previous literature results regarding the effect of this variable on SWB might be found in Dolan et al. (2008). Initially, ***TRAFFIC***, taking the form of total number of vehicles divided by the total number of roads in km, was considered as measure of traffic congestion. However, this variable is highly correlated with ***POPDENSITY*** (0.9711). Thus, it was decided to use only the regional population density as a measure of regional urbanization level.

Furthermore, additional variable, which controls for regional accessibility, is created.[8] As described in Spiekermann and Wegener (2006), area accessibility is constructed from a set of indicators such as the ease of reaching an area, the cost of reaching that area, and the availability of desired activities within the area. The data availability of this indicator is somewhat limited for the countries in the sample. Nevertheless, two indicators of accessibility were retrieved from ESPON database - potential accessibility by air and multimodal potential accessibility. These indicators are available at NUTS3 level, which are then aggregated to NUTS2 level and merged with the data set. However, both indicators are available only for 2006. Under the assumption that regional accessibility might be relatively constant in the medium long-run, values for 2006 are assigned to the corresponding NUTS2 regions for all years.

[8] Definition of accessibility is proposed in Spiekermann and Wegener (2006, pp. 17), referring to Wegener et al. (2002) : *"accessibility indicators describe the location of an area with respect to opportunities, activities or assets existing in other areas and in the area itself, where 'area' may be a region, a city or a corridor"*.

Finally, the newly created data set consists of two variables, which control for air pollution and climate.

The first one is ***AIRPOLLUTION***. Previous literature findings suggest that increased air pollution has an adverse impact on SWB (for example, Smyth et al., 2008; MacKerron and Mourato, 2009; Luechinger, 2009, 2010). Furthermore, Ferreira et al. (2013), combining individual data from the first three waves of ESS with air pollution data from AirBase, find that increased sulfur dioxide levels (SO_2) have negative effect on individuals SWB, *ceteris paribus*. Taking into account the NUTS2 aggregation level of the data set, SO_2 is chosen as an air pollutant. The reasoning behind this decision is three-fold. First, SO_2 has a negative impact on human's health. Second, as compared to other air pollutants, it can be directly noticed by individuals. Lastly, while other pollutants are more localized, SO_2 pollution can spread on a wide regional scale (Ferreira et al., 2013).

Acquiring data on air pollution with SO_2 is a lengthy process involving several steps.[9] First, data on air pollution was retrieved from AirBase - the European air quality database. Second, a spreadsheet is created with pollution levels from all of the available air monitoring stations throughout the countries of interest.[10] After that, information regarding the geographical coordinates of each station is added to the spreadsheet. This data is then analyzed in geo-statistical software ArcGIS. Interpolation is conducted using the inverse-distance weighting (IDW) approach. Finally, the values of air pollution are dissolved to NUTS2 boundaries. This data is then extracted and merged with the data set using NUTS2 codes and years as matching components.[11]

Regional climate is controlled with ***HDD***, which represents the mean heating degree days (HDD) over the period 1980 - 2009 for each NUTS2 region. This climate control is created accordingly to previous findings in Rehdanz and Maddison (2005), which link climate to individuals'

[9] A step-by-step description of the process is available in "Technical report on GIS Analysis, Mapping and Linking of Contextual Data to the European Social Survey" by Brereton, Moro, Ningal and Ferreira, which is a part of the HAPPINESS project of the Cross-National and Multi-level Analysis of Human Values, Institutions and Behaviour (HumVIB) programme (http://www.ucd.ie/happy/index.html). Although authors make their data set publicly available, it consists of only the first three waves of the ESS, which makes it unsuitable for the purposes of this paper.

[10] Location of SO_2 air monitoring stations for 2014 might be seen in figure 2 in the appendix.

[11] Data for 2002 was missing for Bulgarian regions. The missing values were substituted by summing the air pollution values for 2001 and 2003 and dividing them by two.

happiness. Annual HDD data is available from 1980 until 2009 in EUROSTAT.[12] The annual values are then collapsed over the entire period by NUTS2 regions, taking the mean regional values. The reasoning behind the creation of mean values is twofold. First, annual data in EUROSTAT for HDD is available only until 2009. Second, one might reasonably assume that regional climate should be fairly constant in the long run.

Summary statistics of the final data set are available in table 7 in the appendix. Correlation between variables in the final data set is checked and all results suggest that there is no problem to continue with the analysis.[13][14]

3.1.6. Preliminary Evidence

After all of the cleaning and merging processes are finished, the data set is analyzed for any possible preliminary evidence that there is a SWB gap between regions.

Figure 3 presents computed unconditional life satisfaction and happiness means at a country and region (NUTS2) levels. One could observe that the differences between countries are substantial, taking into account that the measurement scale ranges from 0 to 10. The difference between the lowest country mean and the highest country mean in the sample is around 2 points for both, life satisfaction and happiness. The difference is even bigger when observing the NUTS2 means – almost 3 points for life satisfaction and slightly less for happiness.

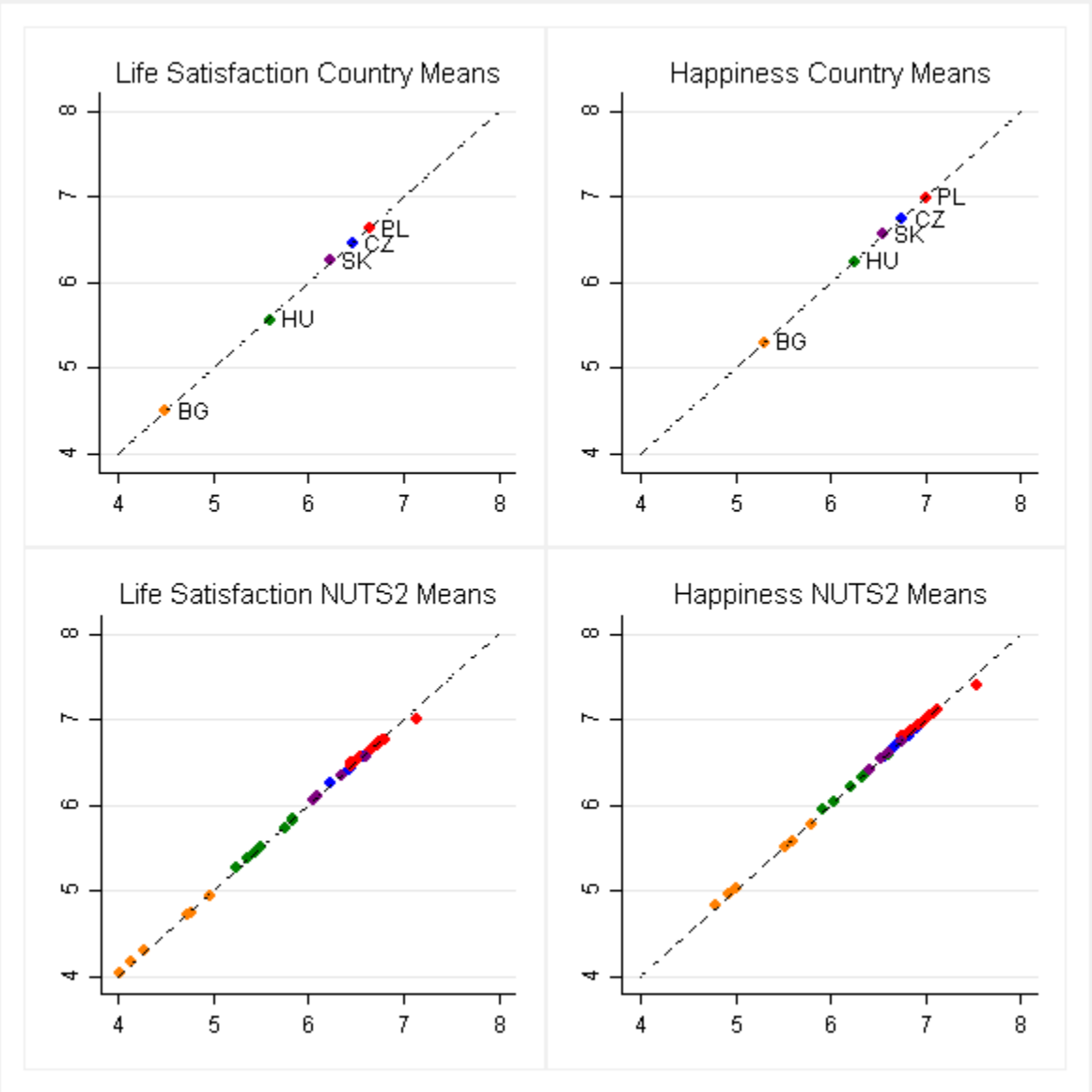
Another interesting observation is that in general happiness means are higher as compared to life satisfaction. This is especially true in the case of Bulgaria, where the difference between life satisfaction and happiness is almost one point.

[12] Definition of HDD is available in EUROSTAT – *"Actual heating degree-days express the severity of the cold in a specific time period taking into consideration outdoor temperature and room temperature. To establish a common and comparable basis, Eurostat defined the following method for the calculation of heating degree days (18 °C - Tm) x d if Tm is lower than or equal to 15 °C (heating threshold) and are nil if Tm is greater than 15 °C where Tm is the mean (Tmin + Tmax / 2) outdoor temperature over a period of d days. Calculations are to be executed on a daily basis (d=1), added up to a calendar month -and subsequently to a year- and published for each Member State separately."*
(EUROSTAT, available online at: http://ec.europa.eu/eurostat/cache/metadata/en/nrg_esdgr_esms.htm)

[13] Scores for age and age^2 are not taken into account as it is expected that they will be highly correlated between each other.

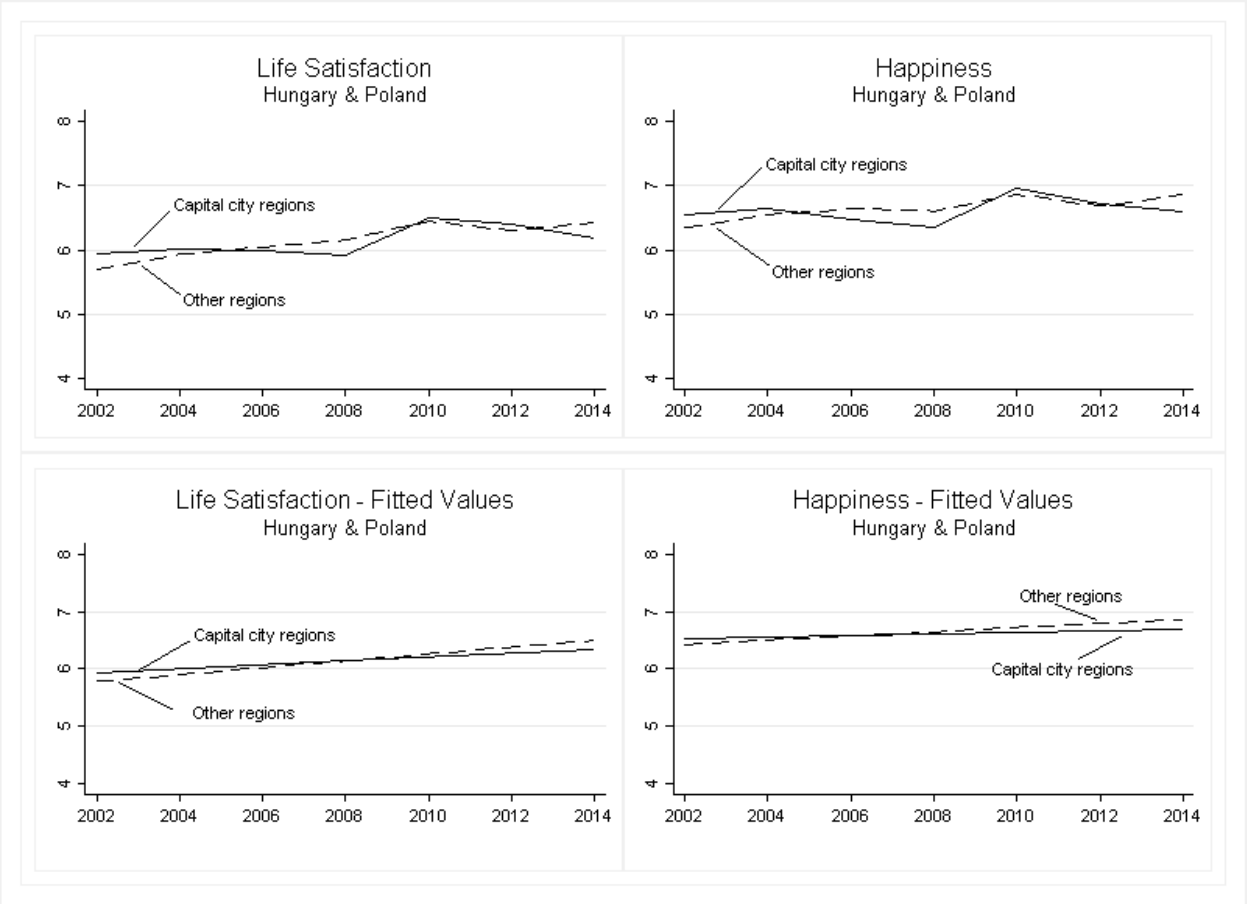
[14] Correlation table is not included due to the substantial number of variables in the final data set. Correlation results are available on request: igeorgiev@student.eur.nl

Figure 3: Unconditional Country and Regional SWB Means



Note: This graph presents life satisfaction and happiness means at a country and NUTS2 level. The color of each region matches with the color of the country. For example, the life satisfaction and happiness means for Bulgaria in the first two graphs are market with orange dots. Accordingly, the regional life satisfaction and happiness means for Bulgaria in the next two graphs are marked with orange dots.

Figure 4: Unconditional SWB levels: Capital City Regions versus Other Regions



Note: This figure shows the mean and fitted values of life satisfaction and happiness for capital city regions and other regions in Hungary and Poland from 2002 to 2014. The data is obtained from European Social Survey.

Figure 4 presents a case study of two countries from the sample – Hungary and Poland. These two countries are selected for drawing figure 4 since all 7 ESS waves contain data for them. On the other hand, including all countries could lead to misleading conclusions while interpreting the figure.

First two graphs present a comparison of the actual unconditional life satisfaction and happiness levels between the capital city regions and other regions. Although reported life satisfaction and happiness in 2002 is higher in the capital city regions, the opposite is true in the period from 2006 until 2010. In the next period, from 2010 until 2012, the reported values in both types of regions are moving closely in the same direction. Interestingly, after this period, the slightly

downward trend for the capital city regions continues, while the trend for other regions is reversed. Moreover, one can observe that in 2014 the life satisfaction and happiness levels in other regions are higher compared to the reported values in the capital city regions.

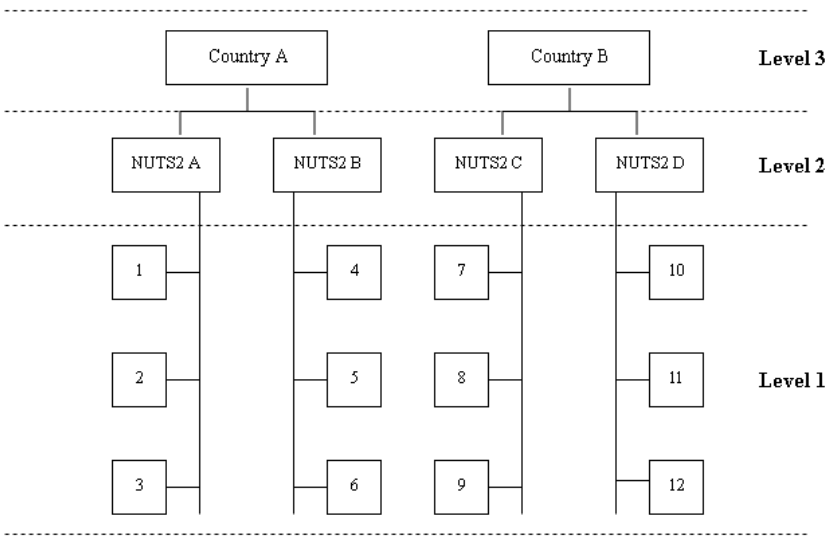
Furthermore, the two graphs in row 2 of figure 4 report the fitted life satisfaction and happiness values for both types of regions. While all of the lines have an upward trend, those for other regions surpasses the capital city regions in the period around 2006 – 2008. Finally, in line with observation made in figure 3, it appears that happiness values are higher as compared to life satisfaction levels. By itself this is an interesting observation and justifies measuring SWB using both variables as a robustness check.

3.2. Methodology

The proposed hypotheses are tested using a 3-level hierarchical model. This model takes into account the possible role of space. Previous findings suggest that subjective well-being might be influenced not only by a set of compositional factors but also by some contextual determinants (Ballas and Tranmer, 2011). For example, it is well documented that macroeconomic indicators such as regional GDP, unemployment and income inequality effect SWB (for example, Alesina et al., 2004; Luttmer, 2004; for summary, Dolan et al., 2008). Moreover, regional characteristics such as air pollution, congestion, region accessibility and climate have been all found to influence to a various extent SWB (for example, Rehdanz & Maddison, 2005; Ferrer-i-Carbonell & Gowdy, 2007; Brereton et al., 2008). Thus, models which take into account the possible interdependence of individual SWB outcomes at various spatial levels should be preferred (Ballas and Tranmer, 2011).

Illustrative example of a 3-level hierarchical model might be seen in figure 5. Individual observations are the first level of the model. These observations are nested into the second level, which consist of NUTS2 regions. Regions are further nested into countries, which represent the third level.

Figure 5: 3-Level Hierarchical Model



Note: This figure presents a graphical illustration of a 3-level hierarchical model. Individual level observations, which represent the first level of the model, are nested into NUTS2 regions. Further, NUTS2 regions, which represent the second level of the model, are nested into the countries. Finally, countries represent the third level of the model.

One feature of a hierarchical model is that it allows the estimation of intra-class correlation coefficients (ICC) for each of its levels. In turn, these coefficients allow to understand how much of the variation of the dependent variable can be attributed to each level of the model. The ICC is computed in several steps. First, an empty model, which is also called a "null" model, is estimated in order to compute the variance components:

$$y_{ijkt} = \alpha_0 + \delta_{kt} + \varepsilon_{jkt} + u_{ijkt} \quad (1)$$

where y is the dependent variable, ijk are the three nested levels: countries are indexed by k (k = 1,…,5); NUTS2 regions are indexed by j (j = 1,…,41); and individuals are indexed by i (i = 1,…,N). t is the time period and α0 is the constant term. δ and ε are the random effect terms for level 3 (countries) and level 2 (NUTS2 regions), respectively. Finally, u is the individual-level error term. δ, ε and u are assumed to be uncorrelated (Ballas and Tranmer, 2011). Further:

$$\mathrm{Var}(y_{ijkt}) = \sigma^2_{\delta} + \sigma^2_{\varepsilon} + \sigma^2_{u} \quad (2)$$

where σ^2_{δ} measures the variation in countries, σ^2_{ε} measures the variation in regions, and σ^2_{u} measures the variation of SWB between individuals. As discussed in Ballas and Tranmer (2011), one of the main assumptions in multilevel modeling is of the interdependence of the error terms

for the different levels of the model. Since δ, ε and u are assumed to be uncorrelated, $\sigma^2_\delta + \sigma^2_\varepsilon + \sigma^2_u$ sum up the total variation in SWB.

The intra-class correlation coefficients (ICC) are then calculated applying the following formulas (Albright & Marinova, 2010):

$$ICC_k = COUNTRYVAR/(COUNTRYVAR+REGIONVAR+INDIVIDVAR) \quad (3)$$

where ICC_k is the intra-class correlation coefficient at the country level (level 3). COUNTRYVAR is the variation at the country level (level 3), REGIONVAR is the variation at the NUTS2 level (level 2), and INDIVIDVAR is the variation at the individual level (level 1).

$$ICC_j = (REGIONVAR)/(COUNTRYVAR+REGIONVAR+INDIVIDVAR) \quad (4)$$

where ICC_j is the intra-class correlation coefficient at the NUTS2 level (level 2). COUNTRYVAR is the variation at the country level (level 3), REGIONVAR is the variation at the NUTS2 level (level 2), and INDIVIDVAR is the variation at the individual level (level 1).

Further, in order to test hypothesis 1, the following model is applied:

$$y_{ijkt} = \alpha_0 + \beta_1 * LOCATION_{jk} + \theta_t + \mu_k + \delta_k + \varepsilon_{jk} + u_{ijk} \quad (5)$$

where y represents the dependent variable, ijk are the three nested levels: countries are indexed by k (k = 1,…,5); NUTS2 regions are indexed by j (j = 1,…,41); and individuals are indexed by i (i = 1,…,N). t is the time period and α0 is the constant term. *LOCATION* is a categorical variable, which test hypothesis 1. θt are time fixed effects and μ_k are country fixed effects. δ and ε are the random effect terms for level 3 (countries) and level 2 (NUTS2 regions), respectively. Finally, u is the individual-level error term.

Hypothesis 2 is testes with the following model:

$$y_{ijkt} = \alpha_0 + \beta_1 * LOCATION_{jk} + \lambda' V_{ijkt} + \gamma' Z_{jkt} + \theta_t + \mu_k + \delta_k + \varepsilon_{jk} + u_{ijk} \quad (6)$$

where y represents the dependent variable, ijk are the three nested levels: countries are indexed by k (k = 1,…,5); NUTS2 regions are indexed by j (j = 1,…,41); and individuals are indexed by i (i = 1,…,N). t is the time period and α0 is the constant term. *LOCATION* is a categorical variable, which test hypothesis 2. V is a matrix of variables controlling for socio-demographic characteristics, Z is a matrix of variables controlling for regional characteristics. θt are time fixed

effects and μ_k are country fixed effects. δ and ε are the random effect terms for level 3 (countries) and level 2 (NUTS2 regions), respectively. Finally, u is the individual-level error term.

Further, hypothesis 3 is tested in two stages. The first stage consists of the following model:

$$y_{ijkt} = \alpha_0 + \beta_1 * LOCATION_{jk} + \eta' ISC_{ijkt} + \nu' RSC_{jkt} + \lambda' V_{ijkt} + \gamma' Z_{jkt} + \\ + \theta_t + \mu_k + \delta_k + \varepsilon_{jk} + u_{ijk} \quad (6)$$

where y represents the dependent variable, ijk are the three nested levels: countries are indexed by k (k = 1,…,5); NUTS2 regions are indexed by j (j = 1,…,41); and individuals are indexed by i (i = 1,…,N). t is the time period and $\alpha 0$ is the constant term. *LOCATION* is a categorical variable, which test hypothesis 3. V is a matrix of variables controlling for socio-demographic characteristics, Z is a matrix of variables controlling for regional characteristics. ISC is a matrix of three index variables *TRUST*, *SOCIALINTERACTION* and *NORMS*, controlling for individual-level social capital. RSC is a matrix of three variables *REG_TRUST*, *REG_SOCIALINTERACTION* and *REG_NORMS*, controlling for regional-level social capital. θt are time fixed effects and μ_k are country fixed effects. δ and ε are the random effect terms for level 3 (countries) and level 2 (NUTS2 regions), respectively. u is the individual-level error term.

Finally, mediation analysis is performed in the second stage. As discussed in Preacher and Hayes (2008), an independent variable might affect the dependent variable directly and indirectly, through a set of mediator variables. Mediation analysis is conducted in two-steps. First, all of the necessary coefficients are obtained performing seemingly unrelated regression (*sureg* command in STATA):

(MV1 IV X) … (MV6 IV X) (DV MV1 MV2 MV3 MV4 MV5 MV6 IV X) (7)

where MV1-MV6 are the six mediators: *TRUST*, *SOCIALINTERACTION*, *NORMS*, *REG_TRUST*, *REG_SOCIALINTERACTION* and *REG_NORMS*. IV is the categorical variable *LOCATION*. DV is one of the two dependent variables *HAPPY* or *LIFESAT*. Finally, X is a set of covariates controlling for socio-demographic characteristics, regional characteristics, time fixed effects and country fixed-effects. The above-described regression is repeated for the two dependent variables.

The second step consists of estimating the indirect effect and the standard errors through the nonlinear combination command *nlcom* in STATA. This step is repeated separately for the indirect effect of each of the mediators, for each of the categories of *LOCATION*, and once for the total indirect effect. The above-described process is then repeated for the two dependent variables.

4. Results

4.1. Happiness

First, an empty multilevel mixed-effects linear regression is estimated (*xtmixed* command in STATA), with added option *mle* which fits the model via maximum likelihood. The model is estimated in order to obtain the variance values at different levels. The dependent variable is ***HAPPY***.

Conducted estimations yield ICC of 0.0671 at the country level and ICC of 0.0084 at the region level. These results might be interpreted as follow. Approximately 6.7% of the variation in SWB might be attributed to the countries and approximately 0.84% of the variation in SWB might be attributed to the regions.

The estimation of ICC and the obtained results have a two-fold implication. First, although the variation at the regional level is relatively low, multilevel model is to be preferred as compared to OLS regression. This is supported by the variation at the country level. Choosing a multilevel model is also supported by the LR test, which rejects the null hypothesis that linear regression is to be preferred at the 1% significance level. Second, the results justify the inclusion of regional control variables in the model, since not all of the variation of SWB can be explained at the individual level.

Next, models are gradually loaded with independent and control variables, which aim to test the three hypotheses. All models are computed with robust standard errors in order to control for possible heteroscedasticity. In line with ESS guidelines weights are applied to all models. Furthermore, the number of observations in all models equals the number of observations in the full model in order to prevent any false conclusions due to the difference in observations between models.

First, a baseline model is estimated with only the categorical variable *LOCATION*, indicating the place of living, time fixed effects and country fixed effects. This model tests hypothesis 1. Further, controls for regional characteristics are introduced in model 2. Hypothesis 2 is tested in model 3, where socio-demographic controls are added. Finally, social capital is added in model 4. This model tests hypothesis 3.

Due to the large number of variables, table 8 presents a short summary of models 1-4. The long specification of table 8 might be found in the appendix – table 9. Model 1 incorporates the categorical variable *LOCATION*, indicating the place of living, time and country fixed effects.

As expected, categories 2-6 have negative coefficient signs. Thus, individuals living in another location then a big city or its outskirts, in the capital city region, are in general less happy, *ceteris paribus*. However, only categories 4 and 5 are statistically significant at the 5% level. Further, the magnitude of the negative coefficient of category 5 is bigger as compared to category 4. These results partially confirm both parts of hypothesis 1.

The next two models help to understand how much of the effect of *LOCATION* is absorbed by regional and socio-demographic characteristics. Regional characteristics are introduced in model 2. Table 9 in the appendix presents the statistics for all of the variables included in the model.

In line with previous literature, air pollution has a negative effect on individual happiness levels, *ceteris paribu*s. This effect is statistically significant at the 10% level. Heating degree days, which control for regional climate and more specifically, how severe are the temperature deviations from previously accepted as a mild conditions, are not statistically significant. Further, happiness inequality has a statistically significant at 1% level negative effect on happiness. This finding is in line with previous literature (for example, Goff, Helliwell and Mayraz, 2016). Regional economic development, measured by the natural logarithm of the GDP per capita at PPP has a positive effect on happiness and is highly significant. Somewhat surprisingly, income inequality is not statistically significant. However, one might hypothesize that this effect is absorbed by happiness inequality, which due to the data constraints is measured more precisely. Further, regional unemployment rate has negative and highly statistically significant effect on happiness. Population density, which measures the level of urbanization, is statistically significant at 1% level and has a negative coefficient. However, the coefficient is 0 to three decimal places. Finally, region accessibility by air is not statistically significant.

Further, categories 2-6 of the categorical variable *LOCATION*, indicating the place of living, are all statistically insignificant. This result might be indicative that the negative and statistically significant effect on happiness of categories 5 and 6 relative to category 1 might be potentially attributed to regional characteristics.

Socio-demographic characteristics are introduced in model 3. In line with previous findings, all of the socio-demographic characteristics have the expected coefficient signs (for example, Dolan et al., 2008). Age has highly statistically significant u-shaped relationship with happiness. The dummy variable, indicating if an individual is unemployed and currently seeking work, has statistically significant negative effect on happiness. Individuals, who have completed more years of education, tend to be happier, *ceteris paribus*. This effect is statistically significant at the 5% level. The effect of the number of people living in the household on happiness is not statistically significant.

Further, individuals who are married or are in a civil partnership tend to be happier as compared to individuals who are divorced or separated, *ceteris paribus*. Individuals who have never been married tend to be happier as compared to individuals who are divorced or separated, *ceteris paribus*. However, this effect is much smaller as compared to the first scenario. In the former case the effect is statistically significant at the 1% level and in the latter case the effect is statistically significant at the 5% level. Finally, individuals who are widowed are relatively happier as compared to the first category, *ceteris paribus*. This effect is statistically significant at the 5% level.

As expected, males are on average less happy compared to females, *ceteris paribus*, and this effect is highly statistically significant. Subjective health has positive and highly significant effect on happiness. Further, individuals who are located in the right political spectrum are relatively happier as compared to individuals in the left political spectrum, *ceteris paribus*. This effect is significant at the 1% level. The coefficients for category 2 and 4 are not statistically significant. Finally, the results for the categorical variable for household income suggest that money might indeed buy happiness in CEE. The positive coefficients for categories 2-10 increase as compared to the previous level. While the coefficient for category 2 is statistically insignificant, the coefficients for categories 3-10 are statistically significant at 1% level. This indicates that moving from one category into a subsequent upper income category, increases

happiness relative to the first category, *ceteris paribus*. Regarding the last two categories, where people refuse to answer or do not know the household income, both are relatively happier as compared to the individuals with lowest income, *ceteris paribus*.

While the magnitude and signs of the coefficients of variables, controlling for region characteristics, are somewhat reduced as compared to model 2, their significance levels are largely preserved. SWB inequality, regional unemployment rate preserve and population density preserve their significance levels. However, the significance of the natural logarithm of GDP per capita is reduced from 1% to 10%. Further, the region accessibility by air is now statistically significant at 5% level. The negative coefficient of the variable is somewhat suspicious and raises potential concerns regarding the specification. One possible explanation might be that the accessibility by air has a negative impact on happiness because the regions with increased accessibility also suffer from an increased congestion. Further, the negative effect of air pollution on happiness increases its statistical significance from 10% to 5%.

Finally, as in model 2, categories 2-6 of *LOCATION* remain statistically insignificant. Thus, hypothesis 2 cannot be confirmed.

Model 4 introduces controls for social capital. In general the coefficient signs and significance levels of the variables controlling for regional and socio-demographic characteristics are robust as compared to model 3. The only notable difference is that the variable measuring air pollution completely loses its statistical significance. Furthermore, while the significance level of category 3 of the variable indicating the marital status is reduced from 5% to 10%, category 4 loses its statistical significance. Lastly, the positive effect of GDP per capita on happiness is now statistically significant at 5%.

In line with previous literature, the index variable measuring social and institutional trust has a positive effect on happiness and it is highly statistically significant. However, somewhat surprising is that the variable, which measures the mean regional value of the index variable *TRUST*, has negative and statistically significant at 5% effect on happiness. Further, the index variable, which measures the individual level of social interactions, has a positive coefficient and it is significant at 5%. However, the variable, which measures the regional mean value of the index variable *SOCIALINTERACTION*, is statistically insignificant. Moreover, the index variable, measuring the individual level of norms and sanctions has a positive effect on

happiness and it is highly statistically significant. On the other hand, the variable, measuring the regional mean level of the index *NORM* is statistically insignificant.

Table 8: Regression Results for Happiness

	(1) HAPPY	(2) HAPPY	(3) HAPPY	(4) HAPPY
LOCATION				
1. Capital city region – Big city/outskirts of big city	Reference	Reference	Reference	Reference
2. Capital city region - Small city or town	-0.142	-0.100	0.048	0.052
	(0.112)	(0.085)	(0.109)	(0.090)
3. Capital city region - Rural village or farm in the countryside	-0.311	-0.269	-0.019	-0.062
	(0.211)	(0.189)	(0.197)	(0.146)
4. Other region - Big city/outskirts of big city	**-0.237****	0.052	0.157	0.189
	(0.095)	(0.168)	(0.159)	(0.132)
5. Other region - Small city or town	**-0.278****	0.018	0.173	0.152
	(0.130)	(0.190)	(0.160)	(0.136)
6. Other region - Rural village or farm in the countryside	-0.269	0.031	0.256	0.158
	(0.189)	(0.244)	(0.182)	(0.165)
Social Capital				YES
Socio-Demographic Controls			YES	YES
Region Characteristics		YES	YES	YES
Time FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Observations	38,214	38,214	38,214	38,214
Number of countries	5	5	5	5
Number of NUTS2 regions	41	41	41	41
Loglikelihood	-65448.07	-65347.43	-62397.79	-61693.97
AIC	130934.1	130748.9	124897.6	123501.9
BIC	131096.6	130979.7	125333.7	123989.3

Note: This table reports the results of the 3-level hierarchical regression analysis. Model 1 tests hypothesis 1, with only *LOCATION*, time fixed effects and country fixed effects included. Hypotheses 2 and 3 are tested in models 3 and 4, respectively. *LOCATION* is the main variable of interest. It has 6 categories. Category 1 indicates a place of living in a big city or the outskirts of a big city in capital city regions. This is the reference category. Category 2 indicates a place of living in a small city or a town in capital city regions. Category 3 indicates a place of living in a country village or a farm/home in the countryside in capital city regions. Category 4 indicates a place of living in a big city or the outskirts of a big city in other regions. Category 5 indicates a place of living in a small city or a town in other regions. Category 6 indicates a place of living in a country village or a farm/home in the countryside in other regions. Further, the dependent variable is *HAPPY*. It indicates the individual level of happiness. The full table might be found in the appendix. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

It is a categorical variable with category 1 representing a big city or the outskirts of a big city. Category 2 represents towns or small cities, and category 3 represents country village or farm/home in the countryside.

Finally, the main variable of interest, namely the categorical variable *LOCATION*, which indicates the place of living, has coefficients, which are fairly close to those observed in model 3. Further, categories 2-6 remain statistically insignificant. Thus, hypothesis 3 cannot be confirmed from the obtained results.

4.2. Life Satisfaction

In order to check if the obtained results are robust, the results are re-estimated using ***LIFESAT*** instead of ***HAPPY*** as the dependent variable. This sub-section replicates the steps done in the previous sub-section. All methods and models are identical except for the dependent variable and the variable, which measures SWB inequality. The latter is substitute for a variable measuring life satisfaction inequality.

First, an empty model, containing only the dependent variable ***LIFESAT*** is estimated in order to obtain the intra-class correlation coefficients.

Conducted estimations yield ICC of 0.0963 at the country level and ICC of 0.0077 at the region level. These results might be interpreted as follow. Approximately 9.6% of the variation in SWB might be attributed to the countries and approximately 0.77% of the variation in SWB might be attributed to the regions.

This suggests that a relatively larger percentage of the variance in life satisfaction levels is explained at the country and region levels as compared to happiness.

Due to the large number of variables, table 10 presents a short summary of models 5-8. The long specification of table 10 might be found in the appendix – table 11. Model 5 incorporates the categorical variable *LOCATION*, indicating the place of living, time and country fixed effects.

As expected, categories 2-6 have negative coefficient signs. Thus, individuals living in another location then a big city or its outskirts, in the capital city region, are in general less satisfied with life, *ceteris paribus*. Similar to model 1, only categories 4 and 5 are statistically significant. Further, while category 4 is statistically significant at 1%, category 5 is statistically significant only at 10%. Interestingly, the magnitude of the negative coefficient of category 5 is smaller as compared to category 4. While these results partially confirm the first part of hypothesis 1, they cannot confirm the second part of the hypothesis.

Model 6 introduces controls for region characteristics. While most of the coefficient signs and significance levels are the same as in model 2, there are some differences worth mentioning. First, it seems that life satisfaction inequality has relatively larger negative effect on life satisfaction as compared to the negative effect of happiness inequality on happiness. Second, although the GDP per capita positive effect on life satisfaction is still highly significant, it is slightly smaller as compared to the positive effect of the same variable on happiness. Third, income inequality has a negative and statistically significant at 10% effect on life satisfaction. This variable is not statistically significant in model 2. Further, the negative effect of air pollution on life satisfaction is statistically insignificant. Finally, categories 2-6 of *LOCATION* become statistically insignificant.

Model 7 introduces controls for socio-demographic characteristics. As compared to model 6, the variables measuring GDP per capita and population density completely lose their significance level. Further, all variables, controlling for socio-demographic characteristics, have the expected coefficient signs. Some of the notable differences as compared to the models for happiness are as follow. First, being unemployed and seeking work has larger negative effect on life satisfaction as compared to its effect on happiness. Second, contrasting to model 3, increasing number of individuals living in the household has a highly significant negative effect on life satisfaction. Third, all of the categories of the variable indicating the political belonging are now statistically significant at 5% level. It seems that relative to being in the left political spectrum, all other individuals are in general more satisfied with life, *ceteris paribus*. This difference is the largest for individuals who belong to the right political spectrum. Finally, all of the categories for the level of household income are highly statistically significant and in general have larger positive coefficients as compared to model 3. This might be indicative that personal income is more important for life satisfaction as compared to its effect on happiness.

Regarding *LOCATION*, only the sixth category is statistically significant at 10% level. After controlling for regional and socio-demographic characteristics, individuals living in the rural areas in a region other than the capital city region are on average more satisfied with life as compared to the reference category, *ceteris paribus*. This result is consistent with the results obtained in Hayo (2007). Although hypothesis 2 is partially confirmed, any conclusions should be approached with extreme caution.

Model 8 introduces controls for social capital. In general the coefficient signs and significance levels of the variables controlling for regional and socio-demographic characteristics are robust as compared to model 7.

Table 10: Regression Results for Life Satisfaction

	(5) LIFESAT	(6) LIFESAT	(7) LIFESAT	(8) LIFESAT
LOCATION				
1. Capital city region - Big city/outskirts of big city	Reference	Reference	Reference	Reference
2. Capital city region - Small city or town	-0.153	-0.116	0.094	0.106
	(0.189)	(0.159)	(0.171)	(0.148)
3. Capital city region - Rural village or farm in the countryside	-0.327	-0.291	0.062	0.018
	(0.289)	(0.265)	(0.264)	(0.206)
4. Other region - Big city/outskirts of big city	**-0.375*** **	-0.119	-0.034	0.053
	(0.099)	(0.143)	(0.132)	(0.124)
5. Other region - Small city or town	**-0.303***	-0.032	0.129	0.151
	(0.164)	(0.177)	(0.136)	(0.119)
6. Other region - Rural village or farm in the countryside	-0.276	-0.002	**0.263***	**0.201**
	(0.218)	(0.229)	(0.145)	(0.130)
Social Capital				YES
Socio-Demographic Controls			YES	YES
Region Characteristics		YES	YES	YES
Time FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Observations	38,385	38,385	38,385	38,385
Number of countries	5	5	5	5
Number of NUTS2 regions	41	41	41	41
Loglikelihood	-69012.71	-68898.91	-65991.02	-65169.02
AIC	138063.4	137851.8	132084	130452
BIC	138226	138082.8	132520.4	130939.7

Note: This table reports the results of the 3-level hierarchical regression analysis. Model 5 tests hypothesis 1, with only *LOCATION*, time fixed effects and country fixed effects included. Hypotheses 2 and 3 are tested in models 7 and 8, respectively. *LOCATION* is the main variable of interest. It has 6 categories. Category 1 indicates a place of living in a big city or the outskirts of a big city in capital city regions. This is the reference category. Category 2 indicates a place of living in a small city or a town in capital city regions. Category 3 indicates a place of living in a country village or a farm/home in the countryside in capital city regions. Category 4 indicates a place of living in a big city or the outskirts of a big city in other regions. Category 5 indicates a place of living in a small city or a town in other regions. Category 6 indicates a place of living in a country village or a farm/home in the countryside in other regions. Further, the dependent variable is *LIFESAT*. It indicates the individual level of life satisfaction. The full table might be found in the appendix. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Regarding the statistically significant social capital components in model 8, they preserve their coefficient signs as compared to models 4. Further, while in general the coefficient magnitude is

relatively similar, the index variable, which measures trust, has larger coefficient in model 8 as compared to model 4. Thus, it seems that individual's social and institutional trust matter more for personal life satisfaction as compared to happiness.

Regarding *LOCATION*, categories 2-6 are all statistically insignificant. Category 6, which is statistically significant at 10% in model 7, is statistically insignificant. Further, the coefficient is lower as compared to the previous model. This result partially confirm hypothesis 3 in the case of life satisfaction. Nevertheless, although this result might be indication that social capital does indeed explain the conditional gap in SWB in the case of category 6 relative to the reference category, it should be approached with reasonable caution.

4.3.Mediation Analysis

The potential mediating influence of social capital on the effect of *LOCATION* on SWB is further explored through mediation analysis.

Although there are user-written commands in STATA, which conduct mediation analysis, these have proven to be unsuitable for the purposes of this work. The reasoning is twofold. First, the user-written commands in STATA, which were tested, cannot incorporate categorical variables as possible covariates in the regressions. Second, several of the reviewed alternatives cannot incorporate more than one mediator in the analysis. Thus, all of the steps, which are necessary for mediation analysis, are done manually in STATA, following Preacher and Hayes (2008) and UCLA (2016a; 2016b).

Table 12 presents the results for mediation of the effect of *LOCATION* on happiness through social capital. The total indirect effect of all categories of *LOCATION* combined is statistically insignificant. Interestingly, the total indirect effect of all social capital components, for each of the categories of *LOCATION*, is statistically significant only for categories 3 and 6. These categories indicate a place of living in the rural areas of both, capital city regions and other regions.

The results for mediation of the effect of *LOCATION* on life satisfaction through social capital are presented in table 13. Similar to the case of happiness, the total indirect effect of all categories of *LOCATION* combined is statistically insignificant. Moreover, the total indirect

effect of all social capital components, for each of the categories of *LOCATION*, is statistically significant for categories 3, 4 and 6.

Table 12: Mediation of the Effect of the Six Categories of *LOCATION* on Happiness through Social Capital

	Cat 2	Cat 3	Cat 4	Cat 5	Cat 6
Social and Institutional Trust (***TRUST***)	-0.011	-0.017*	-0.058***	-0.046***	-0.025**
	(0.010)	(0.010)	(0.013)	(0.013)	(0.013)
Regional mean of TRUST	-0.001	0.000	0.020***	0.021***	0.021***
	(0.001)	(0.001)	(0.007)	(0.008)	(0.007)
Informal and Formal Social Interactions (***SOCIALINTERACTION***)	0.000	0.002*	0.004**	0.004**	0.004**
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Regional mean of SOCIALINTERACTION	0.001	0.001	0.030**	0.030**	0.030**
	(0.001)	(0.001)	(0.012)	(0.013)	(0.012)
Norms & Sanctions (***NORM***)	0.005	0.046***	-0.047***	-0.015	0.050***
	(0.010)	(0.010)	(0.013)	(0.013)	(0.013)
Regional mean of NORM	-0.002	-0.001	0.013	0.013	0.012
	(0.001)	(0.001)	(0.008)	(0.008)	(0.008)
Total Indirect Effect	-0.008	**0.031****	-0.039	0.007	**0.091*****
	(0.015)	(0.015)	(0.025)	(0.025)	(0.025)
Total Indirect Effect - All Categories Combined			0.083		
			(0.084)		

Note: This table reports the results of the mediation analysis and tests hypothesis 3. The multiple mediators are *TRUST*, *SOCIALINTERACTION*, *NORM*, *REG_TRUST*, *REG_SOCIALINTERACTION* and *REG_NORM*. The independent variable is *LOCATION*. Category 1 indicates a place of living in a big city or the outskirts of a big city in capital city regions. This is the reference category. Category 2 indicates a place of living in a small city or a town in capital city regions. Category 3 indicates a place of living in a country village or a farm/home in the countryside in capital city regions. Category 4 indicates a place of living in a big city or the outskirts of a big city in other regions. Category 5 indicates a place of living in a small city or a town in other regions. Category 6 indicates a place of living in a country village or a farm/home in the countryside in other regions. Further, the dependent variable is *HAPPY*. It indicates the individual level of happiness. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 13: Mediation of the Effect of the Six Categories of *LOCATION* on Life Satisfaction through Social Capital

	Cat 2	Cat 3	Cat 4	Cat 5	Cat 6
Social and Institutional Trust (***TRUST***)	-0.012	-0.021*	-0.080***	-0.063***	-0.035**
	(0.013)	(0.013)	(0.017)	(0.017)	(0.017)
Regional mean of TRUST	-0.005**	-0.001	0.043***	0.044***	0.045***
	(0.002)	(0.002)	(0.009)	(0.009)	(0.010)
Informal and Formal Social Interactions (***SOCIALINTERACTION***)	0.000	0.002	0.004*	0.004*	0.004*
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Regional mean of SOCIALINTERACTION	-0.000	-0.001	-0.018	-0.018	-0.018
	(0.001)	(0.001)	(0.013)	(0.013)	(0.013)
Norms & Sanctions (***NORM***)	0.008	0.051***	-0.047***	-0.011	0.059***
	(0.011)	(0.011)	(0.014)	(0.014)	(0.014)
Regional mean of NORM	-0.002	-0.001	0.008	0.008	0.008
	(0.002)	(0.001)	(0.008)	(0.008)	(0.008)
Total Indirect Effect	-0.011	**0.029***	**-0.091*****	-0.036	**0.063****
	(0.018)	(0.018)	(0.028)	(0.029)	(0.029)
Total Indirect Effect - All Categories Combined			-0.046		
			(0.098)		

Note: This table reports the results of the mediation analysis and tests hypothesis 3. The multiple mediators are *TRUST*, *SOCIALINTERACTION*, *NORM*, *REG_TRUST*, *REG_SOCIALINTERACTION* and *REG_NORM*. The independent variable is *LOCATION*. Category 1 indicates a place of living in a big city or the outskirts of a big city in capital city regions. This is the reference category. Category 2 indicates a place of living in a small city or a town in capital city regions. Category 3 indicates a place of living in a country village or a farm/home in the countryside in capital city regions. Category 4 indicates a place of living in a big city or the outskirts of a big city in other regions. Category 5 indicates a place of living in a small city or a town in other regions. Category 6 indicates a place of living in a country village or a farm/home in the countryside in other regions. Further, the dependent variable is *LIFESAT*. It indicates the individual level of life satisfaction. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

However, these results should be approached with reasonable caution. The reasoning behind this statement is as follows. First, in order to prove a causal mediating relationship several assumptions should be met as discussed in Linden and Karlson (2013). The first assumption in mediating analysis is of a sequential ignorability. In the case of this study, this assumption implies that the treatment group is ignorable of potential mediators and outcomes. The treatment group here is the individuals, living in different region types and areas, whose living location is

measured with the categorical variable *LOCATION*. Social capital, measured with the three index variables and their regional means, are tested as potential mediators. In observation studies, the assumption of sequential ignorability is met when no residual confounding is observed after conditioning on covariates. Further, the level of social capital should be independent of the potential happiness or life satisfaction outcomes (Linden and Karlson, 2013). However, as discussed in Helliwell and Putnam (2004), reverse causation and selection bias in the case of social capital and SWB cannot be fully ruled out. Further, the second assumption is of no interaction between the treatment and the mediator, which implies that there is a constant effect going from the independent variable to the dependent one (Linden and Karlson, 2013). However, the nature of the independent variable in this study is such that this assumption might not hold. As further discussed in Linden and Karlson (2013), even the inclusion of an interaction term might not produce reliable results. Moreover, the final step of mediation analysis as discussed in Preacher and Hayes (2008), which consists of computing bootstrap standard errors and confidence intervals cannot be executed due to the limited hardware availability. Thus, the results might be not reliable since the nonlinear combination command *nlcom* computes the standard errors using the delta method. This method is valid only if the estimates of the indirect effects follow normal distribution (UCLA, 2016a; 2016b). Thus, the obtained results should not be interpreted as causal but merely as potential suggestive evidence.

5. Robustness Checks

The advantages of using a hierarchical model, which takes into account the possible interdependence of individual SWB outcomes at various spatial levels, are well explained in Ballas and Tranmer (2011). Nevertheless, the structure of the available data set might cause potential issues. As discussed in Bell et al. (2010), an established rule of thumb calls for at least 30 units in each of the model's levels. While this condition is met at the individual level (level 1) and at the NUTS2 level (level 2), the number of CEE countries in the final data set is limited to only 5. In order to test potential issues, Bell et al. (2010) study the behavior of a 2-level hierarchical model with limited data availability. The authors conclude that a limited number of level-2 units leads to less accurate confidence intervals of level-2 predictors. However, the authors conclude that level-1 estimates remain accurate. Thus, another model is used in order to test if the results obtained in section 4 are accurate.

As discussed in Ferrer-i-Carbonell and Frijters (2004), the most widely used in economics studies models, when dealing with SWB data, are either the ordered logit or ordered probit models. These models assume ordinal comparability across SWB categories. Following among others van Praag et al. (2000), and assuming that the ordered logit and ordered probit models should yield very similar results, this study conducts robustness checks applying ordered probit model. Although the coefficients cannot be compared to the hierarchical model used in section 4, the coefficients' signs and significance levels should provide some informative evidence.

All models are computed with robust standard errors in order to control for possible heteroscedasticity. In line with ESS guidelines weights are applied to all models. Furthermore, the number of observations in all models equals the number of observations in the full model.

Due to the large number of control variables, table 14 presents a short summary of models 9-12, with ***HAPPY*** being the dependent variable. The long specification of table 14 might be found in the appendix – table 15.

In model 9, categories 2-6 has negative coefficient signs. Further, categories 3-6 are statistically significant at 1% and category 2 is statistically significant at 10%. However, it seems that the significance of the categories, indicating the specific place of living, is almost exhausted after the introduction of regional controls. In model 10, only category 3 is significant at 1% level.

Comparison of results presented in table 7 and table 15, reveals that signs and significance levels of all variables, which control for regional characteristics, are relatively robust. The only notable difference is that the negative effect of air pollution on happiness in now insignificant. Moreover, GDP per capita is statistically significant at 1% in all models.

Further, all signs of variables, controlling for socio-demographic characteristics, are robust and as expected. Their significance levels are relatively robust as compared to models 3 and 4, with only minor changes. After the introduction of socio-demographic characteristics, category 3 of the categorical variable *LOCATION* becomes statistically insignificant. On the other hand, category 6 is statistically significant at 10%.

Regarding social capital proxies in model 12, only ***TRUST*** and ***NORM*** retain their significance levels from model 4. On the other hand, the variable, measuring the regional mean of trust, as well as the variable, measuring social interactions, lose their significance. Further, the

introduction of social capital in model 12 reduces the significance level of category 6 of *LOCATION* to insignificant.

Table 14: Robustness Check Using Ordered Probit Regression Model with HAPPY as the Dependent Variable

	(9) HAPPY	(10) HAPPY	(11) HAPPY	(12) HAPPY
LOCATION				
1. Capital city region - Big city/outskirts of big city	Reference	Reference	Reference	Reference
2. Capital city region - Small city or town	**-0.070***	-0.052	0.013	0.015
	(0.038)	(0.042)	(0.042)	(0.043)
3. Capital city region - Rural village or farm in the countryside	**-0.147*****	**-0.133*****	-0.027	-0.051
	(0.039)	(0.042)	(0.041)	(0.042)
4. Other region - Big city/outskirts of big city	**-0.098*****	0.018	0.055	0.075
	(0.026)	(0.054)	(0.054)	(0.055)
5. Other region - Small city or town	**-0.107*****	0.011	0.070	0.063
	(0.024)	(0.055)	(0.055)	(0.057)
6. Other region - Rural village or farm in the countryside	**-0.105*****	0.017	**0.107***	0.062
	(0.024)	(0.055)	(0.056)	(0.057)
Social Capital				YES
Socio-Demographic Controls			YES	YES
Region Characteristics		YES	YES	YES
Time FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Observations	38,214	38,214	38,214	38,214
Number of countries	5	5	5	5
Number of NUTS2 regions	41	41	41	41
Loglikelihood	-62229.11	-62136.41	-59415.02	-58783.51
AIC	124508.2	124338.8	118944	117693
BIC	124722	124621	119431.4	118231.7

Note: This table reports the results of the ordered probit regression analysis. Model 9 tests hypothesis 1, with only *LOCATION*, time fixed effects and country fixed effects included. Hypotheses 2 and 3 are tested in models 11 and 12, respectively. *LOCATION* is the main variable of interest. It has 6 categories. Category 1 indicates a place of living in a big city or the outskirts of a big city in capital city regions. This is the reference category. Category 2 indicates a place of living in a small city or a town in capital city regions. Category 3 indicates a place of living in a country village or a farm/home in the countryside in capital city regions. Category 4 indicates a place of living in a big city or the outskirts of a big city in other regions. Category 5 indicates a place of living in a small city or a town in other regions. Category 6 indicates a place of living in a country village or a farm/home in the countryside in other regions. Further, the dependent variable is *HAPPY*. It indicates the individual level of happiness. The full table might be found in the appendix. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Due to the large number of control variables, table 16 presents a short summary of models 13-16, with ***LIFESAT*** being the dependent variable. The long specification of table 16 might be found in the appendix – table 17.

Table 16: Robustness Check Using Ordered Probit Regression Model with LIFESAT as the Dependent Variable

	(13) LIFESAT	(14) LIFESAT	(15) LIFESAT	(16) LIFESAT
LOCATION				
1. Capital city region - Big city/outskirts of big city	Reference	Reference	Reference	Reference
2. Capital city region - Small city or town	**-0.063***	-0.051	0.042	0.043
	(0.038)	(0.042)	(0.042)	(0.042)
3. Capital city region - Rural village or farm in the countryside	**-0.124*****	**-0.111****	0.034	0.010
	(0.040)	(0.043)	(0.044)	(0.044)
4. Other region - Big city/outskirts of big city	**-0.139*****	-0.054	-0.017	0.015
	(0.026)	(0.054)	(0.054)	(0.055)
5. Other region - Small city or town	**-0.099*****	-0.009	0.062	0.067
	(0.024)	(0.054)	(0.055)	(0.056)
6. Other region - Rural village or farm in the countryside	**-0.085*****	0.004	**0.121****	0.087
	(0.024)	(0.055)	(0.056)	(0.057)
Social Capital				YES
Socio-Demographic Controls			YES	YES
Region Characteristics		YES	YES	YES
Time FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Observations	38,385	38,385	38,385	38,385
Number of countries	5	5	5	5
Number of NUTS2 regions	41	41	41	41
Loglikelihood	-65524.51	-65418.33	-62724.46	-61984.93
AIC	131099	130902.7	125562.9	124095.9
BIC	131312.9	131185	126050.6	124634.9

Note: This table reports the results of the ordered probit regression analysis. Model 13 tests hypothesis 1, with only *LOCATION*, time fixed effects and country fixed effects included. Hypotheses 2 and 3 are tested in models 15 and 16, respectively. *LOCATION* is the main variable of interest. It has 6 categories. Category 1 indicates a place of living in a big city or the outskirts of a big city in capital city regions. This is the reference category. Category 2 indicates a place of living in a small city or a town in capital city regions. Category 3 indicates a place of living in a country village or a farm/home in the countryside in capital city regions. Category 4 indicates a place of living in a big city or the outskirts of a big city in other regions. Category 5 indicates a place of living in a small city or a town in other regions. Category 6 indicates a place of living in a country village or a farm/home in the countryside in other regions. Further, the dependent variable is *LIFESAT*. It indicates the individual level of life satisfaction. The full table might be found in the appendix. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In model 13, categories 2-6 has negative coefficient signs. Further, categories 3-6 are statistically significant at 1% and category 2 is statistically significant at 10%. However, it seems that the significance of the categories, indicating the specific place of living, is almost exhausted after the introduction of regional controls. In model 14, only category 3 is significant at 5% level. These results are very similar to the results obtained for happiness in tables 12 and 13.

As in the case of happiness, comparison of results obtained from the hierarchical model and the ordinal probit model suggests that regional and socio-demographic controls are fairly robust. Again the only notable difference for region controls is that the negative effect of air pollution is reduced to insignificant. Interestingly, GDP per capita is statistically significant only at 10% in model 14 and statistically insignificant in models 15 and 16. This might suggest that for the case of life satisfaction regional economic development is not as important as in the case of happiness. By itself this is an interesting finding and might indicate that life satisfaction and happiness should not be used as interdependent terms. Further, category 3 of *LOCATION* becomes statistically insignificant in model 15. On the other hand, category 6 is now statistically significant at 5%.

Regarding social capital proxies in model 16, all variables retain their significance levels from model 8. Finally, category 6 of *LOCATION*, which is statistically significant at 5% level in model 15, becomes statistically insignificant in model 16.

6. Discussion & Conclusion

Studying the conditional difference in SWB between and within regions in CEE, and its determinants, was inspired by the published happiness reports, which indicate notoriously low levels of happiness in this geographical area. For example, on a possible scale from 0 to 10, the average mean value of happiness in CEE countries is 5.554 as compared to a score of 6.575 in Western Europe (Helliwell et al., 2016 pp. 15). Further, the example of Bulgaria is particularly appealing – the country is ranked 129th in the World Happiness Report 2016, which puts it behind much weaker economies such as Ethiopia, Nepal and Sierra Leone. Moreover, gathered evidence from previous academic literature indicated that one might expect to observe a conditional difference in SWB between various types of regions and communities (for example, Gerdtham and Johannesson, 2001; Andrén and Martinsson, 2006; Lelkes, 2006; Hayo, 2007). Further, the specific history of CEE, marked by the transition from communism and planned

economy to democracy and market economy, justified the need to study in more detail the role of social capital as a possible explanation tool of the conditional difference in SWB. As discussed in Paldam and Svendsen (2000), various forms of dictatorship such as totalitarianism, ordinary dictatorship or absolutism can erode social capital stock within a country. Further, the very strong effect of social capital on SWB is documented (for example, Helliwell et al., 2009; Bartolini et al., 2015). Moreover, results from previous literature indicate that the levels of social capital stock differ between the different types of communities and regions (for example, Onyx and Bullen, 2000; Beugelsdijk and Van Schaik, 2005; Rodríguez-Pose and von Berlepsch, 2014). Finally, as discussed in Helliwell et al. (2009), social capital potentially might be able to explain the differences in SWB between and within countries. These findings from previous academic literature justified the need to study the conditional regional differences in SWB in CEE and the role of social capital.

The main results from the conducted analysis indicate that, before accounting for socio-demographic and regional controls, individuals, living in the big cities or outskirts of the big cities in the capital city regions in CEE, tend to have higher SWB levels as compared to individuals living in other regions and communities. The results from the 3-level hierarchical model indicate that this difference is particularly strong between individuals, living in the big cities or outskirts of the big cities in the capital city regions, and individuals living in the big cities or outskirts of the big cities in other than the capital city regions. However, the difference in SWB between and within regions in CEE is largely explained by the environmental and economic characteristics. In the case of happiness, all categories, indicating the region and community type, become statistically insignificant when region characteristics are pooled into the model. However, in the case of life satisfaction, and after controlling for regional and socio-demographic characteristics, individuals, living in the rural communities of other than the capital city regions, are on average more satisfied with their lives as compared to individuals living in the big city communities in the capital city regions, *ceteris paribus*. This effect becomes insignificant after social capital controls are introduced into the model. Further, the mediation analysis suggests that in the case of rural communities as compared to big city communities, the total indirect effect on SWB through social capital is statistically significant. Thus, it seems that social capital at least partly might be accounted for conditional gap in SWB between the rural community and the big city community within CEE countries. Further, for life satisfaction, the

total indirect effect on SWB through social capital is statistically significant also in the case of big city communities in other regions as compared to big city communities in the capital city regions. The results indicate that the biggest contributors to the mediation effect are the norms and sanctions. On the other hand, the mediation analysis suggests that the combined total indirect effect of the region and community types on SWB through social capital is statistically insignificant. Thus, it seems that the major roles in explaining the difference in SWB between regions and communities in CEE have macroeconomic and environmental characteristics. However, these results should be approached with considerable caution due to the unavailability to fully rule out some of the potential problems.

Further, all three components of social capital, as suggested in Coleman (1988), have positive effect on SWB. However, while the magnitude of the effect of trust and norms and sanctions on SWB is relatively similar, the effect of social interactions on SWB is much smaller. Nevertheless, this might have an alternative explanation. The effect of social interactions on SWB is measured by an index variable, which incorporates Putnam-type informal and Olson-type formal social interactions. As discussed in Rodriguez-Pose and von Berlepsch (2014), referring to Olson (1982), some types of formal social interactions might not necessarily be socially optimal. Thus, the overall positive effect of the index variable measuring social interaction might be reduced. Further, interesting finding is the negative effect of the regional mean value of trust on SWB. Moreover, the magnitude of the negative effect of the regional mean value of trust on SWB is larger in the case of life satisfaction as compared to the case of happiness. Mean values of trust presented in table 1 in the appendix suggest that trust is higher in rural communities as compared to big city communities. One possible explanation for the negative effect of regional trust on SWB could be omitted controls, which effect on SWB is absorbed by regional trust. This is one of the fields, which need to be further researched.

Some of the other main and most surprising findings are as follows. First, as expected, the GDP per capita has a strong positive influence on SWB. Further, as discussed in Nikolova (2015) and Djankov et al. (2016), the strong link between economic development and SWB in transition economies is well-established. Moreover, while economic convergence between CEE and Western Europe is observed, the capital city regions and other regions within CEE are diverging (Dogaru et al., 2015). This might be one of the reasons why people in the big cities or their

outskirts in the capital city regions tend to be happier and more satisfied with their lives as compared to individuals living in other regions and communities. However, the effect of GDP per capita is stronger in the case of happiness as compared to the case of life satisfaction. Moreover, after the introduction of socio-demographic controls, the effect of GDP per capita on life satisfaction becomes statistically insignificant, while it remains statistically significant in the case of happiness.

Second, in line with recent findings by Helliwell et al. (2016) and Goff et al. (2016), SWB inequality has a strong negative impact on SWB. This variable remains statistically significant at the 1% level in all models throughout the study, which highlights its importance for future studies of SWB. However, the negative effect of life satisfaction inequality on life satisfaction is bigger as compared to the negative effect of happiness inequality on happiness. Surprisingly, the Gini coefficient of income inequality has a statistically significant negative effect on SWB only in the case of life satisfaction. From the obtained results it seems that SWB inequality and income inequality are more important in the case of life satisfaction as compared to happiness.

Third, personal unemployment has larger negative effect on SWB as compared to the regional unemployment rate. Further, being unemployed and seeking work has larger negative effect on life satisfaction as compared to its negative effect on happiness. Another interesting finding is that personal income matters more in the case of life satisfaction as compared to the case of happiness. This finding is opposite when compared to the effect of GDP per capita on life satisfaction and happiness. Further, the results suggest that belonging to a particular side of the political spectrum tend to matter more in the case of life satisfaction as compared to happiness. Individuals belonging to the right political spectrum tend to be happier and more satisfied with their lives as compared to individuals, who belong to the left political spectrum, *ceteris paribus*. However, this effect is more than two times larger in the case of life satisfaction as compared to the case of happiness.

The above-described findings have one important implication. Among others Alesina et al. (2004) suggest that answers to happiness and life satisfaction are highly correlated and should yield similar results. However, findings in the present study suggest that, in the case of SWB, happiness and life satisfaction should not be used interchangeably since false generalized conclusions might be reached.

To summarize, on average individuals living in the big cities of the capital city regions are happier and more life satisfied as compared to individuals living in other regions and communities. However, this is largely explained by the macroeconomic and environmental characteristics. Although it seems that social capital mediates the effect of the place of living on SWB in the case of rural communities versus big city communities, the total mediating effect of social capital for all regions and communities is insignificant. These results are supported by the conducted mediation analysis and the robustness checks performed with ordered probit models. Thus, it seems that social capital at least partly might be accounted for the conditional gap in SWB between the rural community and the big city community within CEE countries. Further, results indicate that the biggest contributors to the mediation effect are the norms and sanctions. Although social capital, at least partly, might be accounted for the conditional gap in SWB, it seems that the macroeconomic and environmental characteristics have the biggest contribution for the explanation of the difference in SWB between and within regions in CEE. Further, in line with previous literature, all three components of social capital as suggested in Coleman (1988) have positive effect on SWB. However, while the magnitude of the effect of trust and norms and sanctions on SWB is relatively similar, the effect of social interactions on SWB is much smaller. Further, results suggest that happiness and life satisfaction should not be used interchangeably as this might lead to false generalized conclusions. For example, the effect of GDP per capita is stronger in the case of happiness as compared to the case of life satisfaction. On the other hand, personal income has larger positive effect on life satisfaction as compared to its effect on happiness. Moreover, SWB inequality and income inequality are more important in the case of life satisfaction as compared to happiness.

This paper has an important policy implication since it sheds light to a topic, which was hardly researched before. More precisely, this work studies the conditional difference in SWB between and within the different types of regions in CEE. Further, this paper tries to understand if social capital might be accounted for this conditional difference. The above-presented findings should enable the EU and national governments to improve their current policies towards increasing SWB. As discussed in more detail in section 2, increasing SWB should be one of the prime policy goals. The reasoning behind this is as follows. SWB has been proven to influence positively not only the economy, personal health and the society, but might be also an important factor when making voting decisions during elections. Moreover, this paper is also socially

relevant since SWB and social capital are crucial factors for healthy and active society. Feeling safety after dark, having trust in national institutions, participating in various non-profit organizations, and, in general, being socially active would increase people's self-esteem and quality of life. Furthermore, understanding what affects SWB in CEE will enable policy-makers to build happier and more life satisfied society. Not only that, but a better understanding of the causes for a difference in SWB between and within regions in CEE will also make it possible to achieve SWB cohesion across regions.

Nevertheless, this work is not without its own limitations. First of all, the data availability regarding some of the control variables is limited. Another problem is the number of missing observations for some of the used variables. For example, data for GINI coefficient, which is used as a proxy for income inequality, is available only on a country-level. Another example is the substantial number of missing observations for the variable used as a proxy for household income. Furthermore, household income is used as a proxy of personal income since data for personal income is not available consistently throughout all waves of ESS. Second, ESS does not offer data for all of the countries of interest consistently throughout all of the survey waves. As a consequence of this Romania is excluded from the empirical analysis. Another limitation is the small number of countries, which are left in the final data sample. As already discussed in section 5, this might lead to biased results. In order to deal with this issue, robustness checks are performed, substituting the hierarchical model for an order probit model. Further, the limited hardware availability does not allow computing bootstrap standard errors and confidence intervals for the performed mediation analysis. Further, as already discussed in section 4, some of the assumptions needed for causal interpretation of the performed mediation analysis might not hold in the case of this research. Lastly, as discussed in Helliwell and Putnam (2004), reverse causation and selection bias have been well-acknowledged as one of the main problems when studying the topics of social capital and SWB. Other potential issues are the adaptation and "hedonic treadmill". Thus, using cross-sectional data might yield false conclusions (Helliwell and Putnam, 2004).[15]

As a consequence of the above-mentioned limitations the obtained results should not be interpreted as a causal claims, but just as mere associations.

[15] For more elaboration read Helliwell and Putnam (2004, pp. 1437).

Finally, there are several directions for future research. First of all, the relationship between social capital and SWB in CEE should be further researched. Particular attention should be paid to the results obtained from the mediation analysis. More precisely, the statistically significant mediating effect of social capital in the case of rural communities as compared to big city communities should be further studied. Thus, in order to receive more reliable results, mediation analysis should be re-estimated with bootstrapped standard errors and confidence intervals. Further, more countries from the CEE should be included in the data sample. This will not only yield more trustworthy results, but will also allow conducting subsample analysis, which is not possible with the present data sample. However, this is a lengthy process, involving the collection of data from various sources, due to the limited data availability for this geographical region. Second, the negative effect of the regional mean values of trust on SWB should be further researched. Third, it would be interesting to study more in detail the difference between happiness and life satisfaction in CEE. Present results suggest that these two terms might not be interchangeable, at least in the case of CEE. The following proposition is backed-up by the obtained results, which suggest that various macroeconomic and socio-demographic characteristics influence happiness and life satisfaction with a different magnitude. This finding is particularly interesting in the case of regional GDP per capita and household income. Finally, future research might try to cope with the problem of reverse causation.

References

Adam, F., & Rončević, B. (2003). Social capital: recent debates and research trends. Social Science Information, 42(2), 155-183.

Andrén, D., & Martinsson, P. (2006). What contributes to life satisfaction in transitional Romania?. Review of Development Economics, 10(1), 59-70.

Albright, J. J., & Marinova, D. M. (2010). Estimating multilevel models using SPSS, Stata, SAS, and R. Indiana University, 1-35.

Alesina, A., Di Tella, R., & MacCulloch, R. (2004). Inequality and happiness: are Europeans and Americans different?. Journal of Public Economics, 88(9), 2009-2042.

Ballas, D. (2013). What makes a ‘happy city’?. Cities, 32, 39-50.

Ballas, D., & Tranmer, M. (2011). Happy people or happy places? A multilevel modeling approach to the analysis of happiness and well-being. International Regional Science Review, 1-33.

Bartolini, S., Mikucka, M. & Sarracino, F. (2015). Money, Trust and Happiness in Transition Countries: Evidence from Time Series, Social Indicators Research, 0(0), 1-20.

Becchetti, L., Bruni, L. & Zamagni, S. (2015). Human Values, Civil Economy, and Subjective Well-being. In World Happiness Report 2015, Chapter 7, 132-152.

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, Survey Research Methods Section, 4057-4067.

Beugelsdijk, S., & Van Schaik, T. (2005). Social capital and growth in European regions: an empirical test. European Journal of Political Economy, 21(2), 301-324.

Bolle, F., Okhrin, Y. & Vogel, C. (2009). A note on interdependent happiness. The Journal of Socio-Economics, 38, 713–721.

Bourdieu, P. (1986). The forms of capital, in Richardson, J. Handbook of theory and research for the sociology of education, New York: Greenwood, 241–258.

Brereton, F., Clinch, J. P., & Ferreira, S. (2008). Happiness, geography and the environment. Ecological Economics, 65(2), 386-396.

Bjørnskov, C. (2006). The multiple facets of social capital. European journal of political economy, 22(1), 22-40.

Chapman, S. A., & Meliciani, V. (2012). Income Disparities in The Enlarged EU: Socio-Economic, Specialisation and Geographical Clusters. Tijdschrift voor economische en sociale geografie, 103(3), 293-311.

Coleman, J. S. (1988). Social capital in the creation of human capital. American journal of sociology, 95-120.

De Neve, J. E., Diener, E., Tay, L., & Xuereb, C. (2013). The objective benefits of subjective well-being. World happiness report.

Dogaru, T., Burger, M., Karreman, B., & Oort, F. (2015). Functional and sectoral division of labour within Central and Eastern European countries: evidence from greenfield FDI. Tijdschrift voor economische en sociale geografie, 106(1), 120-129.

Dogaru, T., Burger, M., van Oort, F., & Karreman, B. (2014). The Geography of Multinational Corporations in CEE Countries: Perspectives for Second-Tier City Regions and European Cohesion Policy. Investigaciones Regionales, (29), 193-214.

Dolan, P., Metcalfe, R., & Powdthavee, N. (2008). Electing happiness: Does happiness affect voting and do elections affect happiness. Discussion Papers in Economics, 30.

Dolan, P., Peasgood, T., & White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. Journal of economic psychology, 29(1), 94-122.

Djankov, S., Nikolova, E., & Zilinsky, J. (2016). The happiness gap in Eastern Europe. Journal of Comparative Economics, 44(1), 108-124.

Easterlin, R. A. (2009). Lost in transition: Life satisfaction on the road to capitalism. Journal of Economic Behavior & Organization, 71(2), 130-145.

Ezcurra, R., Pascual, P., & Rapún, M. (2007). The dynamics of regional disparities in Central and Eastern Europe during transition. European Planning Studies, 15(10), 1397-1421.

Ferreira, S., Akay, A., Brereton, F., Cuñado, J., Martinsson, P., Moro, M., & Ningal, T. F. (2013). Life satisfaction and air quality in Europe. Ecological Economics, 88, 1-10.

Ferrer-i-Carbonell, A., & Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness?. The Economic Journal, 114(497), 641-659.

Ferrer-i-Carbonell, A., & Gowdy, J. M. (2007). Environmental degradation and happiness. Ecological Economics, 60(3), 509–516.

Fischer, M. M., & Stirböck, C. (2004). Regional income convergence in the enlarged Europe, 1995-2000: A spatial econometric perspective. ZEW-Centre for European Economic Research Discussion Paper, (04-042).

Gerdtham, U. G., & Johannesson, M. (2001). The relationship between happiness, health, and socio-economic factors: results based on Swedish microdata. The Journal of Socio-Economics, 30(6), 553-557.

Glaeser, E. (2011). Triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier. The Penguin Press, New York.

Glaeser, E. L., Gottlieb, J. D., & Ziv, O. (2014). Unhappy cities. National Bureau of Economic Research, (20291).

Goff, L., Helliwell, J. & Mayraz, G. (2016). The Welfare Costs of Well-being Inequality. NBER Working Paper, (21900).

Glaeser, E., Laibson, D. & Sacerdote, B. (2002). An Economic Approach to Social Capital. The Economic Journal, 112(483), 437-458.

Gross National Happiness Commission. (2007). About Us. Retrieved August 18, 2016, from http://www.gnhc.gov.bt/about-us/

Guriev, S. & Zhuravskaya, E. (2009). (Un)Happiness in Transition. CEPR Discussion Paper, (DP7258).

Hayo, B. (2007). Happiness in transition: An empirical study on Eastern Europe. Economic Systems, 31(2), 204-221.

Helliwell, J. F., & Barrington-Leigh, C. P. (2010). How much is social capital worth? National Bureau of Economic Research, (16025).

Helliwell, J. F., Barrington-Leigh, C. P., Harris, A., & Huang, H. (2009). International evidence on the social context of well-being. National Bureau of Economic Research, (14720).

Helliwell, J. F., Layard, R., & Sachs, J. (2015). Setting the Stage. In World Happiness Report 2015, Chapter 1, 2-7.

Helliwell, J. F., Layard, R., & Sachs, J. (2016). Setting the Stage. In World Happiness Report 2016 Update, Chapter 1, 2-11.

Helliwell, J. F., & Putnam, R. D. (2004). The social context of well-being. Philosophical transactions-royal society of London series B biological sciences, 1435-1446.

Helliwell, J. F., Huang, H. & Wang, S. (2016). The Distribution of World Happiness. In World Happiness Report 2016 Update, Chapter 2, 8-50.

Helliwell, J. F., & Putnam, R. D. (2004). The social context of well-being. Philosophical transactions-royal society of London series B biological sciences, 1435-1446.

Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. The journal of economic perspectives, 20(1), 3-24.

Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. The Quarterly journal of economics, 1251-1288.

LaPorta, R., Lopez-de-Salanes, F., Shleifer, A. & Vishny, R. (1997). Trust in large organisations. American Economic Review Papers and Proceedings, 87, 333-338.

Leung, A., Kier, C., Fung, T., Fung, L., & Sproule, R. (2013). Searching for happiness: The importance of social capital. In The exploration of happiness (pp. 247-267). Springer Netherlands.

Lelkes, O. (2006). Tasting freedom: Happiness, religion and economic transition. Journal of Economic Behavior & Organization, 59(2), 173-194.

Liberini, F., Redoano, M. & Proto, E. (2014). Happy Voters. IZA Discussion Papers, (8498).

Linden, A., & Karlson, K. B. (2013). Using mediation analysis to identify causal mechanisms in disease management interventions. Health Services and Outcomes Research Methodology, 13(2-4), 86-108.

Luechinger, S. (2009). Valuing air quality using the life satisfaction approach. The Economic Journal, 119(536), 482-515.

Luechinger, S. (2010). Life satisfaction and transboundary air pollution. Economics Letters, 107(1), 4-6.

Luttmer, E. F. (2004). Neighbors as negatives: Relative earnings and well-being. National Bureau of Economic Research No. 10667.

MacKerron, G., & Mourato, S. (2009). Life satisfaction and air quality in London. Ecological Economics, 68(5), 1441-1453.

Nikolova, M. (2015). Minding the Happiness Gap: Political Institutions and Perceived Quality of Life in Transition. IZA Discussion Paper No. 9484.

Olson, M. (1982). The rise and decline of nations: Economic growth, stagnation, and social rigidities. New Heaven.

Onyx, J., & Bullen, P. (2000). Measuring social capital in five communities. The journal of applied behavioral science, 36(1), 23-42.

Organization for Economic Co-Operation and Development (OECD). (2001). The Well-being of Nations: The Role of Human and Social Capital. Centre for Educational Research and Innovation, 1-119.

Paas, T., & Schlitte, F. (2006). Regional income inequality and convergence processes in the EU-25. HWWA.

Paldam, M. (2000). Social capital: one or many? Definition and measurement. Journal of economic surveys, 14(5), 629-653.

Paldam, M., & Svendsen, G. T. (2000). Missing social capital and the transition in Eastern Europe. Department of Economics, Faculty of Business Administration, Aarhus School of Business.

Rehdanz, K., & Maddison, D. (2005). Climate and happiness. Ecological Economics, 52(1), 111–125.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior research methods, 40(3), 879-891.

Rodriguez-Pose, A. & Ezcurra, R. (2010). Does decentralization matter for regional disparities? A cross-country analysis. Journal of Economic Geography, 10, 619-644.

Rodríguez-Pose, A., & Maslauskaite, K. (2012). Can policy make us happier? Individual characteristics, socio-economic factors and life satisfaction in Central and Eastern Europe. Cambridge Journal of Regions, Economy and Society, 5(1), 77-96.

Rodríguez-Pose, A., & von Berlepsch, V. (2014). Social capital and individual happiness in Europe. Journal of Happiness Studies, 15(2), 357-386.

Putnam, R. D. (2000). Bowling alone: The collapse and revival of American democracy. Bowling alone: the collapse and the revival of american democracy.

Schimmack, U., Schupp, J., & Wagner, G. G. (2008). The influence of environment and personality on the affective and cognitive component of subjective well-being. Social Indicators Research, 89(1), 41-60.

Smyth, R., Mishra, V., & Qian, X. (2008). The environment and well-being in urban China. Ecological Economics, 68(1), 547-555.

Spiekermanm, K., & Wegener, M. (2006). Accessibility and spatial development in Europe. Scienze Regionali, 5(2), 15-46.

Sutton, S. K., & Davidson, R. J. (1997). Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems. Psychological Science, 8(3), 204-210.

UCLA: Statistical Consulting Group. (2016a). Analysis of Multiple Mediators in Stata. Retrieved 26 July, 2016, from http://www.ats.ucla.edu/stat/stata/faq/mulmediation.html

UCLA: Statistical Consulting Group. (2016b). Mediation Analysis with a Categorical Independent Variable in Stata. Retrieved 10 August, 2016, from http://www.ats.ucla.edu/stat/stata/faq/mediation_cativ.htm

United Nations General Assembly Resolution 66/281. (2012). International Day of Happiness. Retrieved 18 August, 2016, from http://www.happinessday.org/resolutions/

Van Oorschot, W., Arts, W., & Gelissen, J. (2006). Social capital in Europe measurement and social and regional distribution of a multifaceted phenomenon. Acta sociologica, 49(2), 149-167.

Van Praag, B. M., Frijters, P., & Ferrer-i-Carbonell, A. (2000). A structural model of well-being: With an application to German data. FEE, Universiteit Van Amsterdam Working Paper.

Veenhoven, R. (2000). The Four Qualities of Life. Ordering Concepts and Measures of the Good Life. Journal Of Happiness Studies, (1) , 1-39.

Veenhoven, R. (2002). Why Social Policy Needs Subjective Indicators. Social Indicators Research, 58, 33-45.

Veenhoven, R. (2004). Happiness as an aim in public policy, in Positive Psychology in Practice. Chapter 39, Editors: Linley, A., Joseph, S. Publisher: John Wiley and Sons, Inc., Hoboken, N.J., USA. ISBN 0471459062.

Ward, G. (2015). Is Happiness a Predictor of Election Results? CEP Discussion Papers, (1343).

Wegener, M., Eskelinnen, H., Fürst, F., Schürmann, C., & Spiekermann, K. (2002). Criteria for the Spatial Differentiation of the EU Territory: Geographical Position. Study Programme on European Spatial Planning, 1-79.

Wiest, M., Schüz, B., Webster, N., & Wurm, S. (2011). Subjective well-being and mortality revisited: differential effects of cognitive and emotional facets of well-being on mortality. Health Psychology, 30(6), 728.

Woolcock, M. (1998). Social capital and economic development: Toward a theoretical synthesis and policy framework. Theory and society, 27(2), 151-208.

Woolcock, M., & Narayan, D. (2000). Social capital: Implications for development theory, research, and policy. The world bank research observer, 15(2), 225-249.

Appendix

Table 1: Mean Values of Social Capital

	TRUST	SOCIALINTERACTION	NORMS
Capital City Regions			
Big city or outskirts of a big city	-0,01	-0,09	-0,01
Small city or town	0,04	0,00	0,00
Rural village or farm in the countryside	0,00	0,01	0,00
Other Regions			
Big city or outskirts of a big city	-0,07	-0,02	-0,01
Small city or town	0,02	0,03	0,02
Rural village or farm in the countryside	0,00	0,01	-0,01
Total Mean	0,00	0,00	0,00

Note: This table reports the mean values of the index variables *TRUST*, *SOCIALINTERACTION* and NORMS. The data is obtained from European Social Survey.

Table 2: Number of Observations per Country for Each Available Year

	Year							
	2002	2004	2006	2008	2010	2012	2014	
Country								Total
BG	0	0	1,400	2,230	2,434	2,260	0	8,324
CZ	1,356	3,026	0	2,018	2,386	2,009	2,148	12,943
HU	1,685	1,498	1,518	1,544	1,561	2,014	1,698	11,518
PL	2,110	1,716	1,721	1,619	1,751	1,898	1,615	12,430
SK	0	1,505	1,748	1,808	1,856	1,847	0	8,764
Total	5,151	7,745	6,387	9,219	9,988	10,028	5,461	53,979

Note: This table presents a summary of the number of observations of level 1 sampling units for each country and each year. The countries in the data sample are Bulgaria (BG), Czech Republic (CZ), Hungary (HU), Poland (PL) and Slovakia (SK). The data is obtained from European Social Survey. Missing yearly data for a particular country is marked with zero.

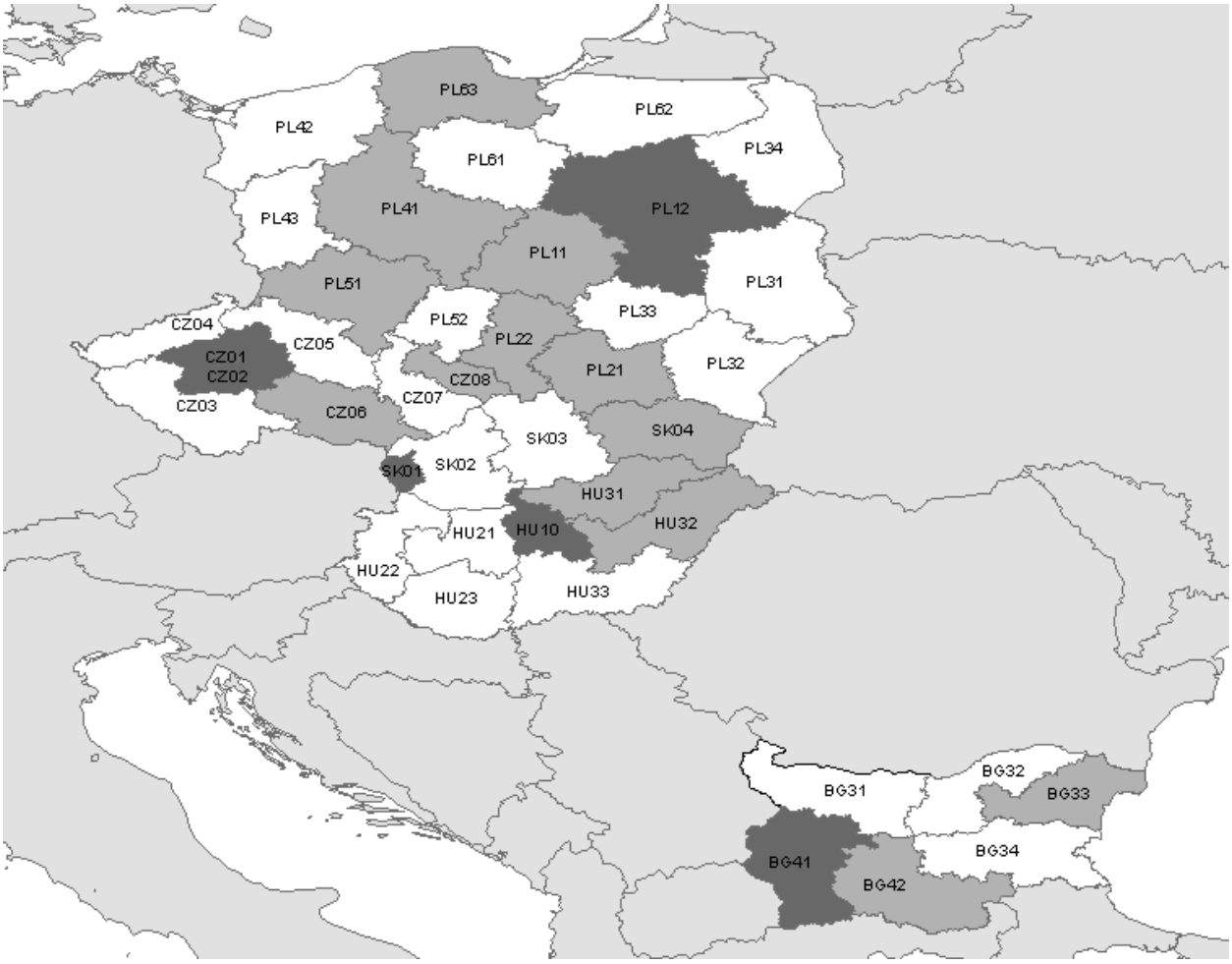
Table 3: NUTS2 Region Classification

		Region Type		
NUTS2 Label	NUTS2 Code	Capital city	Second-tier city	Other
Severozapaden	BG31	0	0	1
Severen tsentralen	BG32	0	0	1
Severoiztochen	BG33	0	1	0
Yugoiztochen	BG34	0	0	1
Yugozapaden	BG41	1	0	0
Yuzhen tsentralen	BG42	0	1	0
Praha	CZ01	1	0	0
Střední Čechy	CZ02	1	0	0
Jihozápad	CZ03	0	0	1
Severozápad	CZ04	0	0	1
Severovýchod	CZ05	0	0	1
Jihovýchod	CZ06	0	1	0
Střední Morava	CZ07	0	0	1
Moravskoslezsko	CZ08	0	1	0
Közép-Magyarország	HU10	1	0	0
Közép-Dunántúl	HU21	0	0	1
Nyugat-Dunántúl	HU22	0	0	1
Dél-Dunántúl	HU23	0	0	1
Észak-Magyarország	HU31	0	1	0
Észak-Alföld	HU32	0	1	0
Dél-Alföld	HU33	0	0	1
Łódzkie	PL11	0	1	0
Mazowieckie	PL12	1	0	0
Małopolskie	PL21	0	1	0
Śląskie	PL22	0	1	0
Lubelskie	PL31	0	0	1
Podkarpackie	PL32	0	0	1
Świętokrzyskie	PL33	0	0	1
Podlaskie	PL34	0	0	1
Wielkopolskie	PL41	0	1	0
Zachodniopomorskie	PL42	0	0	1
Lubuskie	PL43	0	0	1
Dolnośląskie	PL51	0	1	0
Opolskie	PL52	0	0	1
Kujawsko-pomorskie	PL61	0	0	1
Warmińsko-mazurskie	PL62	0	0	1
Pomorskie	PL63	0	1	0
Bratislavský kraj	SK01	1	0	0

Západné Slovensko	SK02	0	0	1
Stredné Slovensko	SK03	0	0	1
Východné Slovensko	SK04	0	1	0

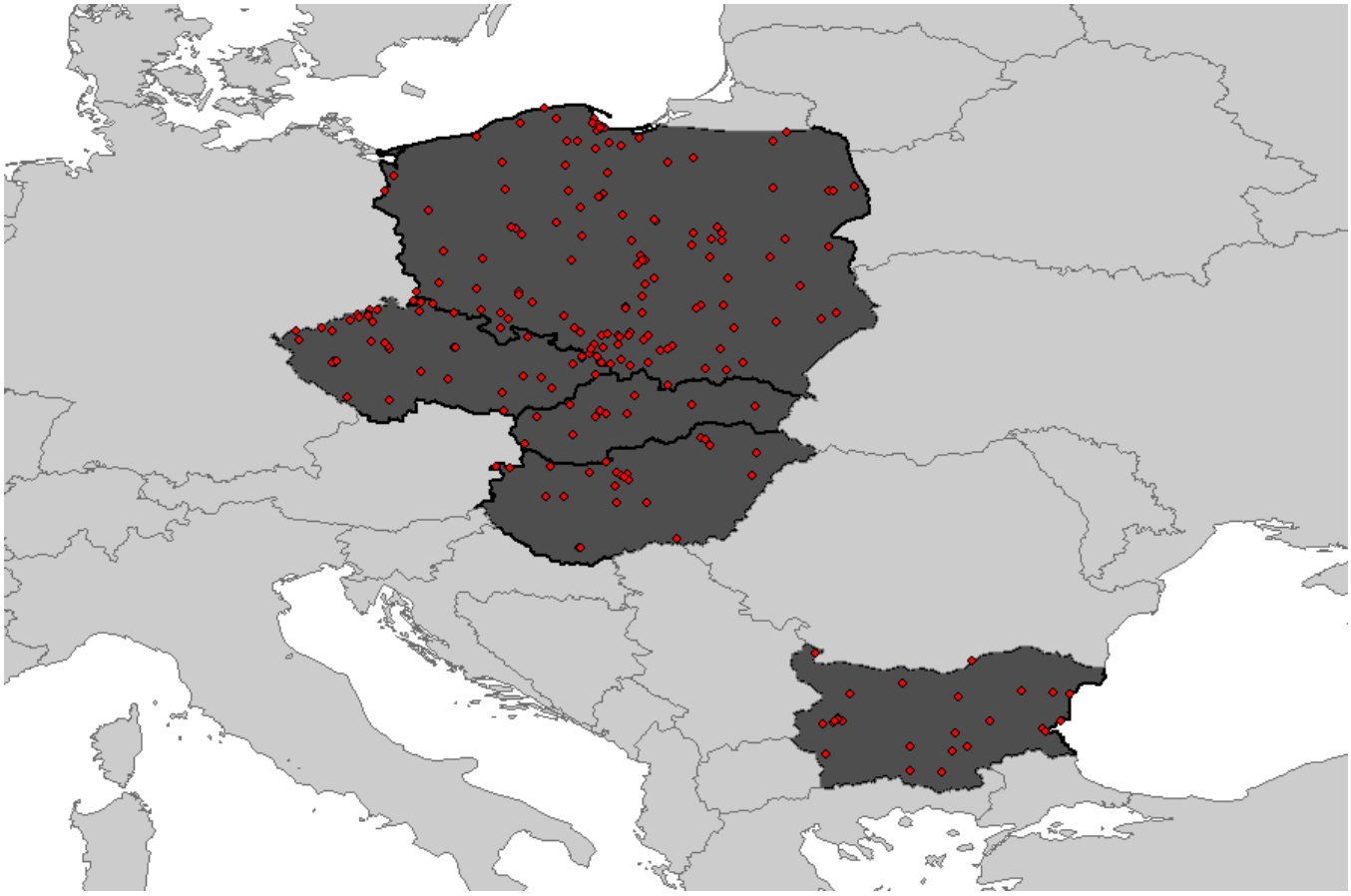
Note: This table presents NUTS2 codes and their corresponding NUTS2 names for each of the 41 NUTS2 regions from the final data sample. Further, each of the regions is classified as a capital city region, second-tier city region or other region according to the classification available in Dogaru et al. (2014, pp. 211). NUTS2 codes and names are matched according to the information available in EUROSTAT.

Figure 1: Mapped NUTS2 Regions According to Their Class



Note: NUTS2 regions which appear in dark grey color are classified as capital city regions. NUTS2 regions which appear in medium grey color are classified as second-tier city regions. Finally, NUTS2 regions appearing in white color are classified as other regions. Each region is assigned to its class accordingly to the classification available in Dogaru et al. (2014, pp. 211). The geo-data was retrieved from GISCO (Geographic Information System of the Commission), EUROSTAT.

Figure 2: Location of SO2 Air Monitoring Stations for 2014



Note: Countries of interest are appearing in dark grey color. Air monitoring stations available in 2014 throughout the countries of interest are marked with red dots. Information regarding stations and their geographical coordinates is retrieved from AirBase. The geo-data used for the creation of this map was retrieved from GISCO (Geographic Information System of the Commission), EUROSTAT.

Table 7: Summary Statistics

Variable	Definition	#Obs	Mean	Std. Dev.	Min	Max
Dependent Variables:						
LIFESAT	Subjective life satisfaction, scale 0-10	53551	5,98	2,51	0,00	10,00
HAPPY	Subjective happiness, scale 0-10	53367	6,45	2,28	0,00	10,00
Independent Variables:						
CAPITAL	Dummy, 1 if region is classified as a capital city region, 0 otherwise	53979	0,20	0,40	0,00	1,00

AREA	Categorical variable indicating place of living within a region , 1 - big city or the outskirts of a big city; 2 - towns or small cities, 3 - country village or farm/home in the countryside	53702	2,05	0,81	1,00	3,00
LOCATION	Categorical variable indicating the place of living.	53702	4,47	1,60	1,00	6,00
Social Capital Indexes:						
TRUST	Index variable measuring trust, which is formed from social trust and institutional trust	49079	-2.74e-10	1,00	-2,48	3,83
SOCIALINTERACTION	Index variable measuring social interactions, which is formed from informal social interactions and formal social interactions	52142	2.38e-09	1,00	-0,59	7,42
NORMS	Index variable measuring norms and sanctions	52024	0,0012	1,00	-2,91	2,61
Regional Social Capital:						
REG_TRUST	Regional mean value of TRUST	53979	-0,0046	0,33	-0,80	1,05
REG_SOCIALINTERACTION	Regional mean value of SOCIALINTERACTION	53979	0,0012	0,18	-0,45	0,60
REG_NORMS	Regional mean value of NORMS	53979	-0,0005	0,23	-0,76	0,88
PPLTRST	“Most people can be trusted in dealing with people”	53664	4,08	2,47	0,00	10,00
PPLHLP	“Would you say that most of the time people try to be helpful or that they are mostly looking out for themselves?”	53538	3,95	2,38	0,00	19,00
PPLFAIR	“Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair?”	53131	4,73	2,38	0,00	10,00
TRUSTPRL	“Please tell me on a score of 0-10 how much you personally trust the legal system?”	52382	3,22	2,45	0,00	10,00
TRUSTLGL	“Please tell me on a score of 0-10 how much you personally trust the legal system?”	51863	3,77	2,56	0,00	10,00
TRUSTPLT	“Please tell me on a score of 0-10 how much you personally trust politicians?”	52533	2,61	2,25	0,00	10,00
TRUSTPLC	“Please tell me on a score of 0-10 how much you personally trust the police?”	52734	4,70	2,63	0,00	10,00
WORKEDPRT	There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you worked in a political party or action group?	53655	0,02	0,15	0,00	1,00

WORKEDORG	There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you worked in worked in another organisation or association?	53606	0,05	0,22	0,00	1,00
CONTACTEDPLT	There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you contacted a politician, government or local government official?	53642	0,10	0,30	0,00	1,00
PETITION	There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you signed a petition?	53533	0,11	0,31	0,00	1,00
BADGE	There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you worn or displayed a campaign badge/sticker?	53584	0,03	0,18	0,00	1,00
STATEHLTH	"Please say what you think overall about the state of health services in [country] nowadays?"	52437	3,99	2,55	0,00	10,00
SOCIAL	"How often do you meet socially with friends, relatives or work colleagues?"	53663	3,34	1,71	0,00	6,00
SAFEDARK	"How safe do you – or would you - feel walking alone in this area after dark?"	52779	1,75	0,74	0,00	3,00
ACHURCH	"Apart from special occasions such as weddings and funerals, about how often do you attend religious services nowadays?"	53488	1,83	1,63	0,00	6,00
TRADEUNION	Are you or have you ever been a member of a trade union or similar organisation? If yes, is that currently or previously?	53507	0,43	0,49	0,00	1,00
Regional Controls:						
GDP	GDP per capita at PPP	53979	15770,79	7698,71	6500,00	47900,00
LNGDP	Log of GDP per capita at PPP	53979	9,57	0,41	8,78	10,78
NUTS2UNEMPL	Regional unemployment % 20-64 yrs	53979	9,97	4,88	42614,00	42608,00
GINI	Country-level Gini coefficient	53979	29,38	3,80	42393,00	36,20
HAPPYINEQ		53979	2,17	0,28	1,60	2,83
LIFESATINEQ		53979	2,36	0,23	1,49	3,09
POPDENSITY	Total regional population divided by the total regional area in km2	53979	199,34	394,49	44,50	2579,70

SO2EMISSION	Regional air pollution with sulfur dioxide	53979	9,46	5,00	1,32	34,75
AIRACCESSIBILITY		53979	73,01	31,26	23,70	160,10
HDD	Mean Heating Degree Days by NUTS2 for the period 1980-2009	53979	3238,70	377,31	2382,49	3908,20
Socio-Demographic Controls:						
AGE	Age of respondent, calculated	53609	47,75	18,23	14,00	97,00
AGE2	AGE*AGE	53609	2612,46	1796,45	196,00	9409,00
INCOME	Household's total net income, categorical variable: 1 being the lowest and 10 being the highest income; 11 - refusal+no answer; 12 - don't know	46738	6,36	3,83	1,00	12,00
UNEMPL	Dummy 1 if unemployed and seeking work	53979	0,06	0,23	0,00	1,00
UNEMPL2	Dummy 1 if unemployed and NOT seeking work	53979	0,03	0,16	0,00	1,00
EDUC	Years of full-time education completed	53256	12,10	3,29	0,00	42,00
HOUSEHOLD	Number of people living regularly as member of household	53919	2,93	1,51	1,00	15,00
MARITAL	Marital status, recoded	53372	2,40	0,83	1,00	4,00
MALE	Dummy 1-male 0-female	53875	0,46	0,50	0,00	1,00
HEALTH	Subjective general health, scale 0-4	53867	2,58	0,96	0,00	4,00
LEFTRIGHT	Placement on left right scale, recoded 4 categories: 1 - left, 2 - middle, 3 - right, 4 - don't know	53227	2,37	0,95	1,00	4,00
IDNO	Respondent's identification number	53979	17245,01	43078,48	1	225010
YEAR	Year of the observation	53979	2008,34	3,67	2002	2014
nutsID	NUTS2 region id	53979	19,84	12,18	1,00	41,00
cntryID	Country id	53979	3,01	1,32	1,00	5,00
DWEIGHT	Design weight	53979	1,00	0,45	0,00	4,08
PWEIGHT	Population size weight	53979	0,72	0,63	0,25	2,25
PSPWGHT	Post-stratification weight including design weight	53979	1,01	0,55	0,00	4,32
WEIGHT	PSPWGHT*PWEIGHT	53979	0,72	0,70	0,00	4,67

Note: This table reports the summary statistics. The main data sources are European Social Survey, EUROSTAT, The World Bank, AirBase and ESPON. The time period is from 2002 to 2014.

Table 8: Regression Results for Happiness

	(1) HAPPY	(2) HAPPY	(3) HAPPY	(4) HAPPY
LOCATION				
1. Capital city region - Big city/outskirts of big city	Reference	Reference	Reference	Reference
2. Capital city region - Small city or	-0.142	-0.100	0.048	0.052

town				
	(0.112)	(0.085)	(0.109)	(0.090)
3. Capital city region - Rural village or farm in the countryside	-0.311	-0.269	-0.019	-0.062
	(0.211)	(0.189)	(0.197)	(0.146)
4. Other region - Big city/outskirts of big city	**-0.237****	0.052	0.157	0.189
	(0.095)	(0.168)	(0.159)	(0.132)
5. Other region - Small city or town	**-0.278****	0.018	0.173	0.152
	(0.130)	(0.190)	(0.160)	(0.136)
6. Other region - Rural village or farm in the countryside	-0.269	0.031	0.256	0.158
	(0.189)	(0.244)	(0.182)	(0.165)
Regional Characteristics:				
Air Pollution		-0.008*	-0.005**	-0.003
		(0.005)	(0.002)	(0.003)
Heating Degree Days		0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
SWB Inequality		-0.550***	-0.450***	-0.387***
		(0.023)	(0.024)	(0.022)
Log GDP per Capita (PPP)		0.503***	0.349*	0.423**
		(0.188)	(0.188)	(0.179)
GINI Coefficient		-0.010	-0.021	-0.012
		(0.034)	(0.023)	(0.024)
Regional Unemployment Rate		-0.040***	-0.032***	-0.025***
		(0.008)	(0.006)	(0.006)
Population Density		-0.000***	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)
Region Accessibility by Air		-0.003	-0.003**	-0.003**
		(0.002)	(0.001)	(0.001)
Socio-Demographic Controls:				
Age			-0.087***	-0.080***
			(0.006)	(0.007)
Age^2			0.001***	0.001***
			(0.000)	(0.000)
Unemployed and Seeking Work (1=Yes)			-0.590***	-0.543***
			(0.032)	(0.022)
Years of Education Completed			0.028**	0.023**
			(0.013)	(0.011)
Household composition			-0.015	-0.019
			(0.012)	(0.014)
Marital Status:				
1. Divorced/Separated			Reference	Reference
2. Married/Civil Partnership			0.915***	0.879***
			(0.195)	(0.210)
3. Never married			0.256**	0.194*
			(0.108)	(0.105)
4. Widowed			0.224**	0.188
			(0.107)	(0.120)

Dummy Male (1=Yes)	-0.245***	-0.270***
	(0.027)	(0.014)
Subjective Health	0.720***	0.627***
	(0.013)	(0.015)
Political spectrum:		
1. Being in the left	Reference	Reference
2. Being in the centre	0.064	0.033
	(0.072)	(0.058)
3. Being in the right	0.289***	0.209***
	(0.080)	(0.064)
4. Don't know	-0.032	0.039
	(0.080)	(0.064)
Household income:		
1. Level 1	Reference	Reference
2. Level 2	0.141	0.135
	(0.116)	(0.118)
3. Level 3	0.380***	0.337***
	(0.081)	(0.088)
4. Level 4	0.503***	0.443***
	(0.096)	(0.099)
5. Level 5	0.535***	0.486***
	(0.091)	(0.091)
6. Level 6	0.570***	0.509***
	(0.090)	(0.093)
7. Level 7	0.555***	0.475***
	(0.142)	(0.145)
8. Level 8	0.579***	0.489***
	(0.105)	(0.108)
9. Level 9	0.713***	0.604***
	(0.087)	(0.100)
10. Level 10	0.916***	0.779***
	(0.103)	(0.104)
11. Refusal/No Answer	0.523***	0.454***
	(0.144)	(0.134)
12. Don't know	0.550***	0.534***
	(0.094)	(0.100)
Social Capital Proxies:		
Social and Institutional Trust (***TRUST***)		0.294***
		(0.021)
Regional mean of TRUST		-0.168**
		(0.080)
Informal and Formal Social		0.017**
Interactions (***SOCIALINTERACTION***)		(0.008)
		0.074
Regional mean of SOCIALINTERACTION		(0.174)
		0.302***
Norms & Sanctions (***NORM***)		(0.023)
		-0.013

Regional mean of NORM				(0.049)
Time FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Observations	38,214	38,214	38,214	38,214
Number of countries	5	5	5	5
Number of NUTS2 regions	41	41	41	41
Loglikelihood	-65448.07	-65347.43	-62397.79	-61693.97
AIC	130934.1	130748.9	124897.6	123501.9
BIC	131096.6	130979.7	125333.7	123989.3

Note: This table reports the results of the 3-level hierarchical regression analysis. Model 1 tests hypothesis 1, with only *LOCATION*, time fixed effects and country fixed effects included. Hypotheses 2 and 3 are tested in models 3 and 4, respectively. *LOCATION* is the main variable of interest. It has 6 categories. Category 1 indicates a place of living in a big city or the outskirts of a big city in capital city regions. This is the reference category. Category 2 indicates a place of living in a small city or a town in capital city regions. Category 3 indicates a place of living in a country village or a farm/home in the countryside in capital city regions. Category 4 indicates a place of living in a big city or the outskirts of a big city in other regions. Category 5 indicates a place of living in a small city or a town in other regions. Category 6 indicates a place of living in a country village or a farm/home in the countryside in other regions. Further, the dependent variable is *HAPPY*. It indicates the individual level of happiness.. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 11: Regression Results for Life Satisfaction

	(5) LIFESAT	(6) LIFESAT	(7) LIFESAT	(8) LIFESAT
LOCATION				
1. Capital city region - Big city/outskirts of big city	Reference	Reference	Reference	Reference
2. Capital city region - Small city or town	-0.153	-0.116	0.094	0.106
	(0.189)	(0.159)	(0.171)	(0.148)
3. Capital city region - Rural village or farm in the countryside	-0.327	-0.291	0.062	0.018
	(0.289)	(0.265)	(0.264)	(0.206)
4. Other region - Big city/outskirts of big city	**-0.375*** **	-0.119	-0.034	0.053
	(0.099)	(0.143)	(0.132)	(0.124)
5. Other region - Small city or town	**-0.303***	-0.032	0.129	0.151
	(0.164)	(0.177)	(0.136)	(0.119)
6. Other region - Rural village or farm in the countryside	-0.276	-0.002	**0.263***	**0.201**
	(0.218)	(0.229)	(0.145)	(0.130)
Regional Characteristics:				
Air Pollution		-0.002	-0.000	0.001
		(0.004)	(0.001)	(0.003)
Heating Degree Days		-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
SWB Inequality		-0.679***	-0.525***	-0.483***
		(0.066)	(0.058)	(0.069)
Log GDP per Capita (PPP)		0.283***	0.044	0.138

	(0.103)	(0.117)	(0.126)
GINI Coefficient	-0.037*	-0.063***	-0.053***
	(0.022)	(0.010)	(0.014)
Regional Unemployment Rate	-0.043***	-0.032***	-0.028***
	(0.008)	(0.005)	(0.004)
Population Density	-0.000***	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)
Region Accessibility by Air	-0.001	-0.002*	-0.002
	(0.001)	(0.001)	(0.001)
Socio-Demographic Controls:			
Age		-0.115***	-0.107***
		(0.008)	(0.009)
Age^2		0.001***	0.001***
		(0.000)	(0.000)
Unemployed and Seeking Work (1=Yes)		-0.868***	-0.812***
		(0.049)	(0.038)
Years of Education Completed		0.040**	0.033**
		(0.016)	(0.014)
Household composition		-0.049***	-0.053***
		(0.010)	(0.008)
Marital Status:			
1. Divorced/Separated		Reference	Reference
2. Married/Civil Partnership		0.886***	0.842***
		(0.199)	(0.221)
3. Never married		0.379**	0.306**
		(0.154)	(0.155)
4. Widowed		0.438***	0.394***
		(0.100)	(0.117)
Dummy Male (1=Yes)		-0.206***	-0.230***
		(0.032)	(0.018)
Subjective Health		0.733***	0.625***
		(0.009)	(0.014)
Political spectrum:			
1. Being in the left		Reference	Reference
2. Being in the centre		0.246***	0.209***
		(0.076)	(0.061)
3. Being in the right		0.609***	0.513***
		(0.086)	(0.070)
4. Don't know		0.169**	0.257***
		(0.074)	(0.054)
Household income:			
1. Level 1		Reference	Reference
2. Level 2		0.322***	0.313***
		(0.086)	(0.086)
3. Level 3		0.606***	0.554***
		(0.050)	(0.053)
4. Level 4		0.774***	0.703***
		(0.046)	(0.050)
5. Level 5		0.841***	0.781***

			(0.049)	(0.047)
6. Level 6			0.919***	0.844***
			(0.061)	(0.060)
7. Level 7			0.936***	0.840***
			(0.067)	(0.068)
8. Level 8			1.031***	0.920***
			(0.093)	(0.090)
9. Level 9			1.061***	0.927***
			(0.063)	(0.067)
10. Level 10			1.420***	1.257***
			(0.129)	(0.114)
11. Refusal/No Answer			0.837***	0.752***
			(0.113)	(0.101)
12. Don't know			0.892***	0.872***
			(0.079)	(0.074)
Social Capital Proxies:				
Social and Institutional Trust (***TRUST***)				0.382***
				(0.038)
Regional mean of TRUST				-0.256**
				(0.128)
Informal and Formal Social				0.023***
Interactions (***SOCIALINTERACTION***)				(0.004)
				-0.049
Regional mean of SOCIALINTERACTION				(0.137)
				0.334***
Norms & Sanctions (***NORM***)				(0.024)
				-0.062
Regional mean of NORM				(0.147)
Time FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Observations	38,385	38,385	38,385	38,385
Number of countries	5	5	5	5
Number of NUTS2 regions	41	41	41	41
Loglikelihood	-69012.71	-68898.91	-65991.02	-65169.02
AIC	138063.4	137851.8	132084	130452
BIC	138226	138082.8	132520.4	130939.7

Note: This table reports the results of the 3-level hierarchical regression analysis. Model 5 tests hypothesis 1, with only *LOCATION*, time fixed effects and country fixed effects included. Hypotheses 2 and 3 are tested in models 7 and 8, respectively. *LOCATION* is the main variable of interest. It has 6 categories. Category 1 indicates a place of living in a big city or the outskirts of a big city in capital city regions. This is the reference category. Category 2 indicates a place of living in a small city or a town in capital city regions. Category 3 indicates a place of living in a country village or a farm/home in the countryside in capital city regions. Category 4 indicates a place of living in a big city or the outskirts of a big city in other regions. Category 5 indicates a place of living in a small city or a town in other regions. Category 6 indicates a place of living in a country village or a farm/home in the countryside in other regions. Further, the dependent variable is *LIFESAT*. It indicates the individual level of life satisfaction. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 15: Robustness Check for Happiness

	(9) HAPPY	(10) HAPPY	(11) HAPPY	(12) HAPPY
LOCATION				
1. Capital city region - Big city/outskirts of big city	Reference	Reference	Reference	Reference
2. Capital city region - Small city or town	**-0.070***	-0.052	0.013	0.015
	(0.038)	(0.042)	(0.042)	(0.043)
3. Capital city region - Rural village or farm in the countryside	**-0.147*****	**-0.133*****	-0.027	-0.051
	(0.039)	(0.042)	(0.041)	(0.042)
4. Other region - Big city/outskirts of big city	**-0.098*****	0.018	0.055	0.075
	(0.026)	(0.054)	(0.054)	(0.055)
5. Other region - Small city or town	**-0.107*****	0.011	0.070	0.063
	(0.024)	(0.055)	(0.055)	(0.057)
6. Other region - Rural village or farm in the countryside	**-0.105*****	0.017	**0.107***	0.062
	(0.024)	(0.055)	(0.056)	(0.057)
Regional Characteristics:				
Air Pollution		-0.002	-0.002	0.000
		(0.002)	(0.002)	(0.002)
Heating Degree Days		-0.000	-0.000	-0.000*
		(0.000)	(0.000)	(0.000)
SWB Inequality		-0.191***	-0.167***	-0.134***
		(0.047)	(0.047)	(0.050)
Log GDP per Capita (PPP)		0.201***	0.146***	0.179***
		(0.053)	(0.053)	(0.055)
GINI Coefficient		-0.008	-0.015**	-0.009
		(0.006)	(0.006)	(0.007)
Regional Unemployment Rate		-0.016***	-0.013***	-0.011***
		(0.003)	(0.003)	(0.003)
Population Density		-0.000***	-0.000**	-0.000**
		(0.000)	(0.000)	(0.000)
Region Accessibility by Air		-0.001	-0.001**	-0.002**
		(0.001)	(0.001)	(0.001)
Socio-Demographic Controls:				
Age			-0.047***	-0.045***
			(0.003)	(0.003)
Age^2			0.000***	0.000***
			(0.000)	(0.000)
Unemployed and Seeking Work (1=Yes)			-0.300***	-0.282***
			(0.033)	(0.033)
Years of Education Completed			0.012***	0.009***
			(0.003)	(0.003)
Household composition			-0.008	-0.010*
			(0.006)	(0.006)

Marital Status:		
1. Divorced/Separated	Reference	Reference
2. Married/Civil Partnership	0.475***	0.468***
	(0.029)	(0.029)
3. Never married	0.105***	0.077**
	(0.034)	(0.034)
4. Widowed	0.109***	0.093**
	(0.039)	(0.039)
Dummy Male (1=Yes)	-0.138***	-0.154***
	(0.015)	(0.015)
Subjective Health	0.376***	0.337***
	(0.011)	(0.011)
Political spectrum:		
1. Being in the left	Reference	Reference
2. Being in the centre	0.026	0.011
	(0.021)	(0.022)
3. Being in the right	0.147***	0.109***
	(0.023)	(0.023)
4. Don't know	-0.007	0.029
	(0.030)	(0.030)
Household income:		
1. Level 1	Reference	Reference
2. Level 2	0.040	0.040
	(0.044)	(0.044)
3. Level 3	0.148***	0.132***
	(0.042)	(0.042)
4. Level 4	0.208***	0.184***
	(0.042)	(0.043)
5. Level 5	0.236***	0.216***
	(0.045)	(0.045)
6. Level 6	0.250***	0.226***
	(0.047)	(0.047)
7. Level 7	0.241***	0.207***
	(0.048)	(0.049)
8. Level 8	0.257***	0.216***
	(0.049)	(0.049)
9. Level 9	0.324***	0.276***
	(0.049)	(0.050)
10. Level 10	0.442***	0.380***
	(0.050)	(0.050)
11. Refusal/No Answer	0.238***	0.209***
	(0.047)	(0.047)
12. Don't know	0.249***	0.249***
	(0.042)	(0.043)
Social Capital Proxies:		
Social and Institutional Trust (***TRUST***)		0.152***
		(0.008)
Regional mean of TRUST		-0.054
		(0.041)
Informal and Formal Social		0.010

Interactions (***SOCIALINTERACTION***)				(0.007)
				0.086
Regional mean of SOCIALINTERACTION				(0.060)
				0.158***
Norms & Sanctions (***NORM***)				(0.009)
				0.001
Regional mean of NORM				(0.059)
Time FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Observations	38,214	38,214	38,214	38,214
Number of countries	5	5	5	5
Number of NUTS2 regions	41	41	41	41
Loglikelihood	-62229.11	-62136.41	-59415.02	-58783.51
AIC	124508.2	124338.8	118944	117693
BIC	124722	124621	119431.4	118231.7

Note: This table reports the results of the ordered probit regression analysis. Model 9 tests hypothesis 1, with only *LOCATION*, time fixed effects and country fixed effects included. Hypotheses 2 and 3 are tested in models 11 and 12, respectively. *LOCATION* is the main variable of interest. It has 6 categories. Category 1 indicates a place of living in a big city or the outskirts of a big city in capital city regions. This is the reference category. Category 2 indicates a place of living in a small city or a town in capital city regions. Category 3 indicates a place of living in a country village or a farm/home in the countryside in capital city regions. Category 4 indicates a place of living in a big city or the outskirts of a big city in other regions. Category 5 indicates a place of living in a small city or a town in other regions. Category 6 indicates a place of living in a country village or a farm/home in the countryside in other regions. Further, the dependent variable is *HAPPY*. It indicates the individual level of happiness.. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 17: Robustness Check for Life Satisfaction

	(13) LIFESAT	(14) LIFESAT	(15) LIFESAT	(16) LIFESAT
LOCATION				
1. Capital city region - Big city/outskirts of big city	Reference	Reference	Reference	Reference
2. Capital city region - Small city or town	**-0.063***	-0.051	0.042	0.043
	(0.038)	(0.042)	(0.042)	(0.042)
3. Capital city region - Rural village or farm in the countryside	**-0.124****	**-0.111****	0.034	0.010
	(0.040)	(0.043)	(0.044)	(0.044)
4. Other region - Big city/outskirts of big city	**-0.139****	-0.054	-0.017	0.015
	(0.026)	(0.054)	(0.054)	(0.055)
5. Other region - Small city or town	**-0.099****	-0.009	0.062	0.067
	(0.024)	(0.054)	(0.055)	(0.056)
6. Other region - Rural village or farm in the countryside	**-0.085****	0.004	**0.121****	0.087
	(0.024)	(0.055)	(0.056)	(0.057)
Regional Characteristics:				

Air Pollution	-0.001	-0.000	0.001
	(0.002)	(0.002)	(0.002)
Heating Degree Days	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
SWB Inequality	-0.222***	-0.181***	-0.164***
	(0.045)	(0.046)	(0.051)
Log GDP per Capita (PPP)	0.099*	0.013	0.041
	(0.052)	(0.053)	(0.054)
GINI Coefficient	-0.017***	-0.031***	-0.027***
	(0.006)	(0.006)	(0.006)
Regional Unemployment Rate	-0.018***	-0.016***	-0.014***
	(0.003)	(0.003)	(0.003)
Population Density	-0.000***	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Region Accessibility by Air	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
Socio-Demographic Controls:			
Age		-0.055***	-0.053***
		(0.003)	(0.003)
Age^2		0.001***	0.001***
		(0.000)	(0.000)
Unemployed and Seeking Work (1=Yes)		-0.382***	-0.368***
		(0.035)	(0.035)
Years of Education Completed		0.016***	0.013***
		(0.003)	(0.003)
Household composition		-0.023***	-0.025***
		(0.006)	(0.006)
Marital Status:			
1. Divorced/Separated		Reference	Reference
2. Married/Civil Partnership		0.410***	0.401***
		(0.029)	(0.030)
3. Never married		0.156***	0.127***
		(0.035)	(0.035)
4. Widowed		0.190***	0.174***
		(0.039)	(0.040)
Dummy Male (1=Yes)		-0.102***	-0.117***
		(0.015)	(0.015)
Subjective Health		0.348***	0.307***
		(0.011)	(0.011)
Political spectrum:			
1. Being in the left		Reference	Reference
2. Being in the centre		0.100***	0.086***
		(0.022)	(0.022)
3. Being in the right		0.276***	0.239***
		(0.023)	(0.023)
4. Don't know		0.086***	0.130***
		(0.030)	(0.030)
Household income:			
1. Level 1		Reference	Reference

2. Level 2			0.121***	0.123***
			(0.043)	(0.044)
3. Level 3			0.242***	0.227***
			(0.042)	(0.042)
4. Level 4			0.310***	0.288***
			(0.042)	(0.043)
5. Level 5			0.357***	0.340***
			(0.045)	(0.045)
6. Level 6			0.389***	0.367***
			(0.047)	(0.047)
7. Level 7			0.389***	0.356***
			(0.048)	(0.049)
8. Level 8			0.436***	0.397***
			(0.048)	(0.048)
9. Level 9			0.455***	0.406***
			(0.050)	(0.051)
10. Level 10			0.636***	0.577***
			(0.051)	(0.051)
11. Refusal/No Answer			0.359***	0.332***
			(0.047)	(0.047)
12. Don't know			0.380***	0.383***
			(0.042)	(0.042)
Social Capital Proxies:				
Social and Institutional Trust (***TRUST***)				0.177***
				(0.009)
Regional mean of TRUST				-0.093**
				(0.041)
Informal and Formal Social				0.013*
Interactions (***SOCIALINTERACTION***)				(0.007)
				-0.014
Regional mean of SOCIALINTERACTION				(0.061)
				0.160***
Norms & Sanctions (***NORM***)				(0.009)
				-0.040
Regional mean of NORM				(0.058)
Time FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Observations	38,385	38,385	38,385	38,385
Number of countries	5	5	5	5
Number of NUTS2 regions	41	41	41	41
Loglikelihood	-65524.51	-65418.33	-62724.46	-61984.93
AIC	131099	130902.7	125562.9	124095.9
BIC	131312.9	131185	126050.6	124634.9

Note: This table reports the results of the ordered probit regression analysis. Model 13 tests hypothesis 1, with only *LOCATION*, time fixed effects and country fixed effects included. Hypotheses 2 and 3 are tested in models 15 and 16, respectively. *LOCATION* is the main variable of interest. It has 6 categories. Category 1 indicates a place of living in a big city or the outskirts of a big city in capital city regions. This is the reference category. Category 2 indicates a place of living in a small city or a town in capital city regions. Category 3 indicates a place of living in a country village or a farm/home in the countryside in capital city regions. Category 4 indicates a place of living in a

big city or the outskirts of a big city in other regions. Category 5 indicates a place of living in a small city or a town in other regions. Category 6 indicates a place of living in a country village or a farm/home in the countryside in other regions. Further, the dependent variable is *LIFESAT*. It indicates the individual level of life satisfaction. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1