



Kent Academic Repository

Mališauskaite, Gintare (2018) *East versus West: did Communist Regimes matter in the Long-run? Essays on the Comparative Economics of the Former Eastern Bloc Countries*. Doctor of Philosophy (PhD) thesis, University of Kent,.

Downloaded from

<https://kar.kent.ac.uk/71525/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.22024/UniKent/01.02.71525>

This document version

Publisher pdf

DOI for this version

Licence for this version

UNSPECIFIED

Additional information

Chapter II has been removed since it was published in the Journal of Comparative Economics: Mališauskaite, G., Klein, A. (2018). Drinking Under Communism: why do Alcohol Consumption Habits in Eastern Europe Differ from the West in the Long-Run? Journal of Comparative Economics, 46(3), 821-837 Available on KAR: <https://kar.kent.ac.uk/69085/>

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

**East versus West: did Communist Regimes matter in the Long-run?
Essays on the Comparative Economics of the Former Eastern Bloc
Countries**

A thesis for a degree of
Doctor of Philosophy in Economics

Gintarė Mališauskaitė

University of Kent
School of Economics

December 2018

Word Count:

43,919

Acknowledgements

It comes as no surprise that the writing of a PhD thesis can be a lonely process. However, it is not stressed enough that frequently it would not start, continue or end without the encouragement and help from some inspiring people. I am very grateful to all of the School of Economics lecturers and staff members who encouraged and supported this endeavour, who sometimes went out of their way to help and gave a smile and a thumbs-up just when it was needed. I am especially appreciative of the help and guidance from Dr. Alex Klein, Dr. Olena Nizalova and Professor Miguel León-Ledesma. This endeavour would not have begun or come to fruition without you. I am also indebted to my family and friends, who believed in me every step of the way and too frequently had to suffer my emotional roller coaster, I love you all.

Table of contents

Introduction	5
Chapter I: Comparing Eastern and Western Europe: has Communism succeeded in increasing Educational Attainments?	11
Abstract.....	11
1.1 Introduction.....	12
1.2 Previous literature.....	14
1.3 Methodology.....	17
1.4 Data and descriptive statistics.....	21
1.5 Results.....	24
1.5.1 Secondary education.....	24
1.5.2 Choice between vocational and university preparation secondary education.....	26
1.5.3 Tertiary education.....	27
1.5.4 Primary education.....	29
1.6 Discussion and conclusion.....	30
References.....	32
Tables	34
Appendix.....	45
Chapter II: Drinking under Communism: why do Alcohol Consumption Habits in Eastern Europe Differ from the West in the Long-Run?	46
Abstract.....	46
2.1 Introduction.....	47
2.2 Literature review.....	48
2.3 Methodology.....	51
2.4 Data and descriptive statistics.....	54
2.5 Regression results.....	56
2.5.1 Communism effect on female drinking habits.....	56
2.5.2 Communism effect on male drinking habits.....	58
2.5.3 Gender comparisons.....	59
2.5.4 Robustness checks.....	61

2.5.5 Changes in drinking habits after the collapse of the communism.....	62
2.6 Discussion and conclusion.....	63
References.....	65
Tables and graphs.....	67
Appendix.....	79
Chapter III: Do People from Eastern Bloc Perceive their Health Differently?	
Determinants and Impact of Health Perception Gap.....	81
Abstract.....	81
3.1 Introduction.....	82
3.2 Literature review.....	83
3.3 Methodology.....	89
3.4 Data and descriptive statistics.....	95
3.5 Results.....	98
3.5.1 Determinants of Self-rated Health and influence of Communist regime.....	98
3.5.2 Relationship between the Communist Regime and Objective Health	
Indicators.....	99
3.5.3 Health Perception Gap: Determinants and Importance.....	100
3.6 Discussion and conclusion.....	103
References.....	106
Tables.....	109
Appendix.....	127
Conclusion.....	133

Introduction

The purpose of this thesis is to contribute to the strand of research regarding the effects of communist regimes in former Eastern Bloc. We explored the areas that were likely to be affected at the time of the communist regimes in Eastern Europe and to test if there are any signs of the impact extending to long-term. This is especially important due to the data scarcity and heavy censorship during the period of the Soviet Union, which does not allow to reliably estimate the contemporaneous effects of the regime. Therefore, the main research questions are:

- Did communist regime education policies create systematic differences in educational attainment levels in comparison to the rest of Europe? Increasing educational attainment level within Eastern Bloc was an important political goal. This is partially due to an aim of developing skilled labour force to man factories and contribute towards advancements in science and technology in the cold war competition with the West, and, in part, due to being means to implant and propagate the regime's ideology.

- Is the popular stereotype of excessive alcohol consumption in communist Eastern Bloc reflected in the behaviour of those who lived through the regime even after its collapse? Alcohol was anecdotally known to be a popular companion for recreation in Eastern Bloc, the idea frequently found in Russian cinematography of the day. Any signs of systematically more frequent or larger intake of alcohol after the collapse of the regime would indicate a combination of at least some or all listed reasons: spreading of cultural drinking norms, drinking preferences becoming habitual, and alcohol being as a coping mechanism for experienced trauma. It is likely the list could be extended by more possible explanations.

- Is living under one of the communist regimes in Eastern Bloc significantly related to any long-term differences in health outcomes in comparison to the rest of Europe? In addition, is there a difference in perception of own health? This could help address a question if the communist regime had an impact on health and perception of people who experienced and survived it. Measuring differences in perception is particularly interesting since perception latently affects behaviour and choices of economic agents.

Each question is addressed in a separate chapter, but the overarching questions are: do educational attainment level, alcohol drinking patterns, health and its perception show signs of the communist regime in Eastern Bloc having a long-term impact? How much, on average, an experience of this regime could contribute to changing and shaping cultural norms, agents' choices, behaviour and perceptions? Answers to these questions would

contribute to the knowledge about measuring impacts of historical experiences, could inspire further research, and potentially could be taken into account when modelling and predicting agents' behaviour.

There are multiple reasons making us believe that a long-term impact of the communist regime exists. The communist regimes in Eastern Europe came as an aftermath of the Second World War, which divided Europe geographically, economically and politically. The ideology of the regime was driven by Vladimir Lenin's lofty and revolutionary interpretation of the philosophical ideas by Karl Marx. Egalitarianism, exposure and investigation of the struggles of capitalism, and prediction of the end of this economic and political system still inspire many up to this day. However, the decision to embark on the pursuit of communism through socialism, starting with the Russian Revolution, followed by the formation of the Soviet Union, the abrupt implementation of the communist ideals, and imposition of communist regimes in Eastern Bloc was much less romantic. Not only communism was highly centralized, came with severely restricted migration outside the Eastern Bloc, rigid censorship, ideological and political propaganda, crippling 'five-year' economic plans and different types of punishments for not complying with the ideology and inexplicit rules of the regime, but it also coincided with an economic post-war hardship and the cold war period. The need to strengthen the economies was obvious, but it was not the main focus of the communist regime. It was considered that by pursuing the communist ideals the economic problems will naturally disappear, as then people would live in a more equal and helpful society, not driven by profits or power struggles. The distinctiveness of this approach, seclusion, and between 40 and 60 years of duration of the communism experiment in Eastern Bloc was bound to leave a mark in the societies that experienced it.

The direct economic consequences of the regime were evident during the transition period. However, it is unlikely the consequences were limited only to economic instability and short-term influence. There is evidence showing long-term effect on preferences and cultural norms due to the communist regime experience in Eastern Europe (Roland 2010; Alesina and Fuchs-Schündeln 2007)¹. This supports the idea that an experience of this political and economic regime could have a long-term impact and affected more areas of socio-economic life: social and human capital, which would result in different preferences,

¹ Roland, G. 2010. The Long-Run Weight of Communism or the Weight of Long-Run History? *Wider Working Papers* [Online] 2010/83;
Alesina, A. and Fuchs-Schündeln, N. 2007. Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences. *American Economic Review*, 97(4), 1507-1528.

behaviour and cultural norms at a country level. The impacts of the communist regimes in former Eastern Bloc and long-term influence on present-day post-communist countries are under-researched and deserve more attention. In addition, this provides a rare opportunity to explore how socio-economic life and changes in it could influence society in the long run. This could reveal some historical lessons to be learned and better explain cultural differences between post-soviet and other European countries.

It is important to acknowledge that former Eastern Bloc was created out of economically and culturally diverse countries. There is little data available allowing their statistical comparison prior to World War II. The GDP per capita estimates from Maddison Historical Statistics Project² show that most Eastern Bloc countries were significantly lagging behind economically, in comparison to the West. However, there were examples like Hungary and Czechoslovakia that were comparative to Ireland, Italy, Finland and Austria at the time, both mentioned countries being significantly above USSR average. In addition, there were significant cultural differences between Eastern Bloc countries. One of them being division in religion: Central European countries with dominant Roman Catholic and Protestant churches, while South-Eastern European countries with dominant Eastern Orthodox churches. Furthermore, multiple language groups were and still are found within Eastern Bloc: Germanic, Slavic, Romance, Baltic, and Finno-Ugric. Not to forget, a North-South divide determining work patterns and cuisine. Even after the communist regime was implemented, it consisted of different communist regimes, which varied in strictness of the regime rules and autonomy levels.

The communist regime attempted to address and even out these original country differences through development of soviet ideology and its dissemination, propaganda, endeavour to eliminate religion, strip off the importance of ethnic heritage and unite Eastern Bloc people under a shared identity – single soviet nation. Despite such attempts mostly being not welcomed by Eastern Bloc countries, there are some noticeable differences between former soviet republics and other European countries after the collapse of the regime. Inefficiency of the communist economic policy: heavy industrialisation, agricultural collectivization, five-year plans, led to a quick abandonment of these policies after the collapse of the regime and sent the former Eastern Bloc countries into an economic decline

² Real GDP per capita estimates of the year 1937 in 2011 US\$, 2011 benchmark, latest estimates (2018). The project estimated various GDP measures for a number of countries in the world throughout history, most estimates being available starting with 19th century. More information available at: <https://www.rug.nl/ggdc/historicaldevelopment/maddison/>

during the transition period. This could explain subsequent Eastern European migration towards Western Europe in large numbers, which included a high proportion of highly skilled labour force. Yet despite the apparent attempt to cut all the ties with communism, research shows that redistribution preferences among the Eastern Europeans are among the highest and tolerated wage inequality is much lower compared to other European or Anglo-American countries (Kuhn 2012)³. There is also statistics showing a significant difference between Central and Eastern European countries and Western Europe regarding attitudes towards same-sex marriage, willingness to accept Muslims into the family, levels of religiosity, belief in fate and importance of ancestry etc⁴. Another striking difference is life expectancy: a 20 year old male from former Eastern Bloc countries was expected to live more than 7 years less in comparison to same age male from other European countries on average in 1990⁵, the life expectancy difference between women of the same age is smaller – 4.5 years less in Eastern Europe. A rapid decrease in adult mortality rates in Europe was observed between years 2000 and 2010. However, the divide between former Eastern Bloc countries and rest of Europe suggests even a gloomier story: male and both sexes' mortality rates in Eastern Europe were more than twice higher in comparison to the West, despite the difference between women being much smaller⁶. A possible explanation for these statistics – different health related habits, which could include preferences for drinking and smoking.

The general analytical approach used in this thesis is quantitative with a comparative perspective. We treat former Eastern Bloc as an area of a natural experiment, thus we compare the outcomes of those individuals we identify as 'affected' by the regime to those that come from other European countries or are recognized as not affected according to our definitions. We identify and treat 'communism' as homogeneous, and even if it is unlikely the regime affected each Eastern Bloc country to the same extent or in the same way, we believe finding an average effect of the regime at this point is sufficient. Further studies

³ Kuhn, A. 2012. Redistributive Preferences, Redistribution, and Inequality: Evidence from a Panel of OECD Countries. *IZA Discussion Paper Series*, [Online] dp6721.

⁴ Saghal, N. et al. 2018. Eastern and Western Europeans Differ on Importance of Religion, Views of Minorities, and Key Social Issues. Pew Research Center, [Online], available at: <http://www.pewforum.org/2018/10/29/eastern-and-western-europeans-differ-on-importance-of-religion-views-of-minorities-and-key-social-issues/>

⁵ Eastern Bloc includes: Czech Republic, Estonia, Hungary, Lithuania, Russia, Slovakia, Slovenia; Western countries: Austria, Spain, Finland, France, Greece, Italy, Luxembourg, Netherlands, Norway, country choice based on availability of data. Human Life Table Database, more information at: <https://www.lifetable.de/cgi-bin/index.php>

⁶ Adult mortality rate – number of deaths between ages 15-60 per 1000 population, data includes 20 former Eastern Bloc countries and 19 Western countries, sample excludes Germany, World Health Organization data.

could explore how much this impact differs depending on other factors like distance from Russia, differences in policies and severity of each separate communist regime, previous history or development level in each country. It was not possible to account for these factors in this research due to the lack of available data but we controlled for country fixed effects with their dummies in order to isolate and measure the average effect of the regime. Country dummies address another concern – possible differences in coefficients of other control variables between Eastern Bloc countries and rest of Europe. However, we did not identify any possible reasons for the control variables related to personal characteristics or their impact to differ significantly between the observations from previous Eastern Bloc countries and the rest of Europe. One possible exception could be income variable, but it is expressed in quintiles relative to native country, which should eliminate the problem. Country selection for each chapter was determined by the dataset and available sample once all control variables were included. In the first chapter, we use European Values Survey, which has data from most European countries. For the next two chapters European Health Interview Survey data was used, which includes a shorter list of European countries, but offers a lot of detailed health-related information, which, to our best knowledge, is not available in a consistent manner from other data sources for European countries.

This thesis is divided into three chapters addressing the questions listed previously. The first chapter is focused on human capital, i.e. the impact of the communist education policy. We look into the effect of the communist regime experience in the early years (0-16 years) on the outcomes of the highest education attainments. In the second chapter, we examine how much the communist regime influenced alcohol consumption frequency and patterns in Eastern Europe comparative to non-Eastern Bloc countries in Europe. We look into the effects of the experience of the communist regime during the formative years (18-25) and the duration of time spent living in the regime respectively on alcohol consumption outcomes: its frequency and incidence of binge drinking. The third chapter is dedicated to the exploration of the differences in health outcomes and changes in social capital, more specifically, we propose a new measure – health perception gap. We look into how the experience of the regime during the formative years is related to the perception of own health and what implications this could have.

I have written this PhD thesis under the supervision of Dr. Alexander Klein, with whom we co-authored the second chapter. His involvement entailed extensive work on co-writing the introduction, conclusions and shaping the article in general. A version of this paper was published in a special issue of the *Journal of Comparative Economics* (2018, 46(3), 821-

837) dedicated to a conference – *People Matter: Quality of Life in Post-Transition Economies*, which was held in Kiev, September 2017, where I presented the article. An earlier version of the first chapter was included in the proceedings of the conference *XXIV Meeting of the Economics of Education Association*, held in Madrid, June 2015 (*Investigaciones de Economía de la Educación*, 2015, vol. 10, chapter 9, p.p. 183-210).

Chapter I

Comparing Eastern and Western Europe: has Communism succeeded in increasing Educational Attainments?

Gintarė Mališauskaitė

University of Kent, Canterbury, UK

Abstract

The paper researches the influence of communist regimes on educational attainments in Eastern Europe in comparison to Western Europe. Education policy in communism influence area had distinctive qualities: it was centrally regulated, free, encouraged to be undertaken by all, supported financially and endorsed equality between genders. The effectiveness of this policy is tested by comparing countries that were subject to regime's influence with those that were not part of it, and searching for observable differences between people who made education choices under the influence of the regime compared to those that were not. This research suggests that communist regime had a significant effect towards educational attainments of people who experienced it in early years. There are observable differences between education completion rates and gender behaviour in the two parts of Europe. Regressions' results support the idea of communism having a positive effect towards primary and especially secondary education completion, but shows the effect towards tertiary education to be insignificant. Data suggests a smaller gender gap in educational attainments in post-communist countries. This would advocate the relative effectiveness of this education policy.

Keywords: educational attainments, communism, Eastern Bloc, Western Europe,

1. Introduction

Educational attainment is an important component in the production function as it represents social capital and skill distribution which in turn are significant determinants of country's labour productivity. That is why increasing educational attainment is an important policy objective in most countries. Throughout the history European countries have undergone different policies and regimes which tackled this issue differently.

It is also agreed that education is an important factor affecting personal income. Raising levels of educational attainment is complex process often limited by the intergenerational transfer of education, occupation mobility or high education costs. Hence it is of an utmost importance to understand the effectiveness of policies attempting to increase educational attainment.

In 19th and 20th centuries most countries in Europe introduced changes in their education systems, aiming at raising educational attainment. The evidence shows that the number of years of compulsory education increased in the first half of the 20th century while the second half witnessed an increase in the secondary and tertiary school participation (Grendler n.d.; Houston 2011). Around 1900 there was regional division by literacy and economic development levels in Western Europe: Protestant north, which was mainly literate and economically developed, and south, which was less literate and underdeveloped (Houston 2011). Eastern Europe was seen to lag behind and was similar to the far south, but progress was still made (Grendler n.d.; Houston 2011). Statistics about education attainment levels and literacy rates especially in Eastern European countries during this period is very scarce. There is some evidence (Foley 2007) that at the time of the Bolshevik Revolution literacy rates in Russia were extremely low: 37.9% for the male population above seven years old and 12.5% for the female population. Unfortunately further comparison of literacy rates in Eastern Europe is not available due to lack and limited availability of resources.

Centrally planned regime governed by communist parties after the Second World War became prevalent in most countries in Eastern Europe for more than half a century. One of the objectives of this regime was an increase in educational attainments' level because of high illiteracy rates, so since the end of 1930 education was compulsory to every adult in USSR (Charque 1932). This regime offered its own original method of achieving this objective: all education was free, encouraged, supported with scholarships to all full-time students, promoted non-discrimination between genders, incentivised education completion by job offers at its end (Charque 1932). Even though social classes were officially abolished

in the communist regime area, children of high ranking officials were more privileged compared to the rest. Regardless of this point, there was still an evident narrowing of the gap in educational attainment between genders and people coming from different social backgrounds (Grendler n.d.). Education in the communist regime area was strictly regulated centrally: religion was taken out of education system, pedagogy was made more uniform between countries, there was not much choice in curriculum of education (Heyneman 1997). Schooling was used to teach certain political and social values, preparing for socialist society; a lot of attention was paid to technical schools to ensure that regime was able to meet the demand for workers needed for plans of heavy industrialisation – education was not liberal and its main goal was to produce socially thinking specialists (Grendler n.d.; Charque 1932).

Research on the effects of the communist regime in Eastern Europe is in short supply. Gerard Roland is one of the main names in this field, he looked into institutional change in post-communist countries (Roland 2002), researched changes in law enforcement in transition economies (Roland 2003), looked into the effect communism had on cultural values and longevity of this change (Roland 2010), etc. Other important studies: Alesina and Fuchs-Schündeln (2007) researched the differences in redistribution preferences between eastern and western Germany; Münich et al. (2005) studies returns to education and wage grit in transition economies in Europe. To our best knowledge, the effect of centralized communist regime on educational attainments in Eastern bloc was not yet researched, even though some studies made some comparison between the educational achievements in Eastern and Western European countries (Ammermüller et al. 2005, Braga et al. 2013). So the aim of this study is to fill this gap in knowledge and to see what the effect of the communist regime's education policy on educational attainments was. Descriptive statistics, data and some literature suggest such reforms could have had a positive effect on average education attainment level in population. This is being tested by comparing countries that have undergone the regime's influence with those that were not part of it and searching for observable differences between people who made education choices under the influence of the regime compared to those that were not. This question should be intriguing because a large number of countries in Europe were subject to the education policy of the communist regime. Results of this research will be useful in gaining insight for education policy decisions.

The data used for this research was taken from European Values Survey (EVS) database providing individual level data. It is available for 48 European and post-soviet countries and

we divided them into two groups: Western and Eastern European countries respectively, where Eastern European countries were subjected to the communist regime in during the second half of 20th century⁷. The effect of the regime's education policy on educational attainments was assessed with a categorical variable 'communism', which is equal to 1 for those respondents who made education decisions subjected to the communist regime. Regressions for secondary, primary and tertiary educations were run separately with three different models: linear probability model, probit, and SNP, following the same specification; bivariate probit and SNP2 models were used for simultaneous estimation of vocational and university preparation secondary educations' regressions.

This paper finds that communist regime had a statistically significant and positive effect on secondary and primary educations attainments, while the effect on tertiary education was found insignificant by regression analysis. The descriptive statistics suggests higher overall education completion rates and a smaller gender gap in educational attainments in post-communist countries. This would advocate the effectiveness of Soviet Union's education policy.

The paper proceeds as follows. The next section presents relevant literature review; section 3 explains methodology used in this study; section 4 examines the data and descriptive statistics; section 5 presents the results and their analysis; section 6 concludes and offers discussion.

2. Previous literature

A number of studies looked into the determinants of educational achievements. The level of educational attainments was found to depend on a number of factors: parental education, parental background, financial constraints, labour market conditions, personal characteristics. The most relevant research for this study will be reviewed in this section.

One of the most important determinants of educational attainments level is found to be parents' background such as education level, financial status, and occupation. Most studies find very low intergenerational mobility between parents and their children in terms of education and occupation. Aina (2013), who studied tertiary education drop-out rates in

⁷ Amount of time spent under the communist regime influence varies: Azerbaijan, Armenia, Belarus, Georgia, Russian Federation 1922-1991, Ukraine 1922-1990, Estonia, Moldova 1940-1991, Latvia 1941-1991, Lithuania, Poland 1944-1990, Albania, former Yugoslavia (Bosnia and Herzegovina, Croatia, Montenegro, Serbia, Slovenia, Macedonia, Kosovo) 1945-1991, Bulgaria, Romania 1946-1990, Hungary 1947-1990, former Czechoslovakia (Czech Republic, Slovak Republic) 1948-1990, Eastern Germany 1949-1990.

Italy, finds strong persistence of intergenerational correlations in education. Her research suggests that children with parents with higher than compulsory education have a higher chance of their child continuing the studies. Dropout rates tend to be higher for males and for children with parents with lower education; this result is not affected by parents' income. Children from better cultured background tend to choose the type of education that is leading towards university, while children from less educated backgrounds tend to choose more labour orientated education. Ben-Halima et al. (2014) found a significant rise of importance of parents' backgrounds on educational attainments, especially through family income, from 1993 to 2003 in France. They claim that intergenerational persistence is higher for males. Similarly, Schütz et al. (2008) made a comparison between 54 countries of how strong is family background influence on children's educational performance. Main indicator for family background was chosen to be the number of books at home. Their results advice that variation between countries is due to systematic failures of country's education system, that the longer pre-primary education the smaller are the effects of parental background. Checchi (1997) discovered that almost half of observed immobility in occupations in Italy, Germany and United States is accountable to educational attainment level. Galindo-Rueda and Vignoles (2005) researched the importance of cognitive ability in comparison to parental background in Britain. They find importance of the former declining while the latter remaining very important – they observed a large increase in educational attainments by children with low ability but good parental background. Triventi and Trivellato (2009) also found family background being very important for university participation rates in Italy. Even though, Checchi et al. (2013) observe an improvement in educational mobility in Italy, children with parents from lower cultural backgrounds remain at disadvantage. Bratti et al. (2008) researched the impact of expansion of higher education in Italy, and found that reduction of inequality of university education access improved enrolment rates but not completion rates. Braga et al. (2013) researched the relationship between education reforms and attainments in 24 European countries and also confirm strong correlation between parents and children's educational attainments. All these studies show how important parental background is and, even though there is small variation, that intergenerational immobility is very persistent.

Intergenerational mobility is directly related to the issue of financial inequality. Gruber and Kosack (2014) find that higher primary enrolment rates are related to slightly higher inequality in the future. It is explained by higher primary education rates reducing wage premia previously enjoyed by a smaller number of people. Tertiary tilt, country focussing its

finances towards tertiary education, is common in a lot of developed countries, while developing countries with high primary education demands tend to focus their finances on primary education due to lack of educated workers. They find that only Eastern Europe and Central Asia did not exhibit clear tertiary bias towards benefiting mainly wealthier citizens which could be due to political history and communism regime effects. Braga et al. (2013) find a positive relationship between education reforms and reduction of economic inequality. Newell and Reilly (1999) ran a cross-country comparison study of rates of return to education in transitional economies ranging through central and Eastern Europe, Russia and former Soviet Union countries in Asia. They find the rise in rates of return during the transition period which comes with an increase in inequality. Triventi and Trivellato (2009) suggest that financial inequalities tend to persist over time with only a slight reduction in Italy. Similarly, Ben-Halima et al. (2014) find a decrease in inequality in general in France since 1970s but higher inter-generational inequality persistence and lower social mobility. There seems to be a general consensus that inequality is persistent, even though it slightly decreased in Western European countries, Eastern European countries experienced the opposite due to transition and raising returns to education.

Another important aspect to consider while comparing education levels between countries – the quality of education. Hanushek (2013) studies differences between educational attainments in developed and developing countries and finds skill deficits in developing countries. Ammermuler (2005) studied schooling quality in seven Eastern European countries using Third International Mathematics and Science Study data. After the collapse of the communist regime, education experienced quick decentralisation and other reforms in education systems and institutions were undertaken in transition economies. Central European countries, that made reforms earlier compared to Eastern European ones, were found to catch up and surpass the Western European countries in terms of test scores, while Baltic states were found lagging behind, with a distribution closer to other post-communist countries. Another factor to consider in quality of education is class size. Lazear (2001) found that larger classes provide better outcomes for better students but discipline plays a more important role than class size. Grouping students by ability was found optimal and naturally happening through self-selection into private schools by more attentive students.

The effect of education policies, reforms and regional effects on educational attainments should also be taken into account. Schooling reforms were found to affect the distribution of educational attainments and to have a positive effect on average years of

education in the population (Braga et al. 2013). The relationship was found between the level of autonomy in a school and the performance in them (Hanushek et al. 2013). This effect was found to be positive in developed countries and negative in developing ones. It is argued that standardisation could be important for decisions related to academic content but less important for process operations and management of employees. Autonomy tends to provide better results where there is external accountability and where opportunistic behaviour is limited. Aina (2013) suggests that geographical area of residence is also important, and finds correlation with educational attainments. Cappellari and Lucifora (2009), evaluating the tertiary education reform in Italy, report that areas with lower unemployment rate and better possibilities tend to discourage university participation. This information suggests that education reforms are likely to have an effect on educational attainments; that there could be a difference between developing and developed countries, also influencing the effectiveness of policies in them; and that smaller regional effects also need to be taken into account.

Referring to the findings in this literature, this study looks into the highest education attainment levels among the individuals from former Eastern Bloc and Western European countries to assess if the education reforms undertaken by the communist regime were successful. Literature hints that Eastern Europe could be subject to a lack of tertiary education tilting, skill deficits and large variance of children’s test scores, increases in rates of return on education and in economic inequality. However, this research only looks into the outcomes recorded after the collapse of the communist regime. The main reason for this was a lack of available and also reliable data from Eastern Europe before the collapse. Which is why to our best knowledge this study is filling a gap in literature by looking into the effect of the communist regime on educational attainment outcomes.

3. Methodology

To analyse the effect of Communist regime on educational attainments in Eastern Europe, we estimate the following regression equation:

$$Educvar_i = \alpha_0 + \alpha_1 CR_i + \sum_{j=1}^n \alpha_j Controls_i^j + \varepsilon_i \quad (1)$$

A few different versions of the specification (1) were run. The dependent variables were created from an ordinal 8 choice variable for ‘*The highest educational level attained by*

*respondent*⁸. Preliminary regressions took into account the ordinal nature of the variable but since it was not possible to claim that the differences between categories were equal it made it difficult to interpret results in an intelligible manner, thus they were not reported here. Instead binary education level indicators were created using this ordinal 8 choice education variable, which *Educvar* stands for: tertiary, secondary and primary. These variables are equal to 1 if individual has attained this level of education, and equal to 0 if not. The implication is made that if respondent indicated achievement of some higher level of education then lower levels of education were attained. Tertiary education takes value of 1 if tertiary (university) education with a degree is completed, and 0 if it is unfinished or no tertiary education. For secondary education 1 means vocational or university preparation education was completed and if tertiary education is either incomplete or finished, 0 – incomplete or no secondary education. Primary education is a binary variable that takes a value of 1 if it is completed and if respondent had any secondary or tertiary education and 0 if unfinished or no primary education. In addition, the choice between vocational (1 if completed and 0 – if incomplete or no secondary education) vs. university preparation secondary education was examined. University preparation is classified as educational tracks that are part of academically oriented general education, completion of which is marked by a satisfactory passing of exams, e.g. Abitur, and provides access to tertiary education; it also includes post-secondary non-tertiary education. Having an incomplete or complete university degree was regarded as having completed preparation for university secondary education. This is a logical assumption but it does not rule out the possibility of a person having completed vocational secondary education and then deciding to go to university, however it is likely there would not be many such cases in the sample. Summary of the variables used in the regression analysis is reported in Table 2 and a table showing how binary dependent variables were constructed is detailed in Table 3.

The main explanatory variable of interest in this study is the experience of communism regime (CR) and its impact was assessed on all three education levels. It is a categorical variable equal to 1 if a country was under the influence of the centralized communist regime prevalent in Eastern Bloc countries and respondent was born between 1940 and 1975, and 0 otherwise. There are two reasons for the choice of the dates. Firstly, by 1945 there was a

⁸ Variable constructed in style of CASMIN (Comparative Analysis of Social Mobility in Industrial Nations) but not in exact accordance since it is also based on educational level in country based on ISCED (International Standard Classification of Education). Variable choices explained in more detail in the Appendix.

clear division between which countries belonged to either the Eastern Bloc or the Western Europe and this remained unchanged until the collapse of the communist regime in 1990-1991. So a person born in 1940 would be of the age of 5 when this division happened and their education decisions and choices would have been made under the influence of the communist education policy. Secondly, according to Giuliano and Spilimbergo (2014), macroeconomic environment experienced when young can affect individual's preferences and beliefs in later years. They found that individuals who experienced a recession when they were young, support more government redistribution and believe success in life is determined more by luck than effort. So a person born in 1940 would have spent at least 11 years influenced by the regime before making education decisions at the age of 16. While a person born in 1975 would have been of the age of 16 in 1991, when the regime fully disintegrated, so would make education related decisions before education reforms by independent new governments taking effect and also being influenced by 16 years spend in the communist regime.

The regression includes different control variables indicated by the vector *Controls*: information about parents⁹, personal characteristics¹⁰, a time trend¹¹ and country dummies. Since data is at individual level, i denotes an individual, and ε indicates an error term.

Several different estimators were used. We start with Linear Probability Model (henceforth LPM), then we relax the assumption of linearity and use probit. Unlike in LPM, marginal effects of the model had to be calculated separately. Average marginal effects were calculated to assess the average impact effect over all the individuals in the sample, which seemed to provide more interpretable information than marginal effects at the mean, which would give the information about the mean observation. This set of regressions was followed by semi-nonparametric (henceforth SNP) equations. SNP estimator was used for two reasons: a) it does not assume error terms to be Gaussian and unknown densities of error

⁹ This includes binary variables for parents' education: tertiary (1 if parent had any tertiary education, 0 otherwise), secondary (1 if either one of secondary education types was completed or tertiary education was started or finished, 0 otherwise) and primary (1 if primary school was completed or higher levels of education started or obtained, 0 if incomplete), and their occupation, an ordered categorical variable. The binary parents' education variables are based on a similar multiple-choice variable indicating the highest education level attained with the same choice categories as respondent's. Parents occupation choices range between 1 and 11, detailed in the Appendix. It is used as a proxy for parental income since it is not available on EVS. The two are expected to be correlated and allow expressing parents' need for education, and its perceived value; this assumes lesser need for higher levels of education the higher the value of parental occupation.

¹⁰ Including: gender (1 if male, 0 otherwise), dummies for age cohorts (15-24, 25-34, 35-44, 45-54, 55-64, 65+) and a dummy for living in a city (50,000 or more inhabitants).

¹¹ Calculated by deducting 1881 (a year of birth of the oldest respondent within the sample) from a year of birth of the respondent.

terms are approximated to derive pseudo-Maximum Likelihood estimator, ensuring a higher accuracy of resulting estimates; and b) due to a high number of individuals having completed primary education, almost 25% of sample observations are excluded in probit estimations of primary education due to perfect predictability, but this is not an issue with SNP. Two regression specifications were reported from each model: one with and one without a dummy indicating communist regime experience. Full available sample was used for each dependent variable estimations.

This analysis was followed by bivariate regressions for choosing between vocational or university preparation secondary education. We investigate the differences between the individuals who made a choice between the two options and look into the effect the communist regime experience in the early years had on this choice. This method allows simultaneous estimation of both dependent variables that are potentially related using the same set of explanatory variables as long as the correlation of error terms of equations is high and significant. Within our sample there are 3 possible outcomes: vocational secondary completed ($y_1=1$) but not university preparation secondary education ($y_2=0$), university preparation secondary education completed, but not vocational ($y_1=0$, $y_2=1$), both types of secondary education are not completed ($y_1=0$, $y_2=0$). Regressions were run using two estimators: bivariate probit and bivariate semi-nonparametric (henceforth SNP2). SNP2 was used so that error terms' densities were estimated instead of just being assumed to be Gaussian (as in bivariate probit). The specification of regression equation is the same as above. Clustered standard errors were used in all of the models mentioned apart from SNP and SNP2 in order to account for possible correlation of the error terms within country level clusters. Estimations were run with full available sample.

The discussed specification presents some challenges. Firstly, the exogeneity of the variable *city* is questionable since it is difficult to determine if living in a city influences educational attainments or educational attainments influence the choice between living in a city or more rural area, both of these effects could be happening at the same time. It is nonetheless included as a control due to the indication in the literature that geography tends to be an important determinant for the educational attainments. Secondly, if the communist regime had an effect on educational attainments of people who lived in it, then the variables for parents' education are likely to be influenced by it since in some regions the regime lasted for 40-60 years. This is not considered a major issue as it would only make the 'communism' effect underestimated and part of it would be captured through parents' education dummies. Lastly, the EVS data we used was collected almost 20 years after the collapse of the regime

and we do not have the information to be able to tell with certainty that the respondents that are identified as ‘affected by the communism regime’ actually lived in the regime during their early years. However, we do not expect migrants to be a large proportion of our sample, and, since migration westwards was more prevalent than the opposite direction, this would signal that any effect of the communist regime found would be likely underestimated too.

4. Data and descriptive statistics

We used European Values Study (henceforth EVS) data. It provides individual level data about opinions and values regarding life, family, work, religion, politics and society in European countries. There are four waves of data available so far: 1981, 1990, 1999, and 2008. The fieldwork for the 5th wave started at the end of 2017, but the data is not released yet. We use only the 4th wave data since it is the only one with providing information about parents’ educational attainments and occupation.

Table 1¹² provides the list of countries and a number of observations in the sample available. The sample of 48 countries consists of two parts: Western European countries (23) and Eastern European countries (25). A country is classified as Eastern European if it was part of the Eastern Bloc. Total number of observations in the sample is 50,865, 23,591 of which are from the Western Europe and 27,274 from the Eastern Europe.

Within the sample used 97.6% of the population have completed primary education; 61.7% have finished one of the two secondary educations; 10.7% have a degree. Vocational secondary education is more than five times less popular compared to university preparation secondary education with 9.3% and 52.4% completion rates respectively. People who are identified as having made their education choices under the influences of the communist regime amount to 30.84% of the sample. Regarding parental educational attainments, 14.3% of them have had some tertiary education or have completed it, 38.3% finished secondary education, and 90.7% are with primary education. These rates are lower compared to respondents’ average achievements, which is expected. There is a higher proportion of parents with tertiary education than their children due to including incomplete university education for parents but not children. This is based on an assumption that most parents would encourage their children to achieve at least the same level of education as they themselves have. Men constitute 44.4% of the sample, 25.6% of respondents live in places

¹² All the tables are reported in Tables at the end of the document.

with at least 50,000 people. The sample is age balanced as represented by the age cohorts (15-24, 25-34, 35-44, 45-54, 55-64, 65+).

Tables 4 and 5 summarize respondents' and their parents' educational attainments by the part of Europe. In both parts of Europe the largest proportion of respondents indicate completing university preparation secondary education as their highest educational attainment (19.65% – West, 31.91% – East). A very small proportion of individuals from Eastern Europe point out primary education as their highest educational attainment (3.6% in comparison to the West – 14.89%). Significantly larger proportions of individuals from Eastern Europe claim completing one of the two types of secondary education or university in comparison to the Western Europe. Similar trends are present in parents' educational attainments. In the Western Europe, the largest proportion of parents (31.09%) list primary education as their highest educational attainment while in the Eastern Europe it is incomplete vocational secondary education with 22.74% of population closely followed by university preparation secondary education (19.74%). Apart from the primary education, the rest of the education categories (vocational and university preparation secondary educations and university education) have almost twice higher completion rates in the Eastern Bloc in comparison to the Western part. This statistics supports the claim that the communist regime could have affected respondents' parents' educational attainments too. It could be argued that higher parental educational attainments propagate higher educational attainments in their children, but this could not explain how such a jump in attainments from the low literacy rates recorded by the literature at the beginning of the 19th century is possible.

Tables 6 and 7 summarize education differences between Western and Eastern Europe by gender and age group. The unconditional means are higher in Eastern Bloc in all age groups and for both sexes. Differences tend to be smallest between primary education rates and grow larger for secondary and tertiary education. Wilcoxon Rank-sum tests were performed to check the statistical significance of the above reported differences¹³. The results show that only in the case of primary education attainments in men's first two age cohorts and men's tertiary education cohorts: 25-34, 35-44, 45-54 the differences are not statistically significant.

These tables suggest some differences by gender and two parts of Europe. For the primary education differences between men and women in the same age group are

¹³ The test is also known as the Mann-Whitney U test, with a null hypothesis of two populations being the same; the alternative is populations being different. It is more efficient than the t-test with populations that have a non-normal distribution and almost as efficient with the ones with a normal distribution.

negligible. The only age group that has more noticeable differences between sexes and geographical locations is 65+ and could be associated with the times new education policies were started in Eastern Bloc. The differences for secondary education attainments are larger. In Eastern Europe women have higher or similar rates compared to men in the first five age groups and the opposite is observed for the last age cohort. In Western Europe women have higher completion rates in the first and third age groups, in the rest of the cohorts the opposite is true and the difference is larger. Rates of completed tertiary education follow a similar pattern: they are higher for women in comparison to men in the middle four age groups (between ages 25-64), and the rate is lower in the last age group. In Western Europe, the rates for women are marginally higher than men's only in the first age group; for the rest of age groups men have higher rates and the difference is more sizeable.

According to these tables, there is a tendency for women in Eastern Europe to be on average better educated than men, unless they belong to the age group of 65+. Men, on the other hand, seem to have better educational attainments in Western Europe in comparison to women in most age groups apart from the 15-24. These numbers support the famous communist regime motto that access to education should be equal between sexes, which could have resulted in higher educational attainments among women. This also suggests the gender gap in education could have started closing in Eastern Europe before Western Europe.

Tables 8 and 9 summarize the two types of secondary education attainments by age, gender and part of Europe. We see that university preparation was chosen more frequently in comparison to vocational training by both genders and among all age groups. Men tend to have higher completion rates for vocational/technical secondary education compared to women in both: Eastern and Western Europe. This difference tends to be larger in Eastern Europe. Completion rates tend to be higher for both genders in Eastern Europe, this difference is only smaller for women in the first two age groups in vocational training. The differences are not significantly different between East and West Europe only for women's 1st and 2nd age cohorts. This points out larger gender differences in vocational secondary education in Eastern Europe, which could be related to previous heavy industrialisation in Eastern Bloc.

The difference between the rates of university preparation tends to be larger between two parts of Europe in comparison to vocational education. The difference by age group is more distinct among women than men. In first five age cohorts more women from Eastern Europe have completed university preparation compared to men; the Western women's ratio

is higher only in the first and third age cohorts. Much larger differences in vocational education between the East and the West could be explained by the communist regime paying a lot of attention to technical professions and jobs.

5. Results

5.1 Secondary education

The results of estimating equation 1 with secondary education as dependent variable are presented in Table 10.

In LPM equations R^2 measures indicate that specification without the communism dummy explains 27.73% of variation in dependent variable and 27.82% with. In probit results inclusion of the communism dummy also improves Pseudo R^2 measure somewhat. Communism dummy has a positive effect towards secondary education attainments and is statistically significant at 1% significance level in all three specifications. LPM estimates show that, *ceteris paribus*, being from the Eastern Europe increase the possibility of 'success' (i.e. completing secondary education) by 0.053. Estimates using probit and SNP show similar results (0.048 and 0.050 respectively). Consistent results across different estimators strongly suggest that communism was a significant factor in increasing secondary education attainments in Eastern Europe.

Other statistically significant explanatory variables in all specifications are parents' secondary and primary education dummies, parents' occupation, age cohorts, city and time trend. Parents' educational attainments have some of the largest effects on secondary education achievements. Their estimated impact size does not change much across models. Parents' tertiary education is not significant in LPM but is highly significant in probit and SNP models, which could be due to LPM failing to capture the variation in the data. It has the smallest marginal effect out of all parents' education dummies, it varies between: 0.04 – 0.048. Their secondary and primary education dummies are significant at 1% level and have quite similar positive coefficients in all specifications, with parents' primary education having a somewhat larger effect. These results show high importance of parents' educational attainments on their children's accomplishments which is in line with our expectations and the literature.

Parents' occupation is statistically significant at 1% significance level in all specifications and has a negative sign of a coefficient or marginal effect. This follows our

expectations, since lower values of this variable represent higher-ranking positions, which should imply on average higher income, and higher values represent more manual labour and lower social status (see the Appendix for the details). *Ceteris paribus* an increase by 1 unit in this variable (marginally lower social status level) decreases the probability of respondent finishing secondary education by 0.017 in each specification in table 10.

Dummy for living in a city is significant at 1% level and has a positive effect towards secondary education attainments in all given estimations. The effect ranges between 0.056-0.064. It is expected for this effect to be positive due to rural-urban educational inequality, which is found to be smaller in countries with greater resources and level of development (Ulubasoglu and Cardak 2007). The magnitude of this effect is close to that of communism.

All age cohorts are statistically significant in provided specifications. First age cohort (15-24) is excluded to avoid dummy variable trap, so the marginal effects of other cohorts show how they compare to the first one. It is noteworthy that estimated effects of age cohorts fall with inclusion of the communism dummy. This would suggest that age dummies capture the communism effect partially, and, since changes in coefficients are different, it hints at communism affecting certain age groups more, the change is most noticeable for 3rd, 4th and 5th age cohorts. All age cohorts coefficients and marginal effects are positive, this is probably related to the fact that some people within the first cohort (comparison group) would be too young to have finished secondary education. The estimated marginal effect for the 4th cohort is largest and only somewhat smaller for the 5th cohort. The age effects are sizeable in comparison to most other explanatory variables.

Gender is only significant in SNP regressions and suggests that being male increases the probability of having secondary education by 0.016. This is in line with the descriptive statistics (tables 6 & 7) for Western Europe, where more men in most age cohorts have completed secondary education compared to women, but do not represent the unconditional averages for the Eastern Bloc, where a higher proportion of women than men tend to have secondary education. This result could suggest that there are marginally more men than women with secondary education in Europe.

Overall, the results from all three models indicate that the communist regime had a statistically significant positive effect on secondary education attainments, which would indicate the effectiveness of the communist education policy¹⁴.

¹⁴ Due to similarity of the results from all three models inefficient SNP estimations might seem unnecessary, but LR test rejects Gaussianity assumption at 1% significance level when comparing SNP model to probit, which justifies the inclusion of these results.

5.2 Choice between vocational and university preparation secondary education

In this section we examine more closely the choice between vocational and university preparation secondary education. The regression specification was used as for other types of education to estimate bivariate probit and SNP2 models. Bivariate probit model assumes that its errors are independent, identically distributed as in probit model. A rho test in bivariate probit reveals that correlation between two regressions residuals is significant at 1% level (Cameron and Trivedi 2010), thus this running bivariate regressions is justified. SNP2 regressions, just as SNP regressions, relax the assumption of gaussianity of residuals. The estimated residuals' correlation coefficient is high and negative at -0.76 in both estimations. Wald Chi2 test shows that all regressors are jointly statistically significant in all sets of equations at 1% significance level. Results are provided in table 11.

The marginal effects from both: bivariate probit and SNP2 models are quite similar for the main variable of interest, communism, but the significance levels are different. This indicator is significant at 5% level for vocational training but not for university preparation in bivariate probit estimation, but it is highly significant at 1% level for both types of education in SNP2. Communism is estimated to increase the possibility of completing vocational school by 1.4-1.6%, *ceteris paribus*, by both models. This is in line with descriptive statistics and the literature, which claims that one of the communist regime goals was to create skilled labour force for factories, manufactories etc. (Charque 1932). SNP2 results show that, *ceteris paribus*, the communist regime increased the probability of completing university preparation secondary education by 2.6%, which is higher than the effect on vocational education. This is consistent with the descriptive statistics as it shows Eastern Europe consistently having higher unconditional averages of university preparation secondary education than Western Europe, while the averages are more similar for vocational education. These results suggest that the communist regime had a positive effect on the completion rates of both types of secondary education, which shows the same trend as with previously discussed levels of education.

Parents' educational attainments are mostly statistically significant predictors. In all specifications, parents' education dummies positively affect choosing university preparation over vocational school. While the marginal effects do not change much in bivariate probit estimations with or without the communism dummy, but the results are different in SNP2 estimations when the communism dummy is included. Since marginal effects of parents' education dummies decrease in the specification with the communism indicator, we might

hypothesize that the estimates of parents' education partially capture the effect of communism in the specifications without this dummy. The magnitude of marginal effects for parents' education suggest that educated parents would very likely encourage their children to attend university preparation secondary school, which could increase their chances of successfully acquiring university degree later in the future. In the case of vocational education only parents' primary education dummy has a positive effect, parents' secondary education is less significant, and their tertiary education has a significant negative effect. These results support general intuition of more educated parents trying to influence their children's choices towards the possibility of higher levels of education.

The effect of parents' occupation is also opposite in the choice between vocational or university preparation in both models. Parents' occupation positions that are higher ranked socially increase the probability of their children choosing university preparation and decrease the chance of vocational studies choice. The marginal effects are very similar in both specifications of bivariate probit equations. The results are different for SNP2 regressions, but after the inclusion of the communism indicator the marginal effects are close to bivariate probit values.

Gender is significant in all bivariate probit equations, but not significant for university preparation in both specifications of SNP2. Being male is more significant and increases the probability of completing vocational studies, and is less significant with a negative effect towards university preparation studies. This is consistent with descriptive statistics in Tables 8 & 9.

Age marginal effects are mostly significant and mostly positive, probably indicating that in comparison to 15-24 aged people other age groups are more likely to hold one of the two types of qualifications. In SNP2 model the last two age cohorts have a negative effect towards university preparation choice, this could suggest an increase of popularity of choosing university preparation secondary education in younger age groups.

Living in a city is statistically significant in all regression specifications. It increases the probability of choosing university preparation but reduces the chance of undertaking vocational studies.

5.3 Tertiary education

The results for tertiary education regressions are presented in Table 11 and include estimations with – LPM, probit and SNP estimators, with and without the communism dummy respectively.

For this set of results R^2 and Pseudo R^2 measure are lower than for the secondary education, which shows a worse fit. The communism dummy does not add much explanation since these measures do not increase with the inclusion of the dummy, even if regressions Wald Chi2 test shows that all of the regressors are jointly significant in SNP estimations. Furthermore, the communism dummy is not significant at any of the estimations. Remembering the descriptive statistics in Tables 6 and 7 which show the differences in tertiary education attainment are not significantly different for 3 out of 6 men's age groups¹⁵ between Eastern and Western Europe could help explain this finding. Nonetheless, in all three levels of education, university completion rates are least different between two parts of Europe for men. It could be speculated this was due to the incentives to undertake university education being quite low then, since the returns to education in communist regime countries were low and higher earnings were foregone during additional years of education (Münich et al. 2005). But since education was promoted equally between sexes in Eastern Europe, this could have encouraged women to enter university, while men were more likely to choose higher earnings.

Except for parents' tertiary and primary education, the marginal effects of the remaining explanatory variables are very similar in all of the specifications. Parents' education is has a positive impact for university completion. The positive effect is largest for their tertiary education in LPM and probit models, while rather similar to parents' primary education in SNP. It is reasonable to expect parents' tertiary education to have a large impact to their children's university completion due to low education intergenerational mobility – there is a high correlation in education attainment between parents and their children, the literature suggests (Aina 2013, Ben-Halima et al. 2014, Checchi 1997, Triventi and Trivellato 2009, Braga et al. 2013).

Parents' occupation is highly statistically significant and higher social standing (lower categorical value of parental occupation variable) is related to higher probability of completing tertiary education. Gender is significant at 1% level only in SNP regressions and has a small positive marginal effect towards tertiary education attainments. This information is probably explained by higher proportion of men with tertiary education in most age cohorts (25+) in Western Europe but less so for Eastern Europe, since tables 6 & 7 show more women completing university than men in our sample. Marginal effects for age cohorts

¹⁵ The differences are statistically significant in all women's age groups for Eastern Europe though, ref. Tables 5.

are highly statistically significant and positive in all of the specifications. This reasonably suggests that older respondents have higher attainments than those in first cohort, since some younger respondents would not have had enough time to finish university¹⁶. Dummy for living in a city is also significant at 1% level and has a positive relationship with tertiary education attainments, which is in line with the observed inequality between rural and urban educational attainments (Ulubasoglu and Cardak 2007).

Overall, the results do not show the communist regime having a significant effect towards tertiary education attainments. Most of the effect is attributed to age cohorts and we can also observe the importance of parents' situation.

5.4. Primary education

The results for primary education are presented in table 13.

In both LPM and probit equations inclusion of the communism dummy marginally improves the goodness of fit measures. The communism dummy is significant and positive in probit and SNP specifications, with the former estimating the increase in the probability of completing primary education by 0.9%, and the latter by 0.6%. This is a much smaller effect in comparison to secondary education, but it is consistent with the descriptive statistics – the difference between two parts of Europe seems much smaller in primary education, as shown in tables 6 & 7.

Parents' education dummies are significant at least at 10% level in LPM and SNP regressions, but their tertiary and secondary education attainments are not significant in probit model. The marginal effect of parents' tertiary education is small and negative, small positive of their secondary education, and larger positive of their primary education. This suggests that parents' tertiary and secondary attainments are not very predictive of their child's primary education completion, probably partially due to little variability in primary education attainment, which could be explained by it being compulsory.

Parents' occupation status is highly statistically significant in all of the specifications and follows a similar trend as in previous estimations with other levels of education as

¹⁶ The same sample was used for all dependent variable estimations (primary, secondary, tertiary and two types of secondary education), which allows comparability between results from each estimation. Even though part of the first cohort (aged 15-24) would be too young for tertiary education completion, but it is only a small part of the sample, and the cohort was not removed from the estimation sample since most of the first cohort are old enough to have completed secondary education, and all are old enough for primary education completion. This choice also allows keeping more individuals from former Eastern Bloc that have not experienced the communist regime in their early years in the sample, which helps with the identification of the 'affected' by communism group. Any resulting issue is also partially addressed with controlling for cohort dummies.

dependent variables. Being male increases the probability of having primary education by 0.011 – 0.013 in all specifications. This is probably explained by higher educational attainment in men than women in older age cohorts, while this difference in younger age cohorts is miniscule. Living in a city is highly significant but the positive effect is much smaller (varies between 0.006 – 0.008) in comparison to secondary (0.056 – 0.064) and tertiary (0.048 – 0.059) education. Intuitively this would suggest that secondary and tertiary education are more important for the life in a city, and that primary education was most likely compulsory everywhere.

Overall, the results indicate that the communist regime had a statistically significant positive effect on primary education attainments in Eastern Europe, even though this effect is small.

6. Discussion and conclusion

This research suggests that the communist regime had a significant effect on the educational attainments of people who experienced it. There are observable differences in education completion rates by gender and part of Europe. The results support the idea of communism having a positive effect towards primary and especially secondary education completion. There was no statistically significant effect found towards tertiary education, even if descriptive statistics suggests higher proportion of people with completed tertiary education in the Eastern Europe.

Parents' education has a significant effect on the respondent's educational attainment – the higher parents' educational attainment, the higher that of the respondent, which is consistent with the findings in the literature of low intergenerational mobility in education (Aina 2013, Ben-Halima et al. 2014, Checchi 1997, Triventi and Trivellato 2009, Braga et al. 2013). However, educational attainment of some parents could have been influenced by the communist regime as well. This would imply an underestimation of the effect of the communist regime on educational attainment since part of the effect would be indirectly captured through parents' education. Distinguishing between these two effects was not explored in this study.

Even if the results of this study show that the encouragement to pursue education expressed during the communist rule in Eastern Europe was relatively successful it is difficult to come up with policy implications. Partially this is due to the methods that were used in the communist regime for encouragement now being no longer considered as ethical:

home checks and penalties for not complying. The difference of circumstances should also be taken into account – at the beginning of 20th century literacy rates and education levels in Eastern Europe were low and education was not considered a priority, there was much greater reliance on agriculture as means for survival, which is not the case anymore. Now secondary education is compulsory in modern countries and we can only think of increasing tertiary education attainment levels. Another reason is financial – it might not be the most cost-efficient approach to force everyone into further levels of state financed education (as in the communist model) if they are not necessary for the individual. Thus in countries with low educational attainment levels, financing and enforcing primary and secondary education enrolment and attainment is probably a reasonable approach; but trying to achieve higher tertiary education completion rates in Western countries might require more innovative solutions, especially keeping in mind that estimations show soviets not succeeding in overtaking the West in the tertiary education race.

The future research could do more in this area. Topics could include: an analysis of the cost efficiency of the education policy, the exploration of the possible ways to assess the quality of education offered back then, etc. This particular study could be expanded through the addition of the indirect effects of the communist regime captured in parents' education through interaction terms or instrumental variables. It could also be interesting to research if there are any other long-lasting effects of the communist centrally governed political regime in previous Eastern Bloc, and especially how it might affect the mentality, behaviour and other preferences of the people affected.

References:

- Aina, C. (2013). Parental background and university dropout in Italy. *Higher Education*, 65(4), 437-456.
- Alesina, A. and Fuchs-Schündeln, N. (2007). Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences. *American Economic Review*, 97(4), 1507-1528.
- Ammermüller, A., Heijke, H. and Wößmann, L. (2005). Schooling quality in Eastern Europe: Educational production during transition. *Economics of Education Review*, 24(5), 579-599.
- Ben-Halima, B., Chusseau, N. and Hellier, J. (2014). Skill premia and intergenerational education mobility: The French case. *Economics of Education Review*, 39, 50-64.
- Braga, M., Checchi, D. and Meschi, E. (2013). Educational policies in a long- run perspective. *Economic Policy*, 28(73), 45-100.
- Bratti, M., Checchi, D. and De Blasio, G. (2008). Does the Expansion of Higher Education Increase the Equality of Educational Opportunities? Evidence from Italy. *Labour*, 22, 53-88.
- Cameron, A. C. and Trivedi, P. K. (2010). Chapter 12 Testing methods; Chapter 14 Binary outcome models; Chapter 15 Multinomial models. In: Trivedi, P. K. ed. *Microeconometrics using Stata*. Revised edition. College Station, Tex.: College Station, Tex.: Stata Press.
- Cappellari, L. and Lucifora, C. (2009). The “Bologna Process” and college enrollment decisions. *Labour Economics*, 16(6), 638-647.
- Checchi, D. (1997). Education and Intergenerational Mobility in Occupations. *American Journal of Economics and Sociology*, 56(3), 331-351.
- Checchi, D., Fiorio, C. V. and Leonardi, M. (2013). Intergenerational persistence of educational attainment in Italy. *Economics Letters*, 118(1), 229-232.
- De Luca, G. (2008). SNP and SML estimation of univariate and bivariate binary-choice models. *The Stata Journal*, 8(2), 190-220.
- Foley, K. (2007). *Literacy and Education in the Early Soviet Union* [Online]. Available from: http://www.portalus.ru/modules/english_russia/print.php?subaction=showfull&id=1190296667&archive=&start_from=&ucat=22& [Accessed 3/24/2015].
- Galindo-Rueda, F. and Vignoles, A. (2005). The Declining Relative Importance of Ability in Predicting Educational Attainment. *Journal of Human Resources*, 40(2), 335-353.
- Giuliano, P. and Spilimbergo A. (2014). Growing up in a Recession. *Review of Economic Studies*, 81(2), 878-817.
- Grendler, P. F. *Education in Europe - Nineteenth-and Twentieth-Century Education - Schools, University, Secondary, and Press - JRank Articles* [Online]. Available from: <http://science.jrank.org/pages/9080/Education-in-Europe-Nineteenth-Twentieth-Century-Education.html> [Accessed 3/18/2015].

- Gruber, L. and Kosack, S. (2014). The Tertiary Tilt: Education and Inequality in the Developing World. *World Development*, 54, 253-272.
- Hanushek, E. A. (2013). Economic growth in developing countries: The role of human capital. *Economics of Education Review*, 37, 204-212.
- Hanushek, E. A., Link, S. and Woessmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, 104, 212-232.
- Heyneman, S. P. (1997). Educational Choice in Eastern Europe and the Former Soviet Union: A Review Essay. *Education Economics*, 5(3), 333-339.
- Houston, R. A. (2011). *Literacy* [Online]. Available from: <http://iegego.eu/en/threads/backgrounds/literacy/robert-a-houston-literacy> [Accessed 3/24/2015].
- Lazear, E. P. (2001). Educational Production. *The Quarterly Journal of Economics*, 116(3), 777-803.
- Münich, D., Svejnar, J. and Terrell, K. (2005). Returns to Human Capital under the Communist Wage Grid and During the Transition to a Market Economy. *Review of Economics and Statistics*, 87(1), 100-123.
- Newell, A. and Reilly, B. (1999). Rates of Return to Educational Qualifications in the Transitional Economies. *Education Economics*, 7(1), 67-84.
- Roland, G. (2010). The Long-Run Weight of Communism or the Weight of Long-Run History? *Wider Working Papers* [Online], 2010/83.
- Roland, G. (2002). The Political Economy of Transition. *Journal of Economic Perspectives*, 16(1), 29-50.
- Roland, G. and Verdier, T. (2003). Law enforcement and transition. *European Economic Review*, 47(4), 669-685.
- Schütz, G., Ursprung, H. W. and Wößmann, L. (2008). Education Policy and Equality of Opportunity. *Kyklos*, 61(2), 279-308.
- Triventi, M. and Trivellato, P. (2009). Participation, performance and inequality in Italian higher education in the 20th century. *Higher Education*, 57(6), 681-702.
- Ulubasoglu, M. A. and Cardak, B. A. (2007). International comparisons of rural-urban educational attainment: Data and determinants. *European Economic Review*, 51(7), 1828-1857.

Tables

Table1: Countries and number of observations

E. Europe country	Nr. of obs.	W. Europe country	Nr. of obs.
Albania	1,143	Austria	1,262
Azerbaijan	1,067	Belgium	1,271
Armenia	1,237	Cyprus	868
Bulgaria	1,284	Northern Cyprus	358
Belarus	1,083	Denmark	1,309
Bosnia Herzegovina	907	Finland	875
Croatia	947	France	1,183
Czech Republic	1,433	W. Germany	823
Estonia	1,298	Greece	1,331
Georgia	1,221	Iceland	610
E. Germany	804	Ireland	731
Hungary	1,289	Italy	1,130
Kosovo	686	Luxembourg	1,296
Latvia	1,159	Malta	1,136
Lithuania	1,081	Netherlands	1,267
Macedonia	861	Norway	950
Moldova	1,235	Portugal	1,238
Montenegro	955	Spain	1,232
Poland	1,163	Sweden	928
Romania	792	Switzerland	1,083
Russian Federation	1,200	Turkey	1,608
Serbia	990	Great Britain	843
Slovak Republic	1,245	Northern Ireland	259
Slovenia	918		
Ukraine	1,276		
Total:	27,274	Total:	23,591

Table 2: Sample summary statistics

Dependent variables:	Mean	Std. Deviation	Min/Max
Primary educ.	0.976	0.152	0/1
Secondary educ.	0.617	0.486	0/1
Tertiary educ.	0.107	0.309	0/1
Secondary education:			0/1
Vocational	0.093	0.291	0/1
University preparation	0.524	0.499	0/1
Explanatory variables:			
Communism	0.317	0.465	0/1
Parental tertiary educ.	0.143	0.350	0/1
Parental secondary educ.	0.383	0.486	0/1
Parental primary educ.	0.907	0.290	0/1
Parental occupation	6.704	3.392	1/11
Sex	0.444	0.497	0/1
City	0.256	0.436	0/1
Time trend	81.221	17.513	19/112
Cohort 1 (15-24)	0.130	0.337	0/1
Cohort 2 (25-34)	0.178	0.383	0/1
Cohort 3 (35-44)	0.184	0.38	0/1
Cohort 4 (45-54)	0.184	0.387	0/1
Cohort 5 (55-64)	0.150	0.357	0/1
Cohort 6 (65+)	0.173	0.379	0/1

Note: sample size 50,865

Table 3: Construction of binary dependent variables

Highest education level attained respondent – categories:	Level of education			Secondary education type	
	Primary	Secondary	Tertiary	Vocational	University preparation
1. Inadequately complete primary	0	0	0	0	0
2. Completed (compulsory) primary	1	0	0	0	0
3. Incomplete secondary: vocational	1	0	0	0	0
4. Complete secondary: vocational	1	1	0	1	0
5. Incomplete secondary: univ. prep.	1	0	0	0	0
6. Complete secondary: univ. prep.	1	1	0	0	1
7. Some university without a degree	1	1	0	0	1
8. University with degree	1	1	1	0	1

Table 4: Respondent's highest educational attainments summarized by region

Highest educational level attained respondent	West Europe		East Europe		Total
	Nr. of Obs.	Ratio with population	Nr. of Obs.	Ratio with population	
Inadequately complete primary educ.	1,058	0.0417	261	0.0093	1,319
Completed (compulsory) primary educ.	3,782	0.1489	1,012	0.0360	4,794
Incomplete secondary vocational educ.	4,929	0.1941	3,889	0.1382	8,818
Complete secondary vocational educ.	1,642	0.0646	3,388	0.1204	5,030
Incomplete secondary: university prep.	2,724	0.1072	3,443	0.1224	6,167
Complete secondary: university prep.	4,992	0.1965	8,979	0.3191	13,971
Some university without a degree	4,254	0.1675	3,653	0.1298	7,907
University with degree	2,019	0.0795	3,515	0.1249	5,534
Total	25,400		28,140		53,540

Table 5: Parental highest educational attainments summarized by region

Highest educational level attained father/mother	West Europe		East Europe		Total
	Nr. of Obs.	Ratio with population	Nr. of Obs.	Ratio with population	
Inadequately complete primary educ.	3,247	0.1364	1,520	0.0554	4,767
Completed (compulsory) primary educ.	7,399	0.3109	4,107	0.1496	11,506
Incomplete secondary vocational educ.	4,404	0.1850	6,245	0.2274	10,649
Complete secondary vocational educ.	1,345	0.0565	3,286	0.1197	4,631
Incomplete secondary: university prep.	2,288	0.0961	2,419	0.0881	4,707
Complete secondary: university prep.	2,229	0.0937	5,420	0.1974	7,649
Some university without degree	1,799	0.0756	2,175	0.0792	3,974
University with degree	1,090	0.0458	2,285	0.0832	3,375
Total	23,801		27,457		51,258

Tables 6: Tertiary, Secondary and Primary education by age group and part of Europe for women

	15-24		25-34		35-44		45-54		55-64		65+	
	East	West	East	West	East	West	East	West	East	West	East	West
Tertiary												
Mean	0.070	0.027	0.182	0.127	0.156	0.105	0.142	0.077	0.139	0.046	0.083	0.021
Std. deviation	0.255	0.161	0.386	0.333	0.363	0.307	0.349	0.266	0.346	0.209	0.277	0.143
Ranksum test	0.000***		0.000***		0.000***		0.000***		0.000***		0.000***	
Secondary												
Mean	0.761	0.650	0.786	0.665	0.745	0.594	0.725	0.512	0.665	0.386	0.475	0.221
Std. deviation	0.427	0.477	0.410	0.472	0.436	0.491	0.447	0.500	0.472	0.487	0.499	0.415
Ranksum test	0.000***		0.000***		0.000***		0.000***		0.000***		0.000***	
Primary												
Mean	0.998	0.994	0.996	0.988	0.998	0.987	0.998	0.972	0.991	0.942	0.952	0.839
Std. deviation	0.042	0.080	0.059	0.110	0.042	0.114	0.048	0.165	0.093	0.234	0.214	0.367
Ranksum test	0.0196**		0.0003***		0.000***		0.000***		0.000***		0.000***	
Obs.	2281	1391	2828	2205	2840	2645	2987	2509	2302	2228	2672	2854

Tables 7: Tertiary, Secondary and Primary education by age group and part of Europe for men

	15-24		25-34		35-44		45-54		55-64		65+	
	East	West	East	West	East	West	East	West	East	West	East	West
Tertiary												
Mean	0.069	0.022	0.150	0.134	0.124	0.118	0.116	0.109	0.113	0.087	0.121	0.066
Std. deviation	0.253	0.148	0.357	0.341	0.330	0.322	0.320	0.311	0.316	0.282	0.326	0.248
Ranksum test	0.000***		0.1315		0.5067		0.4439		0.0116**		0.000***	
Secondary												
Mean	0.709	0.615	0.744	0.680	0.738	0.589	0.707	0.568	0.653	0.488	0.575	0.335
Std. deviation	0.454	0.487	0.436	0.467	0.440	0.492	0.455	0.495	0.476	0.500	0.494	0.472
Ranksum test	0.000***		0.000***		0.000***		0.000***		0.000***		0.000***	
Primary												
Mean	0.996	0.997	0.996	0.994	0.998	0.989	0.996	0.989	0.986	0.968	0.982	0.914
Std. deviation	0.065	0.055	0.065	0.075	0.047	0.105	0.062	0.103	0.117	0.175	0.133	0.280
Ranksum test	0.5962		0.4931		0.0003***		0.008***		0.0004***		0.000***	
Obs.	1889	1304	2385	1948	2223	2059	2318	1970	1670	1859	1666	2346

Tables 8: Vocational and University preparation secondary education by age group and part of Europe for women

	15-24		25-34		35-44		45-54		55-64		65+	
	East	West	East	West	East	West	East	West	East	West	East	West
Vocational/technical secondary education												
Mean	0.074	0.070	0.079	0.069	0.111	0.063	0.123	0.068	0.113	0.053	0.088	0.034
Std. deviation	0.262	0.256	0.270	0.253	0.314	0.243	0.328	0.251	0.317	0.225	0.283	0.182
Ranksum test	0.6805		0.1690		0.000***		0.000***		0.000***		0.000***	
University preparation secondary education												
Mean	0.687	0.579	0.707	0.596	0.634	0.531	0.602	0.444	0.552	0.333	0.387	0.187
Std. deviation	0.646	0.494	0.455	0.491	0.482	0.499	0.490	0.497	0.497	0.471	0.487	0.390
Ranksum test	0.000***		0.000***		0.000***		0.000***		0.000***		0.000***	
Obs.	2281	1391	2828	2205	2840	2645	2987	2509	2302	2228	2672	2854

Tables 9: Vocational and University preparation secondary education by age group and part of Europe for men

	15-24		25-34		35-44		45-54		55-64		65+	
	East	West	East	West	East	West	East	West	East	West	East	West
Vocational/technical secondary education												
Mean	0.096	0.073	0.123	0.084	0.163	0.080	0.179	0.076	0.188	0.077	0.140	0.052
Std. deviation	0.294	0.260	0.329	0.277	0.369	0.271	0.384	0.264	0.391	0.267	0.348	0.221
Ranksum test	0.0232**		0.000***		0.000***		0.000***		0.000***		0.000***	
University preparation secondary education												
Mean	0.614	0.542	0.621	0.597	0.575	0.509	0.527	0.492	0.465	0.410	0.435	0.283
Std. deviation	0.487	0.498	0.485	0.491	0.494	0.500	0.499	0.500	0.499	0.492	0.496	0.451
Ranksum test	0.0001***		0.1007		0.000***		0.0231**		0.0012***		0.000***	
Obs.	1889	1304	2385	1948	2223	2059	2318	1970	1670	1859	1666	2346

Table 10: Results for Secondary education attainments

Model	LPM			Probit		SNP
	1)	2)	3)	4)	5)	6)
Communism		0.053*** (0.008)		0.048*** (0.006)		0.050*** (0.000)
Ptertiary	0.017 (0.146)	0.018 (0.122)	0.040*** (0.001)	0.040*** (0.001)	0.048*** (0.000)	0.047*** (0.000)
Psecondary	0.159*** (0.000)	0.159*** (0.000)	0.148*** (0.000)	0.149*** (0.000)	0.149*** (0.000)	0.152*** (0.000)
Pprimary	0.206*** (0.000)	0.204*** (0.000)	0.189*** (0.000)	0.187*** (0.000)	0.160*** (0.000)	0.157*** (0.000)
Poccupation	-0.017*** (0.000)	-0.017*** (0.000)	-0.017*** (0.000)	-0.017*** (0.000)	-0.017*** (0.000)	-0.017*** (0.000)
Sex	0.013 (0.182)	0.012 (0.185)	0.012 (0.203)	0.011 (0.209)	0.016*** (0.000)	0.016*** (0.000)
Cohort2	0.101*** (0.000)	0.091*** (0.000)	0.106*** (0.000)	0.097*** (0.000)	0.105*** (0.000)	0.094*** (0.000)
Cohort3	0.130*** (0.000)	0.095*** (0.003)	0.130*** (0.000)	0.100*** (0.000)	0.129*** (0.000)	0.096*** (0.000)
Cohort4	0.165*** (0.000)	0.127*** (0.000)	0.156*** (0.000)	0.124*** (0.000)	0.161*** (0.000)	0.122*** (0.000)
Cohort5	0.150*** (0.000)	0.110*** (0.006)	0.141*** (0.000)	0.108*** (0.001)	0.151*** (0.000)	0.108*** (0.000)
Cohort6	0.096** (0.015)	0.071* (0.081)	0.097*** (0.002)	0.074** (0.038)	0.114*** (0.000)	0.081*** (0.002)
City	0.056*** (0.000)	0.056*** (0.000)	0.061*** (0.000)	0.061*** (0.000)	0.063*** (0.000)	0.064*** (0.000)
T	0.004*** (0.000)	0.004*** (0.000)	0.004*** (0.000)	0.004*** (0.000)	0.004*** (0.000)	0.004*** (0.000)
Cons_	-0.245*** (0.001)	-0.199*** (0.007)				
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes
R2/Pseudo R2	0.2773	0.2782	0.2387	0.2395		
Nr. of obs.	50865	50865	50865	50865	50865	50865

Note: Communism defined in methodology section; Ptertiary, Psecondary, Pprimary – dummies for parents' education; Poccupation – proxy for parental income – parental occupation: 1. I: Higher controllers, 2. II: Lower controllers, 3. IIIa: Routine Nonmanual, 4. IIIb: lower sales-service, 5. IAa: Self-employed with employees, 6. IVb: Self-employed with no employees, 7. V: Manual supervisors, 8. VI: Skilled worker, 9. VIIa: unskilled worker, 10. VIIb: farm labour, 11. IVc: self-employed farmer; sex=1 if male and =0 if female; Cohorts 1-6 – by age respectively: 15-24, 25-34, 35-44, 45-54, 55-64, 65+, Cohort1 omitted; City = 1 if has 50,000 people or more, 0 otherwise; T = (year of birth)-1881; LPM – linear probability model, SNP – semi-nonparametric estimations; (p-values in parentheses). * - significant at 10% significance level; ** - significant at 5% significance level; *** - significant at 1% significance level.

Table 11: Results for Vocational and University Preparation secondary educations

Model	Bivariate Probit				Bivariate Semi-nonparametric estimations (SNP2)			
	Vocational	University Preparation	Vocational	University Preparation	Vocational	University Preparation	Vocational	University Preparation
Communism			0.014** (0.047)	0.015 (0.371)			0.016*** (0.000)	0.026*** (0.000)
Ptertiary (D)	-0.044*** (0.000)	0.090*** (0.000)	-0.044*** (0.000)	0.090*** (0.000)	-0.674*** (0.000)	0.575*** (0.000)	-0.043*** (0.000)	0.125*** (0.000)
Psecondary (D)	-0.005 (0.267)	0.139*** (0.000)	-0.006 (0.233)	0.139*** (0.000)	-0.080* (0.073)	0.647*** (0.000)	-0.005* (0.072)	0.140*** (0.000)
Pprimary (D)	0.031*** (0.000)	0.187*** (0.000)	0.030*** (0.000)	0.186*** (0.000)	0.447*** (0.000)	0.643*** (0.000)	0.028*** (0.000)	0.139*** (0.000)
Poccupation	0.005*** (0.000)	-0.019*** (0.000)	0.005*** (0.000)	-0.019*** (0.000)	0.045*** (0.000)	-0.095*** (0.000)	0.003*** (0.000)	-0.020*** (0.000)
Sex (D)	0.018*** (0.000)	-0.020** (0.044)	0.018*** (0.000)	-0.020** (0.044)	0.238*** (0.000)	-0.015 (0.384)	0.015*** (0.000)	-0.003 (0.421)
Cohort2	0.014* (0.096)	0.081*** (0.000)	0.014 (0.117)	0.080*** (0.000)	0.185*** (0.002)	0.223*** (0.000)	0.012*** (0.006)	0.045*** (0.000)
Cohort3	0.037*** (0.006)	0.084*** (0.000)	0.028* (0.053)	0.076*** (0.002)	0.349*** (0.000)	0.095** (0.010)	0.013*** (0.005)	0.008 (0.321)
Cohort4	0.049*** (0.003)	0.096*** (0.000)	0.041** (0.017)	0.089*** (0.001)	0.426*** (0.000)	0.039 (0.335)	0.019*** (0.000)	-0.004 (0.689)
Cohort5	0.053*** (0.007)	0.080*** (0.004)	0.047** (0.016)	0.074** (0.017)	0.313*** (0.000)	-0.132*** (0.004)	0.013** (0.032)	-0.040*** (0.000)
Cohort6	0.027 (0.232)	0.029 (0.403)	0.030 (0.160)	0.031 (0.399)	-0.090 (0.412)	-0.463*** (0.000)	-0.006 (0.427)	-0.102*** (0.000)
City	-0.023*** (0.000)	0.078*** (0.000)	-0.022*** (0.000)	0.078*** (0.000)	-0.278*** (0.000)	0.365*** (0.000)	-0.018*** (0.000)	0.079*** (0.000)
T	0.001 (0.136)	0.003*** (0.000)	0.001** (0.045)	0.003*** (0.000)	-0.001 (0.691)	0.001 (0.251)	0.00003 (0.792)	0.0003 (0.144)
Country dummies		Yes		Yes		Yes		Yes
Athrho		-3.189758 (0.000)***		-3.166512 (0.000)***				
rho						-0.7608118		-0.7638831
Wald Chi2 test		0.0000***		0.0000***		0.0000***		0.0000***
Nr. of obs.		50865		50865		50865		50865

Note: Communism defined in methodology section; Ptertiary, Psecondary, Pprimary – dummies for parents' education; Poccupation – proxy for parental income – parental occupation: 1. I: Higher controllers, 2. II: Lower controllers, 3. IIIa: Routine Nonmanual, 4. IIIb: lower sales-service, 5. IAa: Self-employed with employees, 6. IVb: Self-employed with no employees, 7. V: Manual supervisors, 8. VI: Skilled worker, 9. VIIa: unskilled worker, 10. VIIb: farm labour, 11. IVc: self-employed farmer; sex=1 if male and =0 if female; Cohorts 1-6 – by age respectively: 15-24, 25-34, 35-44, 45-54, 55-64, 65+, Cohort1 omitted; City = 1 if has 50,000 people or more, 0 otherwise; T = (year of birth)-1881; (p-values in parentheses).

* - significant at 10% significance level; ** - significant at 5% significance level; *** - significant at 1% significance level.

Table 12: Results for Tertiary education attainments

Model	LPM		Probit		SNP	
	1)	2)	3)	4)	5)	6)
Communism		0.004 (0.640)		-0.005 (0.440)		-0.006 (0.205)
Ptertiary	0.122*** (0.000)	0.122*** (0.000)	0.070*** (0.000)	0.070*** (0.000)	0.059*** (0.000)	0.058*** (0.000)
Psecondary	0.036*** (0.000)	0.036*** (0.000)	0.037*** (0.000)	0.037*** (0.000)	0.037*** (0.000)	0.037*** (0.000)
Pprimary	0.021** (0.035)	0.021** (0.036)	0.049*** (0.000)	0.049*** (0.000)	0.062*** (0.000)	0.062*** (0.000)
Poccupation	-0.010*** (0.000)	-0.010*** (0.000)	-0.009*** (0.000)	-0.009*** (0.000)	-0.009*** (0.000)	-0.009*** (0.000)
Sex	0.005 (0.210)	0.005 (0.211)	0.007 (0.111)	0.007 (0.112)	0.007*** (0.006)	0.007*** (0.006)
Cohort2	0.117*** (0.000)	0.117*** (0.000)	0.150*** (0.000)	0.152*** (0.000)	0.120*** (0.000)	0.121*** (0.000)
Cohort3	0.111*** (0.000)	0.109*** (0.000)	0.143*** (0.000)	0.149*** (0.000)	0.117*** (0.000)	0.121*** (0.000)
Cohort4	0.109*** (0.000)	0.107*** (0.000)	0.142*** (0.000)	0.148*** (0.000)	0.118*** (0.000)	0.122*** (0.000)
Cohort5	0.103*** (0.000)	0.101*** (0.000)	0.138*** (0.000)	0.144*** (0.000)	0.114*** (0.000)	0.118*** (0.000)
Cohort6	0.086*** (0.000)	0.084*** (0.000)	0.111*** (0.000)	0.114*** (0.000)	0.097*** (0.000)	0.100*** (0.000)
City	0.059*** (0.000)	0.059*** (0.000)	0.052*** (0.000)	0.052*** (0.000)	0.048*** (0.000)	0.048*** (0.000)
T	-0.0001 (0.757)	-0.0001 (0.720)	-0.0001 (0.877)	-0.00003 (0.915)	0.00002 (0.946)	0.00004 (0.914)
Cons_	-0.025 (0.375)	-0.022 (0.458)				
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes
R2/Pseudo R2	0.1645	0.1645	0.2404	0.2404		
Nr. of obs.	50865	50865	50865	50865	50865	50865

Note Communism defined in methodology section; Ptertiary, Psecondary, Pprimary – dummies for parents' education; Poccupation – proxy for parental income – parental occupation: 1. I: Higher controllers, 2. II: Lower controllers, 3. IIIa: Routine Nonmanual, 4. IIIb: lower sales-service, 5. IAa: Self-employed with employees, 6. IVb: Self-employed with no employees, 7. V: Manual supervisors, 8. VI: Skilled worker, 9. VIIa: unskilled worker, 10. VIIb: farm labour, 11. IVc: self-employed farmer; sex=1 if male and =0 if female; Cohorts 1-6 – by age respectively: 15-24, 25-34, 35-44, 45-54, 55-64, 65+, Cohort1 omitted; City = 1 if has 50,000 people or more, 0 otherwise; T = (year of birth)-1881;

LPM – linear probability model, SNP – semi-nonparametric estimations; (p-values in parentheses).

* - significant at 10% significance level; ** - significant at 5% significance level; *** - significant at 1% significance level.

Table 13: Results for Primary education attainments

Model	LPM		Probit		SNP	
	1)	2)	3)	4)	5)	6)
Communism		-0.004 (0.579)		0.009** (0.023)		0.006** (0.046)
Ptertiary	-0.007*** (0.001)	-0.007*** (0.001)	-0.012 (0.126)	-0.012 (0.130)	-0.012* (0.077)	-0.012* (0.071)
Psecondary	-0.011*** (0.002)	-0.012*** (0.002)	0.006 (0.113)	0.006 (0.107)	0.016*** (0.005)	0.015*** (0.005)
Pprimary	0.149*** (0.000)	0.149*** (0.000)	0.059*** (0.000)	0.058*** (0.000)	0.029*** (0.000)	0.030*** (0.000)
Poccupation	-0.002*** (0.001)	-0.002*** (0.001)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Sex	0.011** (0.018)	0.011** (0.018)	0.013*** (0.001)	0.013*** (0.001)	0.013*** (0.000)	0.013*** (0.000)
Cohort2	0.024*** (0.003)	0.024*** (0.005)	0.012*** (0.000)	0.011*** (0.003)	0.011*** (0.006)	0.010** (0.015)
Cohort3	0.049*** (0.003)	0.052*** (0.009)	0.021*** (0.000)	0.018*** (0.000)	0.020*** (0.000)	0.018*** (0.000)
Cohort4	0.072*** (0.004)	0.074*** (0.008)	0.027*** (0.000)	0.023*** (0.000)	0.024*** (0.000)	0.022*** (0.000)
Cohort5	0.084*** (0.006)	0.087** (0.011)	0.024*** (0.001)	0.020 (0.010)	0.019*** (0.000)	0.016*** (0.001)
Cohort6	0.076** (0.016)	0.078** (0.020)	0.023** (0.034)	0.018* (0.092)	0.017*** (0.002)	0.014*** (0.009)
City	0.007*** (0.004)	0.007*** (0.003)	0.008*** (0.002)	0.008*** (0.002)	0.006*** (0.000)	0.006*** (0.000)
T	0.003*** (0.003)	0.003*** (0.003)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Cons_	0.604*** (0.000)	0.601*** (0.000)				
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes
R2/Pseudo R2	0.1769	0.1770	0.4260	0.4268		
Nr. of obs.	50865	50865	38862	38862	50865	50865

Note: Communism defined in methodology section; Ptertiary, Psecondary, Pprimary – dummies for parents' education; Poccupation – proxy for parental income – parental occupation: 1. I: Higher controllers, 2. II: Lower controllers, 3. IIIa: Routine Nonmanual, 4. IIIb: lower sales-service, 5. IAa: Self-employed with employees, 6. IVb: Self-employed with no employees, 7. V: Manual supervisors, 8. VI: Skilled worker, 9. VIIa: unskilled worker, 10. VIIb: farm labour, 11. IVc: self-employed farmer; sex=1 if male and =0 if female; Cohorts 1-6 – by age respectively: 15-24, 25-34, 35-44, 45-54, 55-64, 65+, Cohort1 omitted; City = 1 if has 50,000 people or more, 0 otherwise; T = (year of birth)-1881;

LPM – linear probability model, SNP – semi-nonparametric estimations; (p-values in parentheses).

* - significant at 10% significance level; ** - significant at 5% significance level; *** - significant at 1% significance level.

Appendix

- ‘Highest educational level attained by respondent’ takes these values: 1 – respondent has not had any or has not completed primary education; 2 – completed primary education; 3 – incomplete vocational secondary education; 4 – completed secondary vocational education; 5 – incomplete secondary university preparation education; 6 – completed secondary university preparation education; 7 – incomplete university, university without a degree; 8 - completed university, having a degree.

- ‘Parents’ occupation’ takes these values: 1. Higher controllers, 2. Lower controllers, 3, Routine non-manual, 4. Lower sales-service, 5. Self-employed with employees, 6. Self-employed with no employees, 7. Manual supervisors, 8. Skilled worker, 9. Unskilled worker, 10. Farm labour, 11. Self-employed farmer.

Chapter II

Drinking under Communism: why do Alcohol Consumption Habits in Eastern Europe Differ from the West in the Long-Run?¹⁷

Gintarė Mališauskaitė

University of Kent, Canterbury, UK

Alexander Klein

University of Kent, Canterbury, UK

Abstract

The paper looks into possible explanations for the differences between Eastern and Western Europe alcohol consumption behaviour even twenty years after the collapse of the Soviet regime. It suggests these differences can be viewed as an expression of cultural habits. We explore different ways of defining exposure to the communist regime: using number of years a person spent under the regime and also a dummy indicator for spending formative years (18-25) in it. We find both to be strong factors in explaining alcohol consumption behaviour. We consider differences in frequency of alcohol consumption and binge drinking using European Health Interview Survey (EHIS) micro data from Eurostat. Estimations are run with ordered probit model for men and women separately. Evidence suggests a statistically significant effect of experiencing communist regimes, which is larger for women's alcohol consumption frequency than for men's. It is also the most important factor in explaining more frequent male binge drinking. These effects hold after controlling for socio-economic, country level and time characteristics. This suggests the attitudes towards alcohol consumption could be more permissive in the former Eastern Bloc countries.

Keywords: alcohol consumption, cultural habits, communism, Eastern Bloc, Western Europe

¹⁷ This paper is based on data from Eurostat, *European Health Interview Survey (EHIS), wave 1 (2006-2009)*. The responsibility for all conclusions drawn from the data lies entirely with the authors.

This chapter has been removed since it was published in the Journal of Comparative Economics:

Malisauskaite, G., Klein, A. (2018). Drinking Under Communism: why do Alcohol Consumption Habits in Eastern Europe Differ from the West in the Long-Run? *Journal of Comparative Economics*, 46(3), 821-837

Available on KAR:

<https://kar.kent.ac.uk/69085/>

Chapter III

Do People from Eastern Bloc Perceive their Health Differently? Determinants and Impact of Health Perception Gap¹⁸

Gintarė Mališauskaitė

University of Kent, Canterbury, UK

Abstract

This article looks into the long-term effects of having lived in a communist regime in Eastern Bloc during one's formative years on self-rated health, health measured by Elixhauser Comorbidity Index (ECI) and Charlson Comorbidity Index (CCI), and the derived health perception gap (HPG), the difference between own self-rated health and the average levels of self-rated health by different categories of the two mentioned comorbidity indices. The associations between this gap and a range of health related behaviour outcomes is explored. This study draws attention to a noticeably more negative perception of own health amongst the respondents who have lived in the communist regime during the formative years. The relationships are estimated using different estimators, including ordered probit, OLS and machine learning techniques. We find the effect of communism on perception of own health to be significant. Study uses 1st wave of European Health Interview Survey (EHIS) data.

Keywords: self-rated health, Elixhauser comorbidity index, Charlson comorbidity index, health perception gap, communist regime, Eastern Bloc, Western Europe

¹⁸ This paper is based on data from Eurostat, *European Health Interview Survey (EHIS), wave 1 (2006-2009)*. The responsibility for all conclusions drawn from the data lies entirely with the author.

1. Introduction

This article aims to answer a question whether a political regime impact could have a long-term effect on the perception of own health. To capture the exposure of individuals to communist regimes and the formation of one's perception of health, we use the insights from the psychology literature. Research has shown that the most important period of a person's life towards *visa-a-vis* the adoption of norms and attitudes are so-called the formative years between the age of 18 and 25 (Krosnick and Alwin 1989). Norms and values are found to not change much after this stage later in life. Thus differentiating between Eastern Bloc and Western Europe would not be sufficient in determining the effect of living in the communist regime in Eastern Europe on the perception of own health, and our best guess is to indicate for the exposure during the formative years. This does not mean we assume that other years of a person's life or the length of time spent in the regime were not important, but in this study we focus on the experience of the communist regime during the formative years as we believe this will allow capturing the strongest effect on perception towards health. We control for country-specific conditions, different years and present socio-economic conditions so that we can isolate this effect from other influencing factors as convincingly as possible. It is important to clarify that it is not the experience of economic struggles, insecurity or unemployment during transition period that we are looking into, also this is not about the political beliefs a person holds, or agreement or disagreement with the regime at the time of the regime or later on. We are specifically interested in the latent long-run effect on perception of living in this regime, being exposed to its ideological climate and everything that it comprised.

In order to compare the health perception gap between Eastern Bloc and other European countries, and determine if the group of people who were exposed to communist regime rule during their formative years are any different from the ones that were not, we approached this question in a few stages. Firstly, we looked at the determinants of self-rated health (henceforth SRH) to see if there are any noticeable differences between those who lived in a communist regime during formative years and those that did not. Secondly, we calculated Elixhauser Comorbidity Index (henceforth ECI) and Charlson Comorbidity Index (henceforth CCI), which are frequently used in medical research to allow for a more objective assessment of individual's health than self-assessment. We use these indices as dependent variables to analyse if there are any noticeable differences between those with communist regime experience in formative years and those without. If there are any

noticeable differences between the effects of communist regime experience on self-assessed health and comorbidity indices then there is reason to suspect the existence of perception bias. This was followed by the individual health perception gap calculation (henceforth HPG). It was derived as a difference between own self-rated health and the average levels of self-rated health based on the same category of above mentioned comorbidity indices. The measure was used as a dependent variable to estimate the effect of the exposure to the communist regime. Finally, it was used as an explanatory variable to see if there is any relationship between HPG and variables related to health behaviour. Literature from the fields of economics and psychology suggest that perception biases tend to affect person's behaviour, especially the larger they are.

There are several reasons why an experience of the communist regimes in the former Eastern Bloc is important to research. Due to relatively isolated nature, specific governing principles and policies, and a long period of time the regime lasted, this made its experience a natural economic experiment, the consequences of which were not yet much researched. It is very likely such experience could have caused changes in behaviour, preferences, attitudes, all of which would have an economic impact through influencing social and human capital. Such research then could help explain economic outcomes related to former Eastern Bloc countries, add to the existing knowledge of the long-run effects of political and economic regimes and systems and, more generally, contribute to social capital research in its determinants and effects. To our best knowledge this research question is unique. Consequentially these findings could be of interest to policy makers.

The study uses 1st wave of European Health Interview Survey (henceforth EHIS) data. The following section reviews available literature related to the topic, 3rd discusses methodology, 4th summarizes data and descriptive statistics, 5th analyses the results, and section 6 concludes.

2. Literature review

Background information

Migration between Western Europe and Eastern Bloc was restricted during the times of communism. In addition, the exchange of news and information between the two parts of Europe was limited, making this as close to a natural experiment in history as possible. The characteristics of the communist regimes included exposure to propaganda, censorship, lack

of political transparency, scarcity of some goods and resources, fear of violence or other sanctions for not agreeing with some principles of the regime, or even fear of what could be interpreted as such disagreement. On the other hand, the regime offered a relative job security and easier access to jobs after acquiring a profession, low prices for the available goods, encouraged attainment of free education, higher levels of redistribution by the government, help acquiring housing and promoted spirit of equality and comradeship. There is evidence of traumatic experiences having a long-term scarring effect on the behaviour of the recipient (Desivilya et al. 1996), which suggests that an experience of a different lifestyle in a group of countries for about 40-60 years could also have an effect on cultural norms, values and perceptions.

This article brings together the research from different disciplines: economic studies of the long-run consequences of economic systems and political regimes; psychology research on the process of socialisation, the development of a personality, and the determinants and importance of internal biases; and finally medical research on measuring health. Each strand of literature is presented respectively.

Economics literature

Culture and economics

There are many different suggestions of what at least in part captures cultural differences between countries and how these differences can be explained. Two main strands of this research are: economics and religion and social capital studies. Religious differences between Catholics and Protestants and its effect on economics outcomes were researched by Becker and Woessman (2009), Blum and Dudley (2001) and Grier (1997). Barro and McCleary (2003) suggested a relationship between religious beliefs, church attendance and economics outcomes. Bisin and Verdier (2000) looked into transmission of cultural traits by observing socialisation of immigrants and cultural minorities. Tabellini (2008) was interested in trust and respect and their effect on performance of institutions in countries. Grosjean (2011) stressed the importance of common history and that conditions trust and can reduce cultural distance. Gorodnichenko and Roland (2011) suggested different measures to express culture: individualism, power distance, masculinity and uncertainty avoidance, and found individualism in particular to be important for long-term economic growth in a country. Ananyev and Guriev (2016) explored the effect of economic shocks towards social capital, trust in particular, and find trust responding quickly but the effect being potentially persistent. Social capital is seen as a key determinant of the production of

human capital, which then affects development and growth in a country. This literature is still at its infancy due to measurement difficulties, so adding to the research of what could help understand what factors could shape culture, how common history experiences could help explain perception of a person, and how this translates to economic outcomes.

A small sub-strand of Cultural Economics literature deals with the outcomes of the centralized communist regime in Eastern Europe. There are studies available on institutional changes in post-communist countries (Roland 2002), law enforcement in transition economies (Roland 2003), political system effects on cultural values and the length of this impact (Roland 2010), effect on redistribution preferences between Eastern and Western Germany (Alesina and Fuchs-Schündeln 2007), returns to education (Münich et al. 2005), etc. This study would contribute to this strand of literature by suggesting a new route of influence – through affecting perception of own health.

Financial economics – role of optimism

Research on perceptions, internal biases and their influence is closest to the question of the role of optimism, its measurement and impact. It was mostly researched in Financial Economics literature regarding investments and financial decisions. Puri and Robinson (2007) derived an optimism measure comparing self-reported life expectancy with statistical information and this acquired optimism measure was positively correlated with other indications of positive attitudes. They found that people that were deemed to be more optimistic by this measure worked harder, thought they would retire later, were more likely to remarry, invest in individual stocks and were saving more. It was suggested, a distinction should be made between moderate and extreme optimism, moderate optimists exhibiting more prudent financial and health related behaviour and higher capability of self-control, the case being opposite for extreme optimists. Coval and Thakor (2005) used the idea of distinction between optimists, rational agents, and pessimists to explain financial intermediation, suggesting that optimists would frequently become entrepreneurs and equity holders, rational agents take second ranking type of risk – risky debt, and finally pessimists would take on riskless debt. Wang et al. (2017) look into relationship between optimism and investment decisions and suggest that more optimistic investors would increase portfolio delegation while portfolio-managers would decrease investment in riskier assets which then would result in lower returns for the investor and more certain higher compensation for the managers, overall higher levels of optimism resulting in higher moral hazard. Heaton (2002) explores corporate finance model in a theoretical paper and looks into relationship between

optimism levels of managers and their investment decisions, proposing that optimistic managers underinvest in other capital market firm's risky securities believing them to be overvalued, and overinvest in own firm's projects believing them to be undervalued. Sen and Tumarkin (2015) also research how optimism affects executive managers' investment decisions, concluding that more optimistic executives are more likely to overestimate their company's future, and more likely to short the stock closer to expiration, retain some shares from exercising options in comparison to less optimistic managers. This paper proposes a novel measure of the differences in perception and a new direction for its application – health behaviour related outcomes, which could translate into country level aggregate consequences.

Health economics

Since health is an important determinant of human capital and labour supply, the findings of health and labour economics crossovers should also be considered. Trevisan and Zantomio (2016) looked into the effect of health shocks such as myocardial infarction, stroke or cancer on employment decisions of older workers in Europe and they find that such health shock on average doubles the risk of leaving the labour force and could be followed by deterioration of physical and mental health and shorter perceived life expectancy; the largest reductions were observed in Nordic and Eastern countries and England and authors partially explain it through generosity of social security. García-Gómez (2011) finds that health shocks frequently end up in withdrawal from employment into disability and that these effects tend to be largest in Netherlands, Denmark, Spain, and Ireland, and smallest in France and Italy. Sudden illness and hospitalisation tend to lower employment probability by 7 percent and personal income in two years after the shock – by 5 percent, recovery of income or employment is not found and it effects household income by more than the loss of income of the disabled person in the Netherlands (García-Gómez et al. 2013). Similar findings were made by Jeon and Pohl (2017) who observed that employment and income of Canadians' whose spouses were diagnosed with cancer tend to fall which they explain by increased provision of care and spending more time with the ill spouse. Qin and Wang (2015) find self-reported health affecting employment decisions – urban Chinese aged 40-59 and find that those with self-identified bad health were 34% more likely to retire early, 105% more likely to be unemployed, and 28% more likely to be homemakers in comparison to those who reported good health; they also observed that obesity is not one of the significant factors determining retirement as opposed to the findings for developed countries; significant factors

determining being out of labour force were found to be being female, having lower education, and older age. Pit et al. (2010) look into retirement reasons for 45-64 year old Australians, they discover that those women who have been diagnosed with thrombosis, depression, osteoarthritis or cancer were twice more likely to retire early; for men a greater number of health issues were related to early retirement: stroke, cancer, osteoarthritis, depression, anxiety, heart disease; for both genders the strongest relationship was found between early retirement and self-reported health status. The relationship between employment outcomes and health can cause some concerns, one of them is a likely endogeneity of health which would result in an overestimated effect towards labour outcomes, allowing for the endogeneity of health, a positive relationship is found between health and employment, but the effect of employment is negative for the health of males while positive for females (Cai 2010). Baker et al. (2004) raise concerns regarding objectivity of self-reported health status after comparing it with respondents' medical records and finding large response errors that could result in biased estimates if self-reported health is used as an explanatory variable. Benítez-Silva et al. (2004) try to measure the bias of self-reported health using the data regarding disability status. Health economics is a rich area of research, only a small proportion of available research was discussed here.

Medical and policy research

Self-rated health is frequently taken with a pinch of salt, but there is also evidence of it being a good predictor of health outcomes and mortality, even if taken independently of objective health status. In fact, it was found to be second to age in the strength of its effect (Mossey and Shapiro 1982). A protective effect towards health was observed if health is rated more positively, the association was found between more positive own health ratings and a level of moral and probability of returning to employment (Mossey and Shapiro 1982). Poorer ratings of self-rated health and quality of life also seem to be strongly related to chronic conditions occurring together (McDaid et al. 2013). Other factors associated with worse own health ratings are: being older, born abroad, having low/medium education, low emotional support, economic problems, low trust, being never married or divorced (Lindström 2009). It is also observed that married men tend to report illness more often and engage less in unhealthy practices than the ones that never married, are separated, divorced, or widowed (Bourne 2009). This could be related to men rating their health on average as better than females and females overall more commonly experiencing symptoms related to psychological problems than men (Vaez and Laflamme 2002). There is evidence that living

in a more disadvantaged area could also have an effect on the rating of self-assessed health being worse (Brown 2007), and that there is evidence of a country of birth having a large impact on self-rated physical and mental health (Vaez and Laflamme 2002).

Medical research regarding comorbidity indices is discussed in detail in methodology section.

Psychology research

Research on internal biases received much more attention amongst psychologists. Khallad (2013) found a negative relationship between dispositional optimism and physical symptoms reporting, this research also looked into importance of cultural differences for a sample of American and Jordanian college students, and allowing for gender and socioeconomic differences. Findings include Jordanian women reporting physical symptoms more, and symptom reporting being more common amongst lower socioeconomic status having Jordanians, but overall no significant differences between Americans and Jordanians or American men and women. Though it seems that the measure of optimism itself was not affected by cultural differences, gender or socioeconomic status. Scheier et al. (1994) report a negative relationship between optimism and depression even after controlling for neuroticism, anxiety, self-mastery, and self-esteem. Similar relationship between optimism and depression was reported by Jiang (2016), this study suggests that belief in a just world stimulates optimism and gratitude which in turn affect favourably mental health: subjective well-being and depression. Zhao et al. (2015) found an association between dispositional optimism and self-framing, proposing that more optimistic individuals use more positive meaning having words when describing themselves, opposite being true for less optimistic individuals; also these attitudes were relevant to how individual processes information that relates to risky decision making in regards to health. Optimism was found to have an effect on education as well in Solberg Nes et al. (2009) study. They find that higher dispositional and academic optimism were interconnected with lower chance of dropping out of higher education, better motivation and adjustment; academic optimism also positively relates to grade point average. These studies offer insight how optimistic attitudes might influence the way we live, so it is reasonable to assume they could affect health related behaviours as well.

3. Methodology

We approached our research question in a few steps. Firstly, we looked at determinants of self-rated health (SRH) including the communism dummy. Secondly, we calculated Elixhauser Comorbidity Index (ECI) and Charlson Comorbidity Index (CCI) and studied the factors that impact the comorbidity indices. Thirdly, the individual health perception gap was calculated (HPG) using both indices mentioned, and its determinants explored. Finally, we used the constructed health perception gap as an explanatory variable to see if there is any relationship between HPG and variables related to health behaviour, such as: receiving benefits payments, reported difficulty to do things on their own, receiving personal or household help, times consulted GP per last month, if decided against being hospitalized or against consulting a specialist even if really needed to. Each step is explained in detail.

Self-rated health estimation

Self-rated health equation is determined as follows:

$$SRH_i = \alpha_0 + \alpha_1 CR_i + \sum_{j=1}^n \alpha_j Controls_i^j + \varepsilon_i \quad (1)$$

Self-rated health is a categorical variable ranging from ‘*very bad*’ health to ‘*very good*’ health (1-5), so increase in this measure indicates improvement in perception of own health rating. CR stands for communist regime dummy variable, which specifies those people who were between 18-25 years old, an indication for formative years – the time of increased socialisation and adoption of norms (Krosnick and Alwin 1989), during the time when communist central regime was still in power in Eastern Bloc. The aim of this dummy is to learn how much the experience of living in the communist regime and experiencing everything that defined it, affected person’s perception of own health¹⁹. A lengthy list of control variables was used to help identify the most important factors affecting own health perception and to make sure the effect of communism dummy is not overestimated due to not accounting for a possible route of influence. The full list of controls includes: personal characteristics, socio-economic indicators, health related habits and health compromising

¹⁹ Since the regime was disassembled completely in 1991, people with such experience were in their middle-age years or older during the time of the EHIS study; also since there is variation in time when Eastern European countries were included into Eastern Bloc, the upper age limit for the group differs between countries, but the lower age limit is the same due to the collapse happening more or less at the same time in all the countries in question, here we assume it to be 1990.

conditions²⁰. The coding of the variables is available in the Appendix and descriptive statistics is presented in Table 1 and discussed in data and descriptive statistics section.

Regressions with self-rated health as a dependent variable were run using different sets of explanatory variables identified above. In the first set explanatory variables included only personal and socio-economic characteristics; in addition the second set included the dummy for communist regime exposure; the third set had the same explanatory variables as the first plus health related habits and health compromising conditions; the fourth – same as third with an inclusion of communism dummy; the fifth – the same as previous set only included interactions of all health conditions with communism dummy to see if that changes the average marginal effects of conditions and the communism dummy. All of the specifications account for the country and the year of interview dummies. Due to the ordered categorical nature of the dependent variable the regression was estimated using ordered probit.

Health Status Indices – comparison estimations

To analyse health from a different angle which is more objective and more comparable, medical research on comorbidity indices helping predict of death, use of medicine, hospital resources etc. was explored. Charlson Comorbidity Index (CCI) is one of the most commonly used, it uses weights for experienced health compromising conditions and also accounts for the patient's age. There is evidence of CCI being a good predictor of the risk of death and it also having a negative effect on the quality of life of an individual (Diederichs et al. 2011, Bastian et al. 2017), and is frequently used to identify less co-morbid patients for surgery or chemotherapy (Phillips et al. 2017). Another frequently used index – Elixhauser Comorbidity Index (ECI) (Elixhauser et al. 1998), which was used with administrative data, it takes into account most of the same conditions as CCI but uses a simple sum of the list of determined morbidity conditions instead of using weights, it also accounts for mental health (Bastian et al. 2017). There is some evidence that weighted scores help identify the risk groups of patients better, an example being elderly people, who tend to have multiple chronic conditions that affect their daily life, use of services and mortality (Tooth et al. 2008). These

²⁰ **Personal characteristics:** gender, age (mid-point of each age cohort), age squared; **socio-economic variables:** education, income (country specific quintiles), dummies for being unemployed, in military or retired; **health related habits:** amount of fruit and vegetables a person eats, frequency of smoking and alcohol consumption, days of moderate physical exercise a week; **health compromising conditions:** dummies for being hospitalized in the last 12 months, being disabled, having asthma, bronchitis or other chronic lung diseases, myocardial infarction, coronary heart disease, hypertension (high blood pressure), stroke, rheumatoid arthritis, osteoarthritis, low back disorder or other chronic back defect, neck disorder, diabetes, allergy (not including allergic asthma), stomach ulcer, cirrhosis of liver, any type of cancer, severe headache, urinary incontinence, chronic anxiety, chronic depression, other mental health problems, permanent injury caused by accident.

indices are subject to criticism as the selection of conditions to be included in calculations is usually determined by the high prevalence of these conditions in general population and are known to be related to increased risk of death, also 79.5% of multi-morbidity data is relying on self-reports, and the process of index creation is arbitrary (Diederichs et al. 2011). It is also suggested that other factors need to be accounted for assessing morbidity too - social networks, support, coping strategies, individual preferences, living conditions (Diederichs et al. 2011).

Since comorbidity indices are mostly based on self-reporting, which means the reconstruction of the indices using self-reported data should provide very similar results, and since they are used to identify patients that could withstand operations, chemotherapy, or other types of treatments, it should be a sufficient measure to control for the health status to look into differences in health perception. Data available from the EHIS does not include some of the health conditions that are included in the original calculations of the indices. However, we do not consider this to be a major issue to the reliability of our calculations because relatively few health conditions are missing and the EHIS dataset includes the most prevalent ones, so the missing information would affect the precision of the calculations of the indices for only a small number of people, nonetheless this implies the calculated indices are subject to underestimation which is consistent throughout the sample. The suggested methodology was followed closely and should allow a level of objective comparability between respondents especially since we use both indices. The Elixhauser index is derived by taking a sum of 31 different health conditions, each condition if present is valued as '1'. The Charlson index uses a weighted approach by assigning a weight ranging from 0 to 6 according to the severity of the condition. It also takes into account person's age which is weighted as the rest of the conditions and finally the results are summed up. Detailed account of the methodology for calculations of both indices and how it was addressed in this study is available in the Appendix.

$$OHI_i = \alpha_0 + \alpha_1 CR_i + \sum_{j=1}^n \alpha_j Controls_i^j + \varepsilon_i \quad (2)$$

A range of regressions with both objective health status indicators (OHI) as dependent variables was used to see if the communist regime could have systematically affected the health of people who experienced it to produce long-term differences that cannot be explained by other characteristics. This also allows us to compare the effects of communism on self-rated health and calculated health status based on experienced health compromising

conditions. There are two dependent variables – Elixhauser and Charlson comorbidity indices, and regressions with them were run respectively. The estimation strategy was similar to the one used for the self-rated health, the only difference being – health conditions were no longer included as separate determinants, this is because they are already incorporated in the calculation of the indices. So the first set of regressions only includes personal and socio-economic characteristics; the second set adds the communism dummy; the third set includes health habits without alcohol consumption (since it is included in the Elixhauser index) to the first set variable list; the last set is the same as the third only also includes the communism dummy. All of the regressions include country and year of survey dummies. Due to the continuous nature of the dependent variables, simple OLS estimations were used.

Health Perception Gap and its determinants

Our calculations of Elixhauser and Charlson Comorbidity Indices were used in the construction of the health perception gap measure (HPG). This was done in two different approaches: using European sample averages or within country averages of self-rated health. Two HPG measures, based either on ECI or CCI, were calculated in each approach (derivation summary table in the Appendix). As an example, we will explain the construction of one of the HPG measures step by step.

Calculation of the HPG based on ECI and European self-rated health averages (HPG1):

1. Means of self-rated health (SRH) for the available European sample were taken by each ECI category. ECI ranges from 0 to 9, resulting in 10 different SRH means. These means vary at an individual level due to own SRH being excluded from the European mean calculations;
2. The difference was taken between the derived European SRH mean and own SRH value, provided the underlying ECI is the same. This produces a HPG – the difference between European self-rated health average and own evaluation of SRH based on the same objective health measure – same ECI value.

HPG2 was calculated using the same principle detailed above only based on CCI as an underlying objective health measure (CCI range is 0-15, resulting in 16 SRH averages). HPGI took country level averages instead of European based on ECI. HPGII followed the same principle as HPGI only is based on CCI. The first two measures (HPG1 and HPG2) represent the difference between own and average European health evaluations for the same comorbidity index value (similar level of objectively rated health); the remaining two – the

difference between own and an average at home country. The first two measures would allow noticing if there are any significant differences between the countries in their health perception and if the communist regime experience in formative years could have had a long-lasting impact in former Eastern Bloc countries. The last two measures would help look into more generalized (normalized to own country) measure of perception. Relationship between the experience of communism and within-country HPG would tell us how much the perceptions of those with such experience differ from those without it within country. In addition, distinguishing between the two types of HPG allows exploring the variation in HPG and testing how relationships with HPG vary based on which average is chosen, helping establish more universal associations with attitudes to health.

These calculations are based on the expectation that country-level sample averages are representative of the ‘true’ population averages if the sample is large enough, so the sample average is expected to converge to the ‘true’ population average, an inference based on the weak law of large numbers (Cameron and Trivedi 2009). So we expect that our separate country sample health perception averages are likely to be representative. However, since we do not have all the European countries in our sample, the ‘true’ European average could be different from our European sample average, especially because we assume country perception averages may be different and determined culturally, but by using the available countries’ pooled average we get as close to the European population average as possible. There is also no claim made that the population averages are representative of some ‘true’ value of what being unbiased in perception towards health is, as this is a completely different question. So our main aim is not to refer to some ‘true’ value of what being unbiased is, but explore the deviations from that average (would it be European or would it be within country) and how much history could have impacted it. This would allow noticing tendencies and some extrapolation of the effects of the perception gap, especially when using the within country measures. This reasoning holds as long as we have representative samples of each country populations, i.e. the most basic statistics show a balanced representation of population²¹.

$$HPG_i = \alpha_0 + \alpha_1 CR_i + \sum_{j=1}^n \alpha_j Controls_i^j + \varepsilon_i \quad (3)$$

²¹ Personal and socio-economic descriptive statistics by each country are available in the Appendix.

The chosen regression specification used for estimating the determinants of health perception gap is shown above (3). The first set of regressions for all 4 different measures of health perception gap was run using only personal and socio-economic characteristics as control variables; the second set adds the communism dummy; the third set includes habits to the first set variables; the last set only differs from the third by including the communism dummy. The regressions with health perception gap based on European mean also took into account the country and year dummies, while those based on country-level means took into account only year dummies. Regressions were run using simple OLS and LASSO estimations since the dependent variables are continuous. LASSO is a machine learning algorithm applied for OLS which penalizes the use of many control variables and helps choose the ones that are important to include in the model. We used the STATA command *'lasso2'* which allows the estimation of the elastic net, the coefficient of which, alpha, and penalty coefficient lambda are determined during the process. If alpha is equal to 1 this means estimation of the 'lasso' regression, which reduces the coefficients of the ineffective controls to 0; alpha equal to 0 means the estimation of the 'ridge' regression, where coefficients are penalized and reduced in size but none of them taken out. In our estimations we use alpha equal to 1 and lambda, the penalty coefficient, equal to 100. This is a relatively high value of lambda used in practice for estimations, resulting in higher penalty loadings for separate coefficients. Other values of lambda, lower and much higher, were tested as well, without significant change in results. This approach technically should allow filtering out the most important regressors if LASSO assumptions hold. These assumptions include: a) sparsity – only a small number of variables can be relevant in the model; b) irrepresentable condition – the variables that are relevant in the model cannot be highly correlated with those that are not. At this point it may not be possible so assess with certainty if these assumptions hold in estimated models. However, the main reason for the use of LASSO was an attempt to reduce the number of regressors. Due to the nature of LASSO estimations standard errors are not reported for each explanatory variable respectively, we report penalty loadings in tables 10-13 instead. Higher penalty loadings identify variables that are at a higher risk of being removed from the estimation process. Literature also suggests using Root mean square error (MSE) as a measure of the overall fit of the model.

Application of Health Perception Gap

We use a range of dependent variables here that we assume could be influenced by health perception gap, and could also have an aggregate effect on country-level expenses.

So the vector of dependent variables include, dummies: receiving (any) benefits in the household, receiving personal or household help, not being hospitalized and not consulting a specialist even if really needing to; continuous measures: rated difficulty to do things alone (based on the average of difficulty ratings from 1, ‘no difficulty’, to 4, ‘I can’t achieve it by myself’ on a list of activities: feeding yourself, getting in and out of a bed or chair, dressing and undressing, using toilets, bathing and showering) and a number of times consulted a GP in last four weeks²². For the dependent dummy variables we use probit model and for the continuous variables – simple OLS, both with vce robust errors.

$$HB_i = \alpha_0 + \alpha_1 HPG_i + \sum_{j=1}^n \alpha_j Controls_i^j + \varepsilon_i \quad (4)$$

Our main variables of interest – the vector of different Health Perception Gap (HPG) measures. Regression specification (4) includes personal and socio-economic characteristics mentioned earlier, having a partner, and country and year dummies as control variables. We also control for some indications of health: indication of being disabled and, depending on which comorbidity index our HPG is based on, either ECI or CCI for an overall health status measure. We expect these regressions to be subject to endogeneity issues, but since they are not the main focus of this research and are only used to assess possible routes of applying the HPG measures, more sophisticated econometric models were not developed. This is also the reason why our sample was not adjusted in accordance to this list of dependent variables, details on sample sizes available in Table 1.

4. Data and descriptive statistics

We used European Health Interview Survey (EHIS) wave 1 microdata from Eurostat. The data was collected in 16 countries in Europe over the period of 2006-2009. Due to some information not being collected in all 16 countries but being relevant for the estimations in this study, the final list of included countries from Western European (by our definition) are: Belgium, Cyprus, Greece, Malta, and from former Eastern Bloc: Bulgaria, Czech Republic, Hungary, Latvia, Poland, Romania, Slovenia, and Slovakia.

Table 1 provides descriptive statistics of the sample used in our estimations. The total number of observations in the final sample used consists of 68,852 observations. General

²² More detail regarding the variables mentioned are available in the Appendix.

characteristics suggest that the sample is quite balanced: 45.3% of the respondents are male, average age is 47.5 years (age range: 16 to over 85), average level of education attained is in between lower secondary and upper secondary, closer to the latter, mean income is between second and third quintile (out of five equal divisions of population based on country specific income range), closer to the third, 5.6% of the sample is unemployed, 0.6% works in the military, and 27.1% are retired. The Appendix offers more information on the categories available and coding of each variable used.

Those that are identified as experienced the communist regime during their formative years amount to 49.7% of the sample. The average self-rated health²³ is between 'Fair' and 'Good'. Around 11.2% of respondents stayed in the hospital over the past 12 months as an inpatient, 3.4% of the sample are permanently disabled²⁴. The list of health compromising conditions indicates having the condition or having previously had it. There is more detailed information available if these conditions being diagnosed by a doctor or not, and if the condition was experienced over the past 12 months, but the response rate for these questions falls dramatically so the former indication of conditions was used. Within the sample the rates of health compromising conditions are: asthma (including allergic asthma) – 4.5%, chronic bronchitis (also includes chronic obstructive pulmonary disease, emphysema) – 5.5%, myocardial infarction – 2.9%, coronary heart disease (angina pectoris) – 7.8%, high blood pressure (hypertension) – 24.6%, stroke (cerebral haemorrhage, cerebral thrombosis) – 2.1%, rheumatoid arthritis – 10%, osteoarthritis – 12.3%, low back problems (includes other back defects) – 22.4%, neck problems – 12.3%, diabetes – 6%, allergies – 10%, stomach ulcer – 6.6%, cirrhosis of the liver, liver dysfunction – 1.9%, cancer (includes leukaemia, lymphoma and malignant tumour) – 1.8%, severe headache like migraine – 9.3%, urinary incontinence – 3.8%, chronic anxiety – 3.6%, chronic depression – 3.6%, other mental health problems – 1.4%, permanent injury cause by accident – 4.9%, bad eyesight (having to wear glasses) – 50%, bad hearing (wearing hearing aid) – 3.2. The average BMI is 25.7 which is just over normal healthy level of 18.5 – 25 which means on average sample is mildly overweight; 17.7% of the sample are obese (with BMI over 30); 0.04% of the sample are severely underweight (BMI under 15). On average, respondents consume alcohol between 'Monthly or less' and '2-4 times a month', with 4.1% of the sample drinking alcohol daily; they smoke between 'Never' and 'Occasionally'; engage in mild exercise around 3.3

²³ The categorical EHIS variable based on a request to rate own health over the past 12 months was recoded in reverse, now it ranges from 'very bad health' (1) to 'very good health' (5).

²⁴ Derived from one of the answer categories to a question about the present labour status.

days a week; eat fruit and vegetables between ‘*at least 4 times a week*’ and ‘*once a day*’; 4.5% experienced chronic anxiety or mental health problems, which is identified here as ‘*psychoses*’. Also on average people in the sample experience between no pain to mild pain, closer to the latter.

The mean of Elixhauser comorbidity index (ECI) is 0.95, this index ranges between 0 and 9; Charlson comorbidity index (CCI) mean – 1.32, range: 0-15, in both higher values indicate worse health. The CCI mean is higher than for ECI because of the condition-weighted approach used in the former. This indicates a sample of people with on average 1 health compromising condition. Even if statistics of the prevalence of each health compromising condition in the countries in our sample is not readily available, the EHIS survey design ensures the data is collected from a representative sample of the population in each country.

Table 2 provides a summary of self-rated health responses, divided into two groups – those that are identified as having been exposed to the communist regime during formative years and those that were not of the equivalent age group (40-84). Higher proportions of those who experienced communism rate their health as ‘*very bad*’, ‘*bad*’ or ‘*fair*’ in comparison to those who were not exposed and are of the same age range. The difference for the ‘*very bad*’ health is relatively small, but twice more people with communism experience rate their health as ‘*bad*’ than those without (15.47% and 7.25% respectively). ‘*Fair*’ health is reported by 38.75% and 25.23% respectively, which is around 13% more amongst those exposed to the communist regime. For the last two categories (‘*good*’ and ‘*very good*’ health) proportions are higher for those without communist regime experience: ‘*good*’ health – 42.62% (no communism) and 35.1% (communism), and ‘*very good*’ health – 22.77% and 7.04% respectively. This begs a question if this tilt towards more negative reporting among those who were exposed to communist regime rule can be explained by some characteristics that are not represented in this table, such as: age, socio-economic conditions, or health related habits, or it represents a difference in perception of own health more fundamentally.

Table 3 shows the decomposition of the Elixhauser Comorbidity Index (ECI) by its each value for those affected by the communist and those that were not who are of the same age group. Table 4 shows the same decomposition for the Charlson Comorbidity Index (CCI). There is some variation between the two groups for ECI, but the differences do not seem to be very large. The most noticeable difference is for the values 4-6, where higher proportions of people with the communism experience in formative years fall into (4: 5.11% and 3.92%;

5: 2.16% and 1.83%; 6: 0.9% and 0.56% respectively). The differences between the two groups are more evident with CCI index values, but they do not seem very large or consistent. A proportion of respondents with the index value being 0 – 24.18% in the not affected group and 21.56% in the affected. Higher proportions of individuals with the communism experience have index values of 1-3 or 8 and above, the proportions are larger for the not affected group for the values 4-7. However, the differences are small and seem to follow a similar trend overall. These proportions seem to suggest a similar idea as the ones from ECI – the pattern is less evident than it was in the self-reported health decomposition, and it is not easy to determine if those with or those without the communism experience during formative years are objectively at worse health.

The sample's mean of the calculated health perception gap based on ECI is 0.19, the range: -3.17 to 2.94, the mean for one based on CCI is 0.003, range: -3.2 to 2.75. Negative values of this measure indicate a more negative evaluation of own health than the full sample average relative to the same value of either ECI or CCI, which should indicate a similar objective health status. The perception gap using both calculations is more negative for those people who experienced the communist regime (-0.22 and -0.18) than those of the same age range that did not (0.56 and 0.32). The within country calculations for the perception gap are: -0.005 and -0.0004 based on ECI and CCI indices respectively, this shows health perception differs across countries, but when accounted for the country average the perception gap is much closer to 0, which should mean on average the same own health evaluation as others with similar objective health status. This alone suggests health perception could be defined in each culture.

5. Results

5.1 Determinants of Self-rated Health and influence of Communist regime

Tables 5-7 show the average marginal effects from ordered probit regressions with self-rated health as dependent variable. Results are shown for three categories respectively: 'very bad' health (1), 'fair' health (3), and 'very good' health (5). For each of those categories there are 5 sets of results from different specifications, which are discussed in detail in methodology section. Regressions without including separate health compromising conditions explain over 18% of variation in the dependent variable, as according to Pseudo-R², the fit increases by almost 10% when health conditions are included, which is reasonable

as these conditions should have a large effect on how one evaluates own health. Inclusion of communism dummy contributes to a small improvement in fit.

The average marginal effects for the three categories of self-rated health show that the experience of communism during formative years is consistently highly significant at 1% level. The marginal effects reveal very similar trends as shown in Table 2: communism dummy is an important predictor for a higher probability of rating own health as '*fair*', and lower probability of rating it as '*very good*', it also slightly increases the probability of the rating being '*very bad*'. In the regressions without health compromising conditions communism dummy has the most sizeable effect amongst the personal and socio-economic characteristics dummies. Once controlled for health conditions, the size of the effect is around twice smaller: it produces a 0.4% increase in probability of '*very bad*' rating, 1.4% increase of probability of '*fair*', and 2.3% decrease of probability of '*very good*'.

In the regressions without controlling for health conditions, being unemployed or retired show a strong predictive association with rating health more negatively. This could suggest employment status being important in own health evaluation. Aging also is related to seeing own health as worse. Being male, having higher education attainments, higher income and being in military are associated with increasing the probability of health being rated better. Which is mostly in line with previous findings.

Health compromising conditions have an expected considerable negative effect on own health rating. The largest negative effect is determined by being disabled, followed by having a mental health condition, and then followed by a stroke, cancer and diabetes. Having gone through formative years during communism regime is consistently estimated to have a larger negative impact than neck pain, the effect is comparable to that of bronchitis, having an allergy, having bad eyesight or having a permanent injury. This makes it a sizeable effect, maybe having been exposed to the communist regime could be humorously considered a health compromising condition too.

5.2 Relationship between the Communist Regime and Objective Health Indicators

In tables 8 and 9 the OLS results from the regressions with ECI and CCI as dependent variables are presented, the same regression specifications identified in the methodology section were used for both health indices. The impact of the communist regime is not consistent between the two indices. The communism dummy is highly statistically significant and has a positive sign (increasing the index value indicating worse health) in all specifications, but the size of the effect differs largely between the indices. Regressions with

ECI as dependent variable, show communism dummy having a sizeable effect, especially in comparison to other control variables. Having experienced the communist regime in formative years increases ECI by 0.18. The effect is much smaller for the CCI – 0.026 and many other characteristics that are controlled for have more sizeable coefficients. This could be partially explained by ECI including alcohol abuse as one of the determining factors for the index, which is not included in CCI calculation. In the second chapter of this thesis the communist regime was shown to have a significant impact on increasing the frequency of alcohol consumption and that of binge drinking. This finding could help explain why the effect of the communism dummy is more sizeable on the ECI. In addition, in the CCI regressions over 75% of variation in CCI is explained by the regression specification, while around 31% of variation in ECI is explained by the model, suggesting the CCI estimation producing a better fit.

There are other interesting effects to consider. Being retired significantly adds to the increase of both indices, this effect is larger than that of communism for both indices. Age is another highly statistically significant characteristic but produces mixed results, for the ECI effect of age varies between 0.019 and 0.033 taken into account the lowest age group is 16 this becomes: 0.304-0.528 and much higher for the higher age categories; for CCI the effect is of opposite sign between -0.059 and -0.062 (-0.944 and -0.992 for a 16 year old), and it is puzzling how such difference in results could be explained. Higher levels of education and income have a decreasing effect for both indices and are highly statistically significant. Being unemployed is consistently significant at 5% level in all of the specifications and is negative, suggesting better objective health. This is an opposite finding from the results with self-rated health as dependent variable. This could mean that the negative effect of unemployment towards health evaluation could be more psychological. Finally higher frequency of moderate physical activity has a small but statistically significant beneficial effect on objective health as measured per both indices.

5.3 Health Perception Gap: Determinants and Importance

Health perception gap was derived using ECI and CCI values as a reference. Methodology of the derivation of the HPG is explained in the methodology section, the derivation rules of both indices are provided in the Appendix. This perception gap represents the difference between own health evaluation and the average of the European sample (excluding own value) of the same health index value (measured either by ECI or CCI). To capture the perception gap in more general terms, the perception gap was also calculated

using within country averages. Negative values of this measure indicate more negative evaluation of own health than the average of the reference group. Descriptive statistics show that on average those who have experienced the communist regime during their formative years have a negative health perception gap, while those of the same age that did not have this experience have a positive HPG.

Tables 10 and 11 present the results of OLS and LASSO regressions with health perception gap as dependent variable based either on ECI or CCI and European self-reported health averages. Regardless of the expression of HPG used and the estimating method, the communism dummy was found to be a highly statistically significant predictor with a consistently sizeable negative effect (indicating more negative health perception). For the HPG based on ECI the effect is between -0.107 and -0.127, and for the HPG based on CCI between -0.139 and -0.152 from the OLS regressions. The LASSO estimations show this effect to be larger: between -0.125 and -0.128 for the measure based on ECI and between -0.157 and -0.165 for CCI-based measure. These are large effects knowing both versions of the HPG vary between around -3.2 and 2.9/2.7.

The factors that seem to be statistically significantly associated with more positive HPG are: being male, being more educated, having higher income, eating more fruits, vegetables and engaging in moderate physical activity more frequently. The factors that tend to have a negative effect are: aging and being unemployed. The negative effect of communism is larger than that of any other dummy variable. These results are consistent with those reported in self-rated health equations.

In tables 12 and 13 the with-in country HPG results are presented. This method of measuring HPG compares own health evaluation with that of the average of own country, this allows eliminating any country level cultural effects from the HPG measure. This should mean, if health perception is mainly explained by cultural phenomena, then those people who experienced communist regime during formative years will not be very different from the rest of the population in the country, after controlling for personal and socio-economic characteristics. The effect of the communist regime in this case is considerably smaller, but still mostly statistically significant and negative with both dependent variables. For HPG based on ECI it is between -0.019 and -0.036, and for one based on CCI: -0.017/-0.018 from OLS estimations. LASSO results once again provide somewhat larger coefficients: between -0.03 and -0.042, and between -0.027 and -0.04 respectively. This shows a couple of things: firstly, own health perception is a cultural phenomenon, since the size of the communism dummy effect drops significantly; secondly, even after the adjustment for with-in country

averages, communism experience still maintains a negative effect towards health perception, which means a difference exists between those people who have experienced it during formative years and those that have not even if they are from the same country. These findings suggest that the experience of the communist regime was psychologically traumatic with a long-term effect on perception.

The rest of the control variables maintain very similar coefficients as in tables 10 and 11, the only noticeable difference being – the coefficient of unemployment dummy increasing in size and that of income decreasing. The explanatory power of the with-in country HPG regressions fall in comparison to the HPG based on European averages, which is to be expected knowing part of the variation was removed. LASSO results are very similar to those of OLS overall. In some estimations age squared measure is removed by the LASSO algorithm, but other explanatory variables are kept in and maintain similar coefficients. The only coefficients that change in size even if not substantially are those of the communism dummy, it increases in all of the estimations in comparison to OLS. Root mean square error (henceforth Root MSE) for both estimation methods is very similar, suggesting there is not much of the systematic difference between the two methods for our sample and both provide a similar goodness of fit. There is, however, a small decrease in Root MSE in the estimations that include the communism dummy, which suggests it is an important explanatory variable to be included. LASSO results also indicate that LASSO assumptions might not hold for these estimations since, with an exception of one, none of the regressors were removed from the estimation process. The most likely explanation is HPG is possibly affected by a long list of factors, violating the sparsity condition, but it is also possible that there is no strong disassociation between relevant and irrelevant factors. However, the main variable of interest, experience of communism, as an exogenous indicator is unlikely to be highly correlated with personal and socio-economic indicators in the estimations, and it receiving quite low penalty loadings in comparison and not being removed from the estimations indicates its importance in determining HPG outcomes.

The results with HPG used as an explanatory variable are presented in Table 14. The list of dependent variables include dummy variables for:

1. Claiming any kind of benefits in the household;
2. Respondent receiving help for personal and household activities;
3. Refusing to be hospitalized after a recommendation from a doctor during the past 12 months;

4. Not consulting a specialist even if being in real need at least once over the past 12 months.

Also continuous variables for:

1. Reported average difficulty level for doing a list of activities alone (feeding yourself, getting in and out of bed, dressing and undressing, using toilets, bathing and showering);

2. Number of times consulting a GP or a family doctor on own behalf during the past four weeks.

Regressions were run with each of the health perception gap measures respectively. We controlled for personal, socio-economic, health characteristics²⁵ and country and year dummies. All of the HPG measures are highly statistically significant in all of the regressions run. The results suggest that a more positive evaluation of HPG is associated with a lower probability of benefits being claimed within the household, lower chance of respondent receiving household help, and lower risk of not going to a hospital or not seeing a specialist when it is necessary. Higher HPG measures are related to lower reported levels of experienced difficulty with daily tasks and less frequent GP consultations in the past four weeks. It could be generalized that more positive HPG seems to be related with practices that help maintain health (like going to a hospital or seeing a specialist when necessary), more optimistic assessment of experienced difficulty level, which could explain lower requirement of household help and fewer GP visits, and a lower probability of respondent's household claiming benefits. This could suggest that more optimistic assessment of own health promotes healthy practices and is likely to be beneficial for own pocket and incurred healthcare costs frequently at least partially borne by the country. These relationships were not analysed in detail, since it is not the emphasis of this study, but this shows the possible ways of application of the derived HPG measure, which could be explored in future research.

6. Discussion and conclusion

Findings of this paper suggest that the communist regimes in Eastern Bloc had a long-term effect on the perception of health of people who experienced them during their formative years. This resulted in a tendency to assess self-rated health more negatively. The results from regressions indicate the communist regime experience during formative years had a significant negative effect on the derived health perception gap measure regardless of

²⁵ Controls include: dummy for being disabled, ECI or CCI (based on the HPG measure used).

the reference point: European or own country average. The results also show the communist regimes had a negative long-run effect on objective health too, but this effect is much smaller comparatively. This allows to speculate that the experience of the communist regimes in former Eastern Bloc had a traumatic effect and some of its psychological consequences could be felt up to the present date. Since we also observe country level differences in health perception gap it is possible part of such experience became embedded in culture and could be transmitted between generations.

The effect on health perception gap not only is significant and sizeable, but also we find significant associations between this constructed measure and behaviour relevant to health which could lead to aggregated country level differences in medical costs. So far evidence suggests that more optimistic perception of own health is related to healthier actions like not avoiding going to a hospital or visiting a specialist if it is necessary and potentially lower expenses on healthcare because of lower reported experienced difficulty levels, need for home help, fewer GP visits and a lower probability of claiming benefits. This provides an insight into the additional costs the communist regimes might have inflicted on the future of the countries that experienced them. This also could contribute to the explanation of the mechanism of how the experience of the many could shape cultural norms and internal biases. All of this suggests that 40-60 year long experiment of the communism had a long-term effect on social capital outcomes in former Eastern Bloc, consequences of which could be felt for some time to come.

The communist regimes could have had an effect on perception on own health of those people who have experienced this through these channels:

1. They lived in a system that was based on propaganda, information hiding and corruption. This created a natural sense of mistrust, which could affect life satisfaction and perception of own circumstances, which could have resulted in a generally more negative outlook which then could translate to more negative perception of health as well.
2. There is evidence of a positive relationship between work morale and self-rated health (Mossey and Shapiro 1982). Expecting the relationship between the two to be endogenous, changes in work morale could also impact evaluation of own health. The egalitarian ideology of the communist regimes was translated into compressed wage grid which did not offer much of an economic incentive to keep the work morale high.

This paper addresses a very broad multi-disciplinary question and faces multiple limitations. Firstly, there is a possibility that the health indices do not fully represent the objective health of respondents. It is possible some of those whose objective health indices are equal to 0 could be in denial, refrain from having regular health checks to be diagnosed. On the other hand, the reported conditions could be self-diagnosed, though the same concerns would apply to the whole sample and there is no reason to expect that there would be systematic differences between those with the communism experience and those without. Use of these health indices in medical research also contributes to the credibility of our method. Secondly, it is difficult to define precisely the meaning of the values of both health indices and health perception gap measure, we can only capture the trend of the comparative relationship and show its relative importance. Thirdly, we cannot control for the nationality of the people in the sample, we just have to assume only a small fraction in each country's sample would be foreign; knowing that a lot of eastern Europeans migrated to other western countries while the populations of foreigners in former Eastern Bloc remain small only suggests any found effect of the communist regime would be underestimated.

References:

- Alesina, A. and Fuchs-Schündeln, N. (2007). Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences. *American Economic Review*, 97(4), 1507-1528.
- Ananyev, M. and Guriev, S. (2016). Effect of Income on Trust: Evidence from the 2009 Economic Crisis in Russia. Available at SSRN: <https://ssrn.com/abstract=2542001> or <http://dx.doi.org/10.2139/ssrn.2542001>
- Baker, M., Stabile, M. and Deri, C. (2004). What Do Self-Reported, Objective, Measures of Health Measure?. *Journal of Human Resources*, 39(4), 1067-1093.
- Bastian, L.A., Brandt, C.A., Justice, A.C. (2017). Measuring Multimorbidity: A Risky Business. *Journal of General Internal Medicine*, 32(9), 959-960.
- Barro, R. J. and McCleary, R. M. (2003). "Religion and Economic Growth Across Countries". *American Sociological Review* 68: 760-781.
- Becker, S. O. and Woessmann, L. (2009). "Was Weber Wrong? A Human Capital Theory of Protestant Economic History." *Quarterly Journal of Economics*: 531-596.
- Benítez-Silva, H., et al. (2004). How Large is the Bias in Self-Reported Disability?. *Journal of Applied Econometrics*, 19(6), 649-670.
- Bisin, A. and Verdier, T. (2000). "Beyond the Melting Point: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits." *Quarterly Journal of Economics* 115: 955-988.
- Blum, U. and Dudley, L. (2001). "Religion and Economic Growth: Was Weber Right?" *Journal of Evolutionary Economics* 11: 207-230.
- Bourne, P.A. (2009). Self-Rated Health and Health Conditions of Married and Unmarried Men in Jamaica. *North American Journal of Medical Sciences*, 1(7), 345-352.
- Brown, A.F., Ang, A. and Pebley, A.R. (2007). The Relationship Between Neighbourhood Characteristics and Self-Rated Health for Adults With Chronic Conditions. *American Journal of Public Health*, 97(5), 926-932.
- Cai, L. (2010). The relationship between health and labour force participation: Evidence from a panel data simultaneous equation model. *Labour Economics*, 17(1), 77-90.
- Cameron, A. C. and Trivedi, P. K. (2009). Asymptotic Theory: A.3. Laws of Large Numbers. In: *Microeconometrics: Methods and Applications*. New York: Cambridge University Press, pp. 947-948.
- Coval, J. D. and Thakor, A. V. (2005). Financial intermediation as a beliefs-bridge between optimists and pessimists. *Journal of Financial Economics*, 75(3), 535-569.
- Desivilya, H.S., Gal, R. and Ayalon, O. (1996). Long-term Effects of Trauma in Adolescence: Comparison between Survivors of a Terrorist Attack and Control Counterparts. *Anxiety, Stress & Coping*, 9(2), 135-150.

- Diederichs, C., Berger, K., Bartels, D.B. (2011). The Measurement of Multiple Chronic Diseases – A Systematic Review on Existing Multimorbidity Indices. *Journal of Gerontology: Medical Sciences*, 66A(3), 301-311.
- Elixhauser, A., Steiner, C., Harris, D.R., Coffey, R.M. (1998). Comorbidity measures for use with administrative data. *Medical Care*, 36(1), 8-27
- García-Gómez, P. (2011). Institutions, health shocks and labour market outcomes across Europe. *Journal of Health Economics*, 30(1), 200-213.
- García-Gómez, P., et al. (2013). Long- Term and Spillover Effects of Health Shocks on Employment and Income. *Journal of Human Resources*, 48(4), 873-909.
- Gorodnichenko, Y. and Roland, G. (2011). “What Dimensions of Culture Matter for Long-Run Growth?” *American Economic Review. Papers and Proceedings* 101(3): 492-498.
- Grier, R. (1997). “The Effect of Religion on Economic Development: A Cross National Study of 63 Former Colonies.” *KYKLOS* 50: 47-62.
- Grosjean, P. (2011). “The Weight of History on European Cultural Integration: A Gravity Approach”. *American Economic Review. Papers and Proceedings* 101(3): 504-508.
- Heaton, J. B. (2002). Managerial Optimism and Corporate Finance. *Financial Management (Wiley-Blackwell)*, 31(2), 33.
- Jeon, S. and Pohl, R. V. (2017). Health and work in the family: Evidence from spouses' cancer diagnoses. *Journal of Health Economics*, 52, 1-18.
- Jiang, F., et al. (2016). How Belief in a Just World Benefits Mental Health: The Effects of Optimism and Gratitude. *Social Indicators Research*, 126(1), 411-423.
- Khallad, Y. (2013). Dispositional optimism and physical wellbeing: The relevance of culture, gender, and socioeconomic status. *International Journal of Psychology*, 48(5), 978-985.
- Krosnick, J. A. and Alwin, D. F. (1989). Aging and susceptibility of attitude change. *Journal of Personality and Social Psychology*, 57, 416-425.
- Lindström, M. (2009). Marital Status, Social Capital, Material Conditions and Self-Rated Health: A Population-Based Study. *Health Policy*, 93, 172-179.
- McDaid, O., Hanly, M.J., Richardson, K., Kee, F., Kenny, R.A. and Savva, G.M. (2013). The Effect of Multiple Chronic Conditions on Self-Rated Health, Disability and Quality of Life Among the Older Populations of Northern Ireland and the Republic of Ireland: a Comparison of Two Nationally Representative Cross-Sectional Surveys. *BMJ Open*, 3(6), 1-9.
- Mossey, J.M. and Shapiro, E. (1982). Self-Rated Health: A Predictor of Mortality Among the Elderly. *American Journal of Public Health*, 72(8), 800-808.
- Münich, D., Svejnar, J. and Terrell, K. (2005). Returns to Human Capital under the Communist Wage Grid and During the Transition to a Market Economy. *Review of Economics and Statistics*, 87(1), 100-123.

- Phillips, A., Singh, K., Pounds, R., Sundar, S., Kehoe, S., Nevin, J., Elattar, A. and Balega, J. (2017). Predictive Value of the Age-Adjusted Charlson Comorbidity Index on Peri-Operative Complications, Adjuvant Chemotherapy Usage and Survival in Patients Undergoing Debulking Surgery after Neoadjuvant Chemotherapy for Advanced Epithelial Ovarian Cancer. *Journal of Obstetrics and Gynaecology*, 37(8), 1070-1075.
- Pit, S. W., et al. (2010). Health problems and retirement due to ill-health among Australian retirees aged 45–64 years. *Health Policy*, 94(2), 175-181.
- Puri, M. and Robinson, D. T. (2007). Optimism and economic choice. *Journal of Financial Economics*, 86(1), 71-99.
- Qin, L. and Wang, Z. (2015). Impact of Health Status and Related Factors on Labor Market Positions of Urban Mature Chinese. *Journal of Labor Research*, 36(2), 224-231.
- Roland, G. (2010). The Long-Run Weight of Communism or the Weight of Long-Run History? Wider Working Papers [Online], 2010/83.
- Roland, G. (2002). The Political Economy of Transition. *Journal of Economic Perspectives*, 16(1), 29-50.
- Roland, G. and Verdier, T. (2003). Law enforcement and transition. *European Economic Review*, 47(4), 669-685.
- Scheier, M. F., Carver, C. S. and Bridges, M. W. (1994). Distinguishing Optimism From Neuroticism (and Trait Anxiety, Self-Mastery, and Self-Esteem): A Re-evaluation of the Life Orientation Test. *Journal of Personality & Social Psychology*, 67(6), 1063-1078.
- Sen, R. and Tumarkin, R. (2015). Stocking up: Executive optimism, option exercise, and share retention. *Journal of Financial Economics*, 118(2), 399-430.
- Solberg Nes, L., Evans, D. R. and Segerstrom, S. C. (2009). Optimism and College Retention: Mediation by Motivation, Performance, and Adjustment. *Journal of Applied Social Psychology*, 39(8), 1887-1912.
- Tabellini, G. (2008). "Institutions and Culture. Presidential Address European Economic Association." *Journal of the European Economic Association* 6: 255-294.
- Tooth, L., Hockey, R., Byles, J., Dobson, A. (2008). Weighted Multimorbidity Indexes Predicted Mortality, Health Service Use, and Health-Related Quality of Life in Older Women. *Journal of Clinical Epidemiology*, 61, 151-159.
- Trevisan, E. and Zantomio, F. (2016). The impact of acute health shocks on the labour supply of older workers: Evidence from sixteen European countries. *Labour Economics*, 43, 171-185.
- Vaez, M. and Laflamme, L. (2002). First-Year University Students' Health Status and Socio-Demographic Determinants of Their Self-Rated Health. *Work*, 19(1), 71-80.
- Wang, J., et al. (2017). Optimism bias, portfolio delegation, and economic welfare. *Economics Letters*, 150, 111-113.
- Zhao, X., et al. (2015). Dispositional optimism, self-framing and medical decision-making. *International Journal of Psychology*, 50(2), 121-127.

Tables

Table 1: Descriptive statistics

Variable	Mean	S. D.	Min/Max
Gender (D)	0.453	0.498	0/1
Age	47.59	18.75	16/87
Age squared	2616.53	1853.46	256/7569
Education	3.881	1.277	1/7
Income	2.807	1.354	1/5
Unemployed (D)	0.056	0.230	0/1
Work in military (D)	0.006	0.077	0/1
Retired (D)	0.271	0.445	0/1
Communism (D)	0.497	0.5	0/1
Self-rated Health	3.697	0.99	1/5
Elixhauser Comorbidity Index (ECI)	0.947	1.282	0/9
Charlson Comorbidity Index (CCI)	1.316	1.716	0/15
HPG1 (based on ECI) (C)	0.191	0.804	-3.17/2.941
HPG1 (if Communism=1)	-0.216	0.793	-3.17/2.941
HPG1 (if Communism=0 & age 40-84)	0.255	0.836	-3.17/2.525
HPG2 (based on CCI) (C)	0.003	0.793	-3.204/2.747
HPG2 (if Communism=1)	-0.176	0.8	-3.204/2.609
HPG2 (if Communism=0 & age 40-84)	0.322	0.829	-3.204/2.747
HPGI (in country, based on ECI) (C)	-0.005	0.779	-3.607/3.029
HPGII (within country, based on CCI) (C)	-0.0004	0.76	-3.673/3.115
Hospitalized (D)	0.112	0.316	0/1
Disabled (D)	0.034	0.182	0/1
Asthma (D)	0.045	0.207	0/1
Bronchitis (D)	0.055	0.228	0/1
Infarction (D)	0.029	0.169	0/1
Coronary heart disease (D)	0.078	0.268	0/1
High blood pressure (D)	0.246	0.431	0/1
Stroke (D)	0.021	0.143	0/1
Arthritis (D)	0.1	0.3	0/1
Osteoarthritis (D)	0.123	0.329	0/1

Low back problems (D)	0.224	0.417	0/1
Neck problems (D)	0.123	0.328	0/1
Diabetes (D)	0.06	0.238	0/1
Allergies (D)	0.1	0.301	0/1
Stomach problems (D)	0.066	0.248	0/1
Liver problems (D)	0.019	0.137	0/1
Cancer (D)	0.018	0.134	0/1
Migraine (D)	0.093	0.291	0/1
Urinary problems (D)	0.038	0.191	0/1
Anxiety (D)	0.036	0.186	0/1
Depression (D)	0.036	0.187	0/1
Mental health problems (D)	0.014	0.118	0/1
Permanent injury (D)	0.049	0.216	0/1
Bad eyesight (D)	0.5	0.5	0/1
Bad hearing (D)	0.032	0.175	0/1
BMI (C)	25.736	6.343	2.469/1168
Obese (D)	0.177	0.382	0/1
Severely Underweight (D)	0.0004	0.021	0/1
Alcohol Abuse (D)	0.041	0.199	0/1
Psychoses (D)	0.045	0.208	0/1
Physical Pain	1.902	1.006	1/5
Alcohol frequency	2.347	1.264	1/6
Smoking	1.521	0.843	1/3
Physical exercise	3.306	2.766	0/7
Frequency of Fruit	4.505	1.222	1/6
Frequency of Vegetables	4.56	1.047	1/6
Household benefits (D) (62,755 obs.)	0.693	0.461	0/1
Household help (D) (12,925 obs.)	0.736	0.441	0/1
No hospital (D) (68,472 obs.)	0.031	0.175	0/1
No specialist (D) (68,187 obs.)	0.108	0.311	0/1
Difficulty level (68,796 obs.)	1.07	0.281	1/4
GP visits (38,988 obs.)	0.703	0.99	0/24
Belgium (D)	0.066	0.249	0/1
Bulgaria (D)	0.066	0.248	0/1

Cyprus (D)	0.093	0.29	0/1
Czech Republic (D)	0.02	0.141	0/1
Greece (D)	0.00007	0.009	0/1
Hungary (D)	0.065	0.247	0/1
Latvia (D)	0.09	0.286	0/1
Malta (D)	0.031	0.172	0/1
Poland (D)	0.331	0.471	0/1
Romania (D)	0.155	0.362	0/1
Slovenia (D)	0.022	0.148	0/1
Slovakia (D)	0.06	0.238	0/1
<hr/>			
Year 2006 (D)	0.0001	0.01	0/1
Year 2007 (D)	0.021	0.142	0/1
Year 2008 (D)	0.422	0.494	0/1
Year 2009 (D)	0.557	0.497	0/1

Note: Sample size - 68,852; S.D. - Standard deviation; D – denotes a dummy variable, C – a continuous variable.

Table 2: Self-rated health by Communism (for the age group: 40-84)

Self-rated Health	Communism = 0	Communism = 1
1 – Very bad health	173 (2.13%)	1,244 (3.64%)
2 – Bad	588 (7.25%)	5,292 (15.47%)
3 – Fair	2,045 (25.23%)	13,257 (38.75%)
4 – Good	3,455 (42.62%)	12,009 (35.1%)
5 – Very good health	1,846 (22.77%)	2,407 (7.04%)
Total	8,107	34,209

Note: The chosen age group represents the main age group for the respondents who experienced the communist regime in the Eastern Bloc in their formative years, individuals of same age group were chosen for comparison.

Table 3: Elixhauser Comorbidity Index by Communism (for the age group: 40-84)

Elixhauser Comorbidity Index	Communism = 0	Communism = 1
0	2,834 (34.96%)	11,836 (34.6%)
1	2,460 (30.34%)	9,647 (28.2%)
2	1,502 (18.53%)	6,269 (18.33%)
3	777 (9.58%)	3,535 (10.33%)
4	318 (3.92%)	1,747 (5.11%)
5	148 (1.83%)	739 (2.16%)
6	45 (0.56%)	307 (0.9%)
7	18 (0.22%)	94 (0.27%)
8	5 (0.06%)	29 (0.08%)
9	-	6 (0.02%)
Total	8,107 (100%)	34,209 (100%)

Note: Higher values of the index represent worse health status. Derivation of the index explained in detail in the methodology section, calculation tables available in the Appendix. The chosen age group represents the main age group for the respondents who experienced the communist regime in the Eastern Bloc in their formative years, individuals of same age group were chosen for comparison.

Table 4: Charlson Comorbidity Index by Communism (for the age group: 40-84)

Charlson Comorbidity Index	Communism = 0	Communism = 1
0	1,960 (24.18%)	7,375 (21.56%)
1	1,858 (22.92%)	8,602 (25.15%)
2	1,544 (19.05%)	6,879 (20.11%)
3	1,144 (14.11)	5,664 (16.56%)
4	894 (11.03%)	3,096 (9.05%)
5	421 (5.19%)	1,455 (4.25%)
6	178 (2.2%)	588 (1.72%)
7	61 (0.75%)	222 (0.65%)
8	24 (0.3%)	136 (0.4%)
9	8 (0.1%)	79 (0.23%)
10	9 (0.11%)	51 (0.15%)
11	4 (0.05%)	30 (0.09%)
12	1 (0.01%)	18 (0.05%)
13	1 (0.01%)	9 (0.03%)
14	-	1 (0.00%)
15	-	4 (0.01%)
Total	8,107 (100%)	34,209 (100%)

Note: Higher values of the index represent worse health status. Derivation of the index explained in detail in the methodology section, calculation tables available in the Appendix. The chosen age group represents the main age group for the respondents who experienced the communist regime in the Eastern Bloc in their formative years, individuals of same age group were chosen for comparison.

Table 5: Determinants of Self-rated health, average marginal effects for the outcome – ‘Very bad health’ (1) from ordered probit estimation

Variable	1) SES	2) SES + Com	3) SES+habits +conditions	4) SES+habits +conditions+Com
Communism.D		0.009*** (0.001)		0.004*** (0.001)
Male.D	-0.006*** (0.0004)	-0.006*** (0.0003)	-0.002*** (0.0003)	-0.002*** (0.0003)
Age	0.003*** (0.000)	0.002*** (0.0001)	0.002*** (0.0001)	0.002*** (0.0001)
Age-squared	-0.00001*** (0.000)	-9.2e-06*** (0.000)	-0.00001*** (0.000)	-9.14e-06*** (0.000)
Education	-0.006*** (0.0002)	-0.006*** (0.0002)	-0.004*** (0.0002)	-0.004*** (0.0002)
Income	-0.005*** (0.0002)	-0.005*** (0.0002)	-0.002*** (0.0001)	-0.002*** (0.0001)
Unemployed.D	0.005*** (0.001)	0.005*** (0.001)	0.007*** (0.001)	0.007*** (0.001)
Military.D	-0.006*** (0.002)	-0.007*** (0.002)	-0.007*** (0.002)	-0.007*** (0.002)
Retired.D	0.005*** (0.001)	0.005*** (0.001)	0.006*** (0.001)	0.006*** (0.001)
Fruits			-0.002*** (0.0002)	-0.002*** (0.0002)
Vegetables			-0.001*** (0.0002)	-0.001*** (0.0002)
Smoke			0.001*** (0.0002)	0.001*** (0.0002)
Alcohol			-0.002*** (0.0001)	-0.002*** (0.0001)
Phys. activity			-0.0005*** (0.0001)	-0.0005*** (0.0001)
Hospitalized.D			0.02*** (0.001)	0.02*** (0.001)
Disabled.D			0.05*** (0.002)	0.05*** (0.002)
Asthma.D			0.014*** (0.001)	0.014*** (0.001)
Bronchitis.D			0.005*** (0.001)	0.005*** (0.001)
Infarction.D			0.010*** (0.001)	0.01*** (0.001)
Heart disease.D			0.009*** (0.001)	0.009*** (0.001)
High blood pres.D			0.012*** (0.0004)	0.012*** (0.0004)
Stroke.D			0.023*** (0.002)	0.023*** (0.002)
Arthritis.D			0.009*** (0.001)	0.009*** (0.001)

Arthrosis.D			0.008*** (0.001)	0.008*** (0.001)
Low back pain.D			0.012*** (0.0005)	0.012*** (0.0005)
Neck pain.D			0.003*** (0.001)	0.003*** (0.001)
Diabetes.D			0.019*** (0.001)	0.019*** (0.001)
Allergy.D			0.005*** (0.001)	0.005*** (0.001)
Stomach.D			0.011*** (0.001)	0.011*** (0.001)
Liver.D			0.017*** (0.002)	0.016*** (0.002)
Cancer.D			0.022*** (0.002)	0.022*** (0.002)
Migraine.D			0.009*** (0.001)	0.009*** (0.001)
Bladder.D			0.001 (0.001)	0.001 (0.001)
Anxiety.D			0.007*** (0.001)	0.007*** (0.001)
Depression.D			0.013*** (0.001)	0.013*** (0.001)
Mental health.D			0.026*** (0.003)	0.026*** (0.003)
Injury.D			0.006*** (0.001)	0.006*** (0.001)
Bad eyes.D			0.004*** (0.0003)	0.004*** (0.0003)
Bad hearing.D			0.0002 (0.001)	0.0004 (0.001)
Country and Time Ds	Yes	Yes	Yes	Yes
No of obs.	68,852	68,852	68,852	68,852
Pseudo R2	0.1849	0.1859	0.2756	0.2758

Note: Standard errors in parentheses. D – denotes a dummy variable.

Specification (1) includes only personal and socioeconomic characteristics as control variables; (2) adds communism dummy for experience during formative years; (3) also includes health habits and health compromising conditions without communism dummy; (4) same as (3) with inclusion of communism dummy. All specifications include country and year of interview dummies. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Determinants of Self-rated health, average marginal effects for the outcome – ‘Fair health’ (3) from ordered probit estimation

Variable	1) SES	2) SES + Com	3) SES+habits +conditions	4) SES+habits +conditions+Com
Communism.D		0.030*** (0.002)		0.014*** (0.002)
Male.D	-0.018*** (0.001)	-0.018*** (0.001)	-0.009*** (0.001)	-0.009*** (0.001)
Age	0.008*** (0.0002)	0.007*** (0.0002)	0.007*** (0.002)	0.006*** (0.0002)
Age-squared	-0.00004*** (0.000)	-2.04e-06*** (0.000)	-0.00004*** (0.000)	-0.00003*** (0.000)
Education	-0.016*** (0.0005)	-0.016*** (0.0005)	-0.013*** (0.001)	-0.013*** (0.001)
Income	-0.015*** (0.0005)	-0.014*** (0.001)	-0.008*** (0.001)	-0.008*** (0.001)
Unemployed.D	0.012*** (0.002)	0.013*** (0.002)	0.024*** (0.002)	0.024*** (0.002)
Military.D	-0.021** (0.009)	-0.022** (0.009)	-0.026*** (0.008)	-0.027*** (0.008)
Retired.D	0.016*** (0.002)	0.015*** (0.002)	0.024*** (0.002)	0.024*** (0.002)
Fruits			-0.006*** (0.001)	-0.006*** (0.001)
Vegetables			-0.003*** (0.001)	-0.003*** (0.001)
Smoke			0.002*** (0.001)	0.003*** (0.001)
Alcohol			-0.008*** (0.001)	-0.007*** (0.0004)
Phys. activity			-0.002*** (0.0002)	-0.002*** (0.0002)
Hospitalized.D			0.062*** (0.002)	0.061*** (0.002)
Disabled.D			0.101*** (0.002)	0.101*** (0.002)
Asthma.D			0.043*** (0.003)	0.042*** (0.003)
Bronchitis.D			0.016*** (0.003)	0.016*** (0.003)
Infarction.D			0.031*** (0.003)	0.031*** (0.003)
Heart disease.D			0.031*** (0.002)	0.031*** (0.002)
High blood pres.D			0.05*** (0.002)	0.05*** (0.002)
Stroke.D			0.061*** (0.004)	0.06*** (0.004)
Arthritis.D			0.031*** (0.002)	0.031*** (0.002)

Arthrosis.D			0.029*** (0.002)	0.029*** (0.002)
Low back pain.D			0.044*** (0.002)	0.044*** (0.002)
Neck pain.D			0.009*** (0.002)	0.009*** (0.002)
Diabetes.D			0.056*** (0.002)	0.055*** (0.002)
Allergy.D			0.016*** (0.002)	0.016*** (0.002)
Stomach.D			0.036*** (0.002)	0.035*** (0.002)
Liver.D			0.047*** (0.004)	0.047*** (0.004)
Cancer.D			0.059*** (0.004)	0.059*** (0.004)
Migraine.D			0.03*** (0.002)	0.03*** (0.002)
Bladder.D			0.003 (0.003)	0.003 (0.003)
Anxiety.D			0.022*** (0.004)	0.022*** (0.004)
Depression.D			0.039*** (0.003)	0.039*** (0.003)
Mental health.D			0.066*** (0.005)	0.065*** (0.005)
Injury.D			0.019*** (0.003)	0.019*** (0.003)
Bad eyes.D			0.015*** (0.001)	0.014*** (0.001)
Bad hearing.D			0.001 (0.003)	0.001 (0.003)
Country and Time Ds	Yes	Yes	Yes	Yes
No of obs.	68,852	68,852	68,852	68,852
Pseudo R2	0.1849	0.1859	0.2756	0.2758

Note: Standard errors in parentheses. D – denotes a dummy variable.

Specification (1) includes only personal and socioeconomic characteristics as control variables; (2) adds communism dummy for experience during formative years; (3) also includes health habits and health compromising conditions without communism dummy; (4) same as (3) with inclusion of communism dummy. All specifications include country and year of interview dummies. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Determinants of Self-rated health, average marginal effects for the outcome – ‘Very good health’ (5) from ordered probit estimation

Variable	1) SES	2) SES + Com	3) SES+habits +conditions	4) SES+habits +conditions+Com
Communism.D		-0.047*** (0.003)		-0.023*** (0.003)
Male.D	0.031*** (0.002)	0.031*** (0.002)	0.015*** (0.002)	0.015*** (0.002)
Age	-0.014*** (0.0003)	-0.011*** (0.0003)	-0.012*** (0.0003)	-0.01*** (0.0004)
Age-squared	0.00007*** (0.000)	0.00004*** (0.000)	0.0001*** (0.000)	0.0001*** (0.000)
Education	0.027*** (0.001)	0.027*** (0.001)	0.022*** (0.001)	0.022*** (0.001)
Income	0.025*** (0.001)	0.025*** (0.001)	0.014*** (0.001)	0.014*** (0.001)
Unemployed.D	-0.021*** (0.004)	-0.023*** (0.004)	-0.039*** (0.004)	-0.04*** (0.004)
Military.D	0.036** (0.015)	0.038** (0.015)	0.047*** (0.015)	0.048*** (0.015)
Retired.D	-0.026*** (0.003)	-0.024*** (0.003)	-0.038*** (0.003)	-0.037*** (0.003)
Fruits			0.011*** (0.001)	0.01*** (0.001)
Vegetables			0.005*** (0.001)	0.005*** (0.001)
Smoke			-0.004*** (0.001)	-0.004*** (0.001)
Alcohol			0.013*** (0.001)	0.013*** (0.001)
Phys. activity			0.003*** (0.0004)	0.003*** (0.0004)
Hospitalized.D			-0.094*** (0.002)	-0.094*** (0.002)
Disabled.D			-0.149*** (0.003)	-0.149*** (0.003)
Asthma.D			-0.068*** (0.004)	-0.068*** (0.004)
Bronchitis.D			-0.027*** (0.004)	-0.027*** (0.004)
Infarction.D			-0.051*** (0.005)	-0.051*** (0.005)
Heart disease.D			-0.05*** (0.003)	-0.05*** (0.003)
High blood pres.D			-0.073*** (0.002)	-0.073*** (0.002)
Stroke.D			-0.094*** (0.005)	-0.094*** (0.005)
Arthritis.D			-0.05*** (0.003)	-0.049*** (0.003)

Arthrosis.D			-0.046*** (0.003)	-0.046*** (0.003)
Low back pain.D			-0.068*** (0.002)	-0.069*** (0.002)
Neck pain.D			-0.016*** (0.003)	-0.016*** (0.003)
Diabetes.D			-0.085*** (0.003)	-0.086*** (0.003)
Allergy.D			-0.027*** (0.003)	-0.027*** (0.003)
Stomach.D			-0.057*** (0.003)	-0.057*** (0.003)
Liver.D			-0.075*** (0.006)	-0.075*** (0.006)
Cancer.D			-0.092*** (0.006)	-0.093*** (0.006)
Migraine.D			-0.049*** (0.003)	-0.05*** (0.003)
Bladder.D			-0.005 (0.005)	-0.006 (0.005)
Anxiety.D			-0.036*** (0.006)	-0.037*** (0.006)
Depression.D			-0.062*** (0.005)	-0.063*** (0.005)
Mental health.D			-0.102*** (0.007)	-0.102*** (0.007)
Injury.D			-0.031*** (0.004)	-0.031*** (0.004)
Bad eyes.D			-0.025*** (0.002)	-0.024*** (0.002)
Bad hearing.D			-0.001 (0.005)	-0.002 (0.005)
Country and Time Ds	Yes	Yes	Yes	Yes
No of obs.	68,852	68,852	68,852	68,852
Pseudo R2	0.1849	0.1859	0.2756	0.2758

Note: Standard errors in parentheses. D – denotes a dummy variable.

Specification (1) includes only personal and socioeconomic characteristics as control variables; (2) adds communism dummy for experience during formative years; (3) also includes health habits and health compromising conditions without communism dummy; (4) same as (3) with inclusion of communism dummy. All specifications include country and year of interview dummies. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Regressions with Elixhauser Comorbidity Index (ECI) as dependent variable, OLS

Variables	1) SES	2) SES + Com	3) SES + habits	4) SES + habits + Com
Communism.D		0.183*** (0.012)		0.181*** (0.012)
Male.D	0.0001 (0.008)	-0.0003 (0.008)	-0.003 (0.008)	-0.003 (0.008)
Age	0.03*** (0.001)	0.019*** (0.001)	0.033*** (0.001)	0.021*** (0.001)
Age-squared	-1.68e-05 (1.50e-05)	7.38e-05*** (1.51e-05)	-4.32e-05*** (1.51e-05)	4.63e-05*** (1.53e-05)
Education	-0.062*** (0.004)	-0.06*** (0.004)	-0.063*** (0.004)	-0.061*** (0.004)
Income	-0.063*** (0.004)	-0.063*** (0.004)	-0.063*** (0.004)	-0.063*** (0.004)
Unemployed.D	-0.04** (0.016)	-0.035** (0.016)	-0.045*** (0.016)	-0.04** (0.016)
Military.D	0.09* (0.051)	0.075 (0.051)	0.097* (0.051)	0.082 (0.051)
Retired.D	0.259*** (0.018)	0.25*** (0.018)	0.26*** (0.018)	0.251*** (0.018)
Fruits			-0.008* (0.004)	-0.006 (0.004)
Vegetables			-9.84e-05 (0.005)	0.001 (0.005)
Phys. activity			-0.024*** (0.002)	-0.024*** (0.002)
Constant	-0.024 (0.028)	0.171*** (0.029)	0.043 (0.034)	0.227*** (0.034)
Country and year Ds	Yes	Yes	Yes	Yes
Observations	68,852	68,852	68,852	68,852
R-squared	0.313	0.315	0.315	0.317

Note: Standard errors in parentheses. D – denotes a dummy variable.

Specification (1) includes only personal and socioeconomic characteristics as control variables; (2) also includes communism dummy indicating the experience during formative years; (3) includes health habits, but excludes communism dummy; (4) same as (3) with inclusion of communism dummy. All specifications include country and year of interview dummies. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Regressions with Charlson Comorbidity Index (CCI) as dependent variable, OLS

Variables	1) SES	2) SES + Com	3) SES + habits	4) SES + habits + Com
Communism.D		0.026*** (0.01)		0.025*** (0.01)
Male.D	0.014** (0.006)	0.014** (0.006)	0.014** (0.007)	0.014** (0.007)
Age	-0.06*** (0.001)	-0.062*** (0.001)	-0.059*** (0.001)	-0.06*** (0.001)
Age-squared	0.001*** (1.35e-05)	0.001*** (1.31e-05)	0.001*** (1.36e-05)	0.001*** (1.32e-05)
Education	-0.017*** (0.003)	-0.017*** (0.003)	-0.019*** (0.003)	-0.019*** (0.003)
Income	-0.041*** (0.003)	-0.041*** (0.003)	-0.041*** (0.003)	-0.041*** (0.003)
Unemployed.D	-0.026** (0.011)	-0.025** (0.011)	-0.029** (0.011)	-0.028** (0.011)
Military.D	0.02 (0.038)	0.018 (0.038)	0.025 (0.038)	0.023 (0.038)
Retired.D	0.327*** (0.016)	0.325*** (0.016)	0.326*** (0.015)	0.325*** (0.015)
Fruits			-0.004 (0.004)	-0.004 (0.004)
Vegetables			0.009** (0.004)	0.009** (0.004)
Phys. activity			-0.02*** (0.001)	-0.02*** (0.001)
Constant	0.835*** (0.022)	0.863*** (0.022)	0.848*** (0.027)	0.874*** (0.027)
Country and year Ds	Yes	Yes	Yes	Yes
Observations	68,852	68,852	68,852	68,852
R-squared	0.753	0.753	0.754	0.754

Note: Standard errors in parentheses. D – denotes a dummy variable.

Specification (1) includes only personal and socioeconomic characteristics as control variables; (2) also includes communism dummy indicating the experience during formative years; (3) includes health habits, but excludes communism dummy; (4) same as (3) with inclusion of communism dummy. All specifications include country and year of interview dummies. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Regressions with Health Perception Gap (HPG) based on Elixhauser Comorbidity Index (ECI) as dependent variable

Variables	OLS				LASSO			
	1) SES	2) SES + Com	3) SES + habits	4) SES + habits + Com	1) SES	2) SES + Com	3) SES + habits	4) SES + habits + Com
Communism.D		-0.114*** (0.01)		-0.107*** (0.01)		-0.128 [0.500]		-0.125 [0.500]
Male.D	0.105*** (0.006)	0.105*** (0.006)	0.116*** (0.006)	0.116*** (0.006)	0.104 [0.498]	0.104 [0.498]	0.115 [0.498]	0.115 [0.498]
Age	-0.019*** (0.001)	-0.012*** (0.001)	-0.02*** (0.001)	-0.013*** (0.001)	-0.016 [18.745]	-0.009 [18.745]	-0.017 [18.745]	-0.009 [18.745]
Age-squared	8.22e-05*** (9.81e-06)	2.55e-05** (1.06e-05)	9.46e-05*** (9.85e-06)	4.14e-05*** (1.06e-05)	- [1853.44]	- [1853.44]	- [1853.44]	- [1853.44]
Education	0.059*** (0.003)	0.058*** (0.003)	0.055*** (0.003)	0.054*** (0.003)	0.056 [1.277]	0.056 [1.277]	0.053 [1.277]	0.052 [1.277]
Income	0.053*** (0.003)	0.052*** (0.003)	0.05*** (0.003)	0.05*** (0.003)	0.052 [1.354]	0.051 [1.354]	0.049 [1.354]	0.049 [1.354]
Unemployed.D	-0.068*** (0.013)	-0.071*** (0.013)	-0.057*** (0.013)	-0.06*** (0.013)	-0.068 [0.230]	-0.071 [0.230]	-0.057 [0.230]	-0.061 [0.230]
Military.D	0.07** (0.034)	0.079** (0.034)	0.067** (0.034)	0.076** (0.034)	0.064 [0.077]	0.074 [0.077]	0.062 [0.077]	0.072 [0.077]
Retired.D	-0.027** (0.011)	0.032*** (0.011)	0.025** (0.011)	0.03*** (0.011)	0.032 [0.445]	0.035 [0.445]	0.030 [0.445]	0.036 [0.445]
Fruits			0.025*** (0.003)	0.024*** (0.003)			0.024 [1.222]	0.023 [1.222]
Vegetables			0.02*** (0.003)	0.02*** (0.003)			0.02 [1.047]	0.020 [1.047]
Phys. activity			0.012*** (0.001)	0.012*** (0.001)			0.011 [2.766]	0.011 [2.766]
Constant	0.19*** (0.023)	0.068*** (0.024)	-0.009 (0.026)	-0.118*** (0.027)	0.351	0.085	0.194	-0.088
Country and year Ds	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	68,852	68,852	68,852	68,852	68,852	68,852	68,852	68,852
R-squared	0.157	0.159	0.162	0.164	0.1572	0.159	0.162	0.1634
Root MSE	0.738	0.737	0.736	0.735	0.745	0.737	0.747	0.735

Note: Standard errors in parentheses for OLS regressions, and predictor specific penalty loadings in squared brackets for the lasso regressions. D – denotes a dummy variable. Derivation of ECI and HPG explained in methodology section, technical calculation details of ECI available in the Appendix.

Specification (1) includes only personal and socioeconomic characteristics as control variables; (2) also includes communism dummy indicating the experience during formative years; (3) includes health habits, but excludes communism dummy; (4) same as (3) with inclusion of communism dummy. All specifications include country and year of interview dummies. *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Regressions with Health Perception Gap (HPG) based on Charlson Comorbidity Index (CCI) as dependent variable

Variables	OLS				LASSO			
	1) SES	2) SES + Com	3) SES + habits	4) SES + habits + Com	1) SES	2) SES + Com	3) SES + habits	4) SES + habits + Com
Communism.D		-0.147*** (0.01)		-0.139*** (0.01)		-0.165 [0.500]		-0.157 [0.500]
Male.D	0.106*** (0.006)	0.106*** (0.006)	0.119*** (0.006)	0.119*** (0.006)	0.105 [0.498]	0.105 [0.498]	0.117 [0.498]	0.117 [0.498]
Age	-0.042*** (0.001)	-0.033*** (0.001)	-0.044*** (0.001)	-0.035*** (0.001)	-0.040 [18.745]	-0.030 [18.745]	-0.041 [18.745]	-0.032 [18.745]
Age-squared	0.0004*** (9.79e-06)	0.0004*** (1.05e-05)	0.0004*** (9.80e-06)	0.0004*** (1.05e-05)	0.0004 [1853.44]	0.0003 [1853.44]	0.0004 [1853.44]	0.0003 [1853.44]
Education	0.077*** (0.003)	0.075*** (0.003)	0.073*** (0.003)	0.072*** (0.003)	0.075 [1.277]	0.073 [1.277]	0.071 [1.277]	0.070 [1.277]
Income	0.069*** (0.003)	0.068*** (0.003)	0.066*** (0.003)	0.065*** (0.003)	0.068 [1.354]	0.067 [1.354]	0.065 [1.354]	0.064 [1.354]
Unemployed.D	-0.058*** (0.013)	-0.062*** (0.013)	-0.045*** (0.013)	-0.049*** (0.013)	-0.058 [0.230]	-0.062 [0.230]	-0.045 [0.230]	-0.050 [0.230]
Military.D	0.039 (0.034)	0.052 (0.034)	0.036 (0.034)	0.047 (0.034)	0.034 [0.077]	0.047 [0.077]	0.03 [0.077]	0.043 [0.077]
Retired.D	0.017 (0.011)	0.024** (0.011)	0.015 (0.011)	0.022** (0.011)	0.022 [0.445]	0.030 [0.445]	0.020 [0.445]	0.027 [0.445]
Fruits			0.029*** (0.003)	0.028*** (0.003)			0.029 [1.222]	0.028 [1.222]
Vegetables			0.023*** (0.003)	0.022*** (0.003)			0.022 [1.047]	0.022 [1.047]
Phys. activity			0.016*** (0.001)	0.016*** (0.001)			0.016 [2.766]	0.016 [2.766]
Constant	0.247*** (0.023)	0.09*** (0.024)	0.014 (0.026)	-0.128*** (0.027)	0.334	0.169	0.112	-0.038
Country and year Ds	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	68,852	68,852	68,852	68,852	68,852	68,852	68,852	68,852
R-squared	0.131	0.133	0.138	0.141	0.1304	0.1333	0.1379	0.1405
Root MSE	0.739	0.738	0.736	0.735	0.739	0.738	0.736	0.735

Note: Standard errors in parentheses for OLS regressions, and predictor specific penalty loadings in squared brackets for the lasso regressions. D – denotes a dummy variable. Derivation of ECI and HPG explained in methodology section, technical calculation details of ECI available in the Appendix.

Specification (1) includes only personal and socioeconomic characteristics as control variables; (2) also includes communism dummy indicating the experience during formative years; (3) includes health habits, but excludes communism dummy; (4) same as (3) with inclusion of communism dummy. All specifications include country and year of interview dummies. *** p<0.01, ** p<0.05, * p<0.1.

Table 12: Regressions with within country Health Perception Gap (HPG) based on Elixhauser Comorbidity Index (ECI) as dependent variable

Variables	OLS				LASSO			
	1) SES	2) SES + Com	3) SES + habits	4) SES + habits + Com	1) SES	2) SES + Com	3) SES + habits	4) SES + habits + Com
Communism.D		-0.019** (0.008)		-0.032*** (0.008)		-0.03 [0.500]		-0.042 [0.500]
Male.D	0.107*** (0.006)	0.107*** (0.006)	0.116*** (0.006)	0.116*** (0.006)	0.105 [0.498]	0.105 [0.498]	0.115 [0.498]	0.115 [0.498]
Age	-0.017*** (0.001)	-0.016*** (0.001)	-0.019*** (0.001)	-0.017*** (0.001)	-0.015 [18.745]	-0.013 [18.745]	-0.016 [18.745]	-0.014 [18.745]
Age-squared	5.45e-05*** (9.56e-06)	4.4e-05*** (1.03e-05)	7.41e-05*** (9.63e-06)	5.77e-05*** (1.03e-05)	- [1853.44]	- [1853.44]	- [1853.44]	- [1853.44]
Education	0.058*** (0.003)	0.058*** (0.003)	0.056*** (0.003)	0.056*** (0.003)	0.056 [1.277]	0.056 [1.277]	0.054 [1.277]	0.054 [1.277]
Income	0.038*** (0.002)	0.038*** (0.002)	0.038*** (0.002)	0.037*** (0.002)	0.038 [1.354]	0.037 [1.354]	0.038 [1.354]	0.036 [1.354]
Unemployed.D	-0.089*** (0.013)	-0.09*** (0.013)	-0.077*** (0.013)	-0.079*** (0.013)	-0.088 [0.230]	-0.09 [0.230]	-0.076 [0.230]	-0.079 [0.230]
Military.D	0.062* (0.034)	0.062* (0.034)	0.058* (0.034)	0.059* (0.034)	0.055 [0.077]	0.057 [0.077]	0.052 [0.077]	0.054 [0.077]
Retired.D	0.06*** (0.011)	0.062*** (0.011)	0.055*** (0.011)	0.059*** (0.011)	0.064 [0.445]	0.068 [0.445]	0.06 [0.445]	0.064 [0.445]
Fruits			0.017*** (0.003)	0.017*** (0.003)			0.017 [1.222]	0.016 [1.222]
Vegetables			0.023*** (0.003)	0.023*** (0.003)			0.023 [1.047]	0.022 [1.047]
Phys. activity			0.015*** (0.001)	0.016*** (0.001)			0.014 [2.766]	0.015 [2.766]
Constant	0.246*** (0.019)	0.228*** (0.02)	0.048** (0.023)	0.018 (0.024)	0.289	0.215	0.143	0.046
Year Ds	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	68,852	68,852	68,852	68,852	68,852	68,852	68,852	68,852
R-squared	0.111	0.111	0.116	0.117	0.1109	0.111	0.1165	0.1167
Root MSE	0.734	0.734	0.732	0.732	0.736	0.735	0.737	0.733

Note: Standard errors in parentheses for OLS regressions, and predictor specific penalty loadings in squared brackets for the lasso regressions. D – denotes a dummy variable. Derivation of ECI and HPG explained in methodology section, technical calculation details of ECI available in the Appendix.

Specification (1) includes only personal and socioeconomic characteristics as control variables; (2) also includes communism dummy indicating the experience during formative years; (3) includes health habits, but excludes communism dummy; (4) same as (3) with inclusion of communism dummy. All specifications include country and year of interview dummies. *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Regressions with within country Health Perception Gap (HPG) based on Charlson Comorbidity Index (CCI) as dependent variable

Variables	OLS				LASSO			
	1) SES	2) SES + Com	3) SES + habits	4) SES + habits + Com	1) SES	2) SES + Com	3) SES + habits	4) SES + habits + Com
Communism.D		5.79e-05 (0.008)		-0.017** (0.008)		-0.011 [0.500]		-0.027 [0.500]
Male.D	0.109*** (0.006)	0.109*** (0.006)	0.12*** (0.006)	0.12*** (0.006)	0.108 [0.498]	0.108 [0.498]	0.118 [0.498]	0.118 [0.498]
Age	-0.04*** (0.001)	-0.04*** (0.001)	-0.042*** (0.001)	-0.041*** (0.001)	-0.037 [18.745]	-0.037 [18.745]	-0.039 [18.745]	-0.038 [18.745]
Age-squared	0.0004*** (9.59e-06)	0.0004*** (1.04e-05)	0.0004*** (9.64e-06)	0.0004*** (1.03e-05)	0.0004 [1853.44]	0.0004 [1853.44]	0.0004 [1853.44]	0.0004 [1853.44]
Education	0.076*** (0.003)	0.076*** (0.003)	0.073*** (0.003)	0.073*** (0.003)	0.074 [1.277]	0.074 [1.277]	0.071 [1.277]	0.071 [1.277]
Income	0.046*** (0.002)	0.046*** (0.002)	0.046*** (0.002)	0.046*** (0.002)	0.045 [1.354]	0.045 [1.354]	0.046 [1.354]	0.045 [1.354]
Unemployed.D	-0.091*** (0.013)	-0.091*** (0.013)	-0.076*** (0.013)	-0.077*** (0.013)	-0.090 [0.230]	-0.091 [0.230]	-0.076 [0.230]	-0.077 [0.230]
Military.D	0.032 (0.035)	0.032 (0.035)	0.028 (0.034)	0.028 (0.034)	0.026 [0.077]	0.026 [0.077]	0.022 [0.077]	0.023 [0.077]
Retired.D	0.078*** (0.011)	0.078*** (0.011)	0.073*** (0.011)	0.075*** (0.011)	0.083 [0.445]	0.084 [0.445]	0.078 [0.445]	0.081 [0.445]
Fruits			0.021*** (0.003)	0.02*** (0.003)			0.02 [1.222]	0.02 [1.222]
Vegetables			0.023*** (0.003)	0.023*** (0.003)			0.023 [1.047]	0.023 [1.047]
Phys. activity			0.02*** (0.001)	0.02*** (0.001)			0.019 [2.766]	0.019 [2.766]
Constant	0.341*** (0.019)	0.341*** (0.02)	0.115*** (0.023)	0.1*** (0.024)	0.306	0.295	0.084	0.059
Year Ds	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	68,852	68,852	68,852	68,852	68,852	68,852	68,852	68,852
R-squared	0.061	0.061	0.07	0.07	0.0611	0.0611	0.0695	0.0696
Root MSE	0.737	0.737	0.734	0.734	0.737	0.737	0.734	0.733

Note: Standard errors in parentheses for OLS regressions, and predictor specific penalty loadings in squared brackets for the lasso regressions. D – denotes a dummy variable. Derivation of ECI and HPG explained in methodology section, technical calculation details of ECI available in the Appendix.

Specification (1) includes only personal and socioeconomic characteristics as control variables; (2) also includes communism dummy indicating the experience during formative years; (3) includes health habits, but excludes communism dummy; (4) same as (3) with inclusion of communism dummy. All specifications include country and year of interview dummies. *** p<0.01, ** p<0.05, * p<0.1.

Table 14: Effects of Health Perception Gap on various health habits related variables

	Probit – Average marginal effects				OLS	
	Benefits	Help	No to hospital	Not see specialist	Difficulty	GP consultations
HPG1 (ECI)	-0.012** (0.002)	-0.064** (0.005)	-0.022** (0.001)	-0.048** (0.002)	-0.08** (0.002)	-0.148** (0.007)
HPG2 (CCI)	-0.016** (0.002)	-0.064** (0.005)	-0.024** (0.001)	-0.054** (0.002)	-0.083** (0.002)	-0.162** (0.007)
HPGI (wc ECI)	-0.011** (0.002)	-0.064** (0.005)	-0.021** (0.001)	-0.047** (0.002)	-0.081** (0.002)	-0.147** (0.007)
HPGII (wc CCI)	-0.016** (0.002)	-0.063** (0.005)	-0.023** (0.001)	-0.052** (0.002)	-0.084** (0.002)	-0.161** (0.007)

Note: Robust standard errors in parentheses. Additional explanatory variables included in the specifications but not reported here: personal and socio-economic characteristics, country and year dummies, indications of health: being disabled and, depending on which comorbidity index HPG is based on, either ECI or CCI for an overall health status measure. ** p<0.01, * p<0.05, + p<0.1.

Appendix

Information on variables

Variable/name in EHIS dataset	Answer choices/ explanation
Being in hospital as an inpatient over past 12 months/ hc01 (recoded)	0. Not hospitalized; 1. Hospitalized
Permanently disabled/ hh08 (category)	0. Not disabled; 1. Disabled
Experience of physical pain or discomfort/ sf01	1. None; 2. Mild; 3. Moderate; 4. Severe; 5. Extreme
Self-rated Health/ hs01 (recoded)	1. Very bad, 2. Bad, 3. Fair, 4. Good, 5. Very good
Health conditions/ hs04a-u (recoded), and pl01, pl04 (all recoded)	Asthma, Bronchitis, Infarction, Coronary heart disease, High blood pressure, Stroke, Arthritis, Osteoarthritis, Low back problems, Neck problems, Diabetes, Allergies, Stomach problems, Liver problems, Cancer, Migraine, Urinary tract problems, Anxiety, Depression, Mental health problems, permanent injury, bad eyesight, bad hearing, dummies (0. no condition, 1. condition present)
Psychoses/ hs04r, hs04t (if either is present)	0. If no chronic anxiety or other mental health problems present; 1. If has chronic anxiety or other mental health problems
Obese/ bmi03 (if >30)	0. If not obese; 1. If considered obese
Severely underweight/ bmi03 (if <15)	0. If not severely underweight; 1. If considered severely underweight
Age	Mid-points of available age groups: 16, 18.5, 22, 27, 32, 37, 42, 47, 52, 57, 62, 67, 72, 77, 82, 85
Age squared	Derived from Age
Education / hh07	1. No formal education, 2. Primary, 3. Lower secondary, 4. Upper secondary, 5. Post-secondary but non-tertiary, 6. First stage of tertiary, 7. Second stage of tertiary
Gender	0. Female, 1. Male
Occupation dummies/ hh08 (separate categories)	Work in military, Unemployed, Retired (1. If person belong to the category and, 0. If not)

Income/ in04	1. Below 1 st quintile, 2. Between 1 st and 2 nd quintiles, 3. Between 2 nd and 3 rd quintiles, 4. Between 3 rd and 4 th quintiles, 5. Above 4 th quintile (country specific range and currency)
Alcohol consumption frequency/ al01	1. Never, 2. Monthly or less, 3. 2 to 4 times a month, 4. 2 to 3 times a week, 5. 4 to 6 times a week, 6. Every day
Alcohol abuse/ al01 (if drinking daily)	0. If not considered abusing alcohol; 1. If consumes alcohol daily
Smoking/sk01 (recoded)	1. Not at all, 2. Yes, occasionally, 3. Yes, daily
Moderate physical exercise (days per week) / pe03	1-7 days
How much fruit in diet/ fv01 (recoded)	1. Never, 2. Less than once a week; 3. Less than 4 times a week, but at least once a week; 4. Less than once a day but at least 4 times a week, 5. Once a day, 6. Twice or more a day
How much vegetables in diet/ fv02 (recoded)	1. Never, 2. Less than once a week; 3. Less than 4 times a week, but at least once a week; 4. Less than once a day but at least 4 times a week, 5. Once a day, 6. Twice or more a day
Claiming any benefits in household/ in01b-in01h (recoded)	0. No benefits in the household claimed; 1. Claiming any type of benefits in the household
Receiving help for personal or household activities/ pc02 and ha03 (recoded)	0. No help received; 1. Any type of help received
Not being hospitalized even if advised to over 12 months/hc06 (recoded)	0. Going to the hospital if needed; 1. Not being hospitalized even if necessary
Not visiting a specialist even if in need over 12 months / hc14 (recoded)	0. Seeing a medical specialist if necessary; 1. Not seeing a medical specialist even if necessary
Average difficulty level in performing tasks alone/pc01a-pc01e (recoded)	In original questions: 1. No difficulty; 2. Yes, some difficulty; 3. Yes, a lot of difficulty; 4. I can't achieve it by myself; the average of responses to: feeding yourself; getting in and out of bed or chair; dressing and undressing; using toilets; bathing and showering
GP visits over last 4 weeks/ hc11	Number of times visited

Elixhauser Comorbidity Index

Condition	How addressed
1. Congestive Heart Failure	
2. Cardiac Arrhythmia	
3. Valvular Disease	
4. Pulmonary Circulation Disorders	
5. Peripheral Vascular Disorders	
6. Hypertension without complications	Not possible to differentiate between types of Hypertension, but these two categories should be mutually exclusive thus included with indication of Hypertension
7. Hypertension with complications	
8. Paralysis	N/A
9. Other Neurological Disorders	N/A, used indication of Stroke (cerebral haemorrhage, cerebral thrombosis) instead
10. Chronic Pulmonary Disease	The category available besides Chronic pulmonary disease also includes Chronic bronchitis and Emphysema
11. Diabetes without complications	Not possible to differentiate between types of Diabetes, but these two categories should be mutually exclusive thus included with indication of Diabetes
12. Diabetes with complications	
13. Hypothyroidism	N/A
14. Renal Failure	N/A
15. Liver Disease	Cirrhosis of liver, liver dysfunction
16. Peptic Ulcer Disease excluding bleeding	Stomach ulcer (gastric or duodenal ulcer), no indication of bleeding available
17. HIV/AIDS	N/A
18. Lymphoma	Not possible to differentiate between types of Cancer, but indication of Cancer includes tumour, Leukaemia and Lymphoma
19. Metastatic Cancer	
20. Solid Tumour without Metastasis	
21. Rheumatoid Arthritis/Collagen	No information on Collagen but Rheumatoid Arthritis indicated
22. Coagulopathy	N/A
23. Obesity	If BMI>30

24. Weight Loss	If BMI<15
25. Fluid and Electrolyte Disorders	N/A
26. Blood Loss Anemia	N/A
27. Deficiency Anemia	N/A
28. Alcohol Abuse	Indicated as present if alcohol consumed daily
29. Drug Abuse	N/A
30. Psychoses	If either Chronic Anxiety or Other mental problems present
31. Depression	Indicated

Charlson Comorbidity Index (sum of weighted conditions)

Rule	How addressed
Age: <50=0; 50-59=1; 60-69=2; 70-79=3; 80+=4;	As indicated
Diabetes: none=0; uncomplicated=1; end-organ damage=2;	As indicated and for end-organ damage - since diabetes can cause hypertension and strokes if both present assumed=2;
Liver disease: none=0; mild=1; moderate to severe=3;	As indicated, severe case assumed if pain experienced is severe or extreme (sf01)
Tumour: none=0; localized=2; metastatic=6;	Available as Cancer, metastatic case assumed if pain experienced is severe or extreme (sf01)
AIDS: no=0; yes=6;	Not available
Chronic Kidney Disease: no=0; moderate to severe=2;	Not available
Congestive Heart Failure: no=0; yes=1;	Not available, assumed present if hypertension and coronary heart disease together are present because both are causes
Myocardial Infarction: no=0; yes=1;	As indicated
Chronic Obstructive Pulmonary Disease: no=0; yes=1;	As indicated, the category available also includes Chronic bronchitis and Emphysema
Peripheral Vascular Disease: no=0; yes=1;	Not available
Cerebrovascular accident or Transient Ischemic Attack: no=0; yes=1;	Indication of Stroke used instead
Dementia: no=0; yes=1;	Not available
Hemiplegia: no=0; yes=2;	Not available, but is considered to be caused by the factors that are included
Connective tissue disease: no=0; yes=1;	Not available
Leukemia: no=0; yes=2;	Available as Cancer only, included previously
Malignant lymphoma: no=0, yes=2;	Available as Cancer only, included previously
Peptic ulcer disease: no=0; yes=1;	As indicated

Derivation differences in HPG measures

HPG measure:	ECI	CCI	European avg.	Within-country avg.
HPG1	✓		✓	
HPG2		✓	✓	
HPGI	✓			✓
HPGII		✓		✓

Basic descriptive characteristics by country

Variable/Country	Belgium	Bulgaria	Cyprus	Czech Republic	Greece	Hungary	Latvia	Malta	Poland	Romania	Slovenia	Slovakia
Male (D)	0.468	0.465	0.473	0.475	0.4	0.452	0.44	0.480	0.432	0.464	0.473	0.469
Age	48.8	51.76	46.02	48.64	58	47.97	46.3	47.02	48.16	47.08	46.29	44.29
Education	4.363	3.9	3.728	4.049	1.4	4.048	4.068	4.015	3.78	3.667	3.498	4.232
Income	3.535	3.158	2.897	3.305	2.6	2.857	2.77	3.128	3.056	1.49	3.09	3.064
Unemployed (D)	0.064	0.082	0.044	0.029	0	0.075	0.067	0.034	0.06	0.028	0.074	0.061
Military (D)	0	0.008	0.019	0.004	0	0.007	0.002	0	0.006	0.004	0.006	0.003
Retired (D)	0.26	0.359	0.195	0.319	0.4	0.266	0.239	0.149	0.298	0.29	0.268	0.216
SRH	3.943	3.64	4.17	3.748	1.6	3.469	3.374	3.986	3.526	3.91	3.673	3.717
Total no of obs.	4,565	4,533	6,405	1,406	5	4,509	6,180	2,100	22,806	10,667	1,539	4,137

Conclusion

The aim of this thesis was to look into different outcomes likely affected by the communist regime in former Eastern Bloc. This research shows that the communism experiment in the former Eastern Bloc had a long-term impact and could still be indirectly influencing economic, social capital, human capital, preferences and perceptions. However, it is not easy to claim this impact is solely negative or positive. We explored a range of different ways to define the experience or exposure to the communism, including different periods of a person's life and the amount of time spent in the regime.

In the first chapter we looked into how early life (0-16 years) spent in the communist regime affected the highest education attainment at the individual level. Findings suggest that the communist regimes' attempt to increase the education attainments in general public was quite successful, showing evidence of secondary education completion rates being even higher than those of other European countries and people who were not affected by the regime during the early years. However, there is not enough evidence to claim this education policy affected tertiary attainment also.

In the second chapter, we find evidence supporting the anecdotal claims about Eastern Europeans drinking more frequently and being more likely to binge drink. We observe some gender-specific differences in drinking behaviour. Surprisingly, contradicting the existing literature, results suggest that years spent in the communist regime and its experience during the formative years respectively increased the probability of alcohol consumption being more frequent among women, this effect also being larger in comparison to men even if both are highly significant. However, we did not find that either of those communism measures would have an impact on the incidence of binge drinking for women, but the effect is comparatively large and significant for men. Especially spending the formative years in the communist regime seems to be an important factor in increasing the probability of male binge drinking. This suggests that the number of years spent under communism and formative years experience could be seen as representations of two different transmission mechanisms: the habitual and the socialization respectively.

Finally, in the last chapter, we find that living in the communist regime during the person's formative years had a long-term effect on own health perception gap in

comparison to those who did not have this experience during these years. Our findings show that the communism experience negatively affected the evaluation of own health but its effect on the more objective health measures – Elixhauser and Charlson comorbidity indices is much smaller even though still negative, which indicates perception to health potentially being more negatively biased. This encouraged looking into calculating the health perception gap and examining the effect of the communism on it. We indeed find the health perception gap to be more negative for the people who lived in the communist regime in their formative years. This finding is not only curious on its own but it could be suspected to have an effect on a person's behaviour. Preliminary findings show that more positive health perception gap measures are significantly associated with lower reported difficulty levels of engaging in personal activities without help, lower probability of people receiving personal and household help, fewer monthly GP visits, and lower probability of claiming any kind of benefits by the household. On the other hand, more positive perception is related to a higher chance of not refusing to be an inpatient in the hospital if required by a doctor or seeing a health specialist if necessary. Even though these findings need to be examined in more detail to draw more robust conclusions, but it can be suspected such behaviour on the aggregate level could affect a country's spending on healthcare and welfare.

We could draw a range of policy suggestions from these findings. Firstly, it is probably not surprising that promoting education and reducing its cost on young people is likely to help increase overall attainment level. Secondly, more attention should be paid to young people, and our research so far confirms the importance of the formative years towards behaviour, habits and even perception for the rest of a person's life. In relation to our findings more specifically, by helping avoid or deal with the triggers increasing the probability of alcohol abuse in the formative years, could help avoid this becoming a long-run habit, which could lead to more health related problems in later years. This probably implies more support and attention being paid to young people in need, those experiencing any kind of financial or psychological problems during the formative years that could leave a scarring effect and influence the choice of the crowd to socialize with. These factors are likely to have an impact on the general adopted attitude during the formative years, which would determine attitudes and perception in later years. This research supports the idea that such 'soft' measures are important in economics and should probably receive more

attention in policymaking decisions. Lastly, we find evidence that the communist regime experience had a long-term effect on a range of outcomes and there is no reason to think more areas related to social and human capital or cultural norms were not influenced. Based on this research, we can infer that large scale political and economic changes are likely to leave an impact which might be felt for some time. This might sound like a quite vague idea but it should probably be routinely reminded to the policy makers frequently seeking short-term gains, and it should be seen as one of the fundamental elements in the way we think about the policymaking and its implications, to help avoid the lack of accountability. Finally, to all the supporters of communism – revolution and imposition of a new lifestyle might not be the best way to make Marx’s vision reality.