Università
Ca' Foscari
Venezia

Scuola Dottorale di Ateneo
Graduate School

Dottorato di Ricerca in Economia
Ciclo XXVI
Anno di Discussione 2015

Three essays on socio-economic determinants of health: a life-course approach

settore scientifico disciplinare di afferenza: SECS-P/01, SECS-P/05

Tesi di dottorato di **Liudmila Antonova**

Matricola 955824

Coordinatore del Dottorato	Tutore del dottorando
Prof. Michele Bernasconi	**Prof. Giacomo Pasini**

The undersigned Liudmila Antonova, in his quality of doctoral candidate for a Ph.D. degree in Economics granted by the Advanced School of Economics of Ca' Foscari University of Venice, attests that the research exposed in this dissertation is original and that it has not been and it will not be used to pursue or attain any other academic degree of any level at any other academic institution, be it foreign or Italian.

Abstract

The thesis is composed of three chapters that analyse the influence of different socio-economic factors on health. The common theoretical background of these studies is the Grossman model (1972) that considers health to be a stock of capital in a similar manner as financial and human capital. Health gets depreciated with age and can be increased by health investments, such as medical care and health behavior. Adverse events at one period of life can have an immediate impact on the stock of health and, if not being compensated by extra investment into health, a negative effect can persist for a long period of time.

Adverse socio-economic events related to health can be caused by the external circumstances or an individual decision to engage into health-damaging activities because they induce utility or are associated with monetary gains. I consider three types of socio-economic factors affecting health. The first one is macroeconomic crisis. Since crises do not depend on individual decisions they represent the perfect case of completely exogenous events that affect the whole population of the country. The second factor is rules of health behavior that can be followed without extra money investments. They are on the opposite side of the scale "exogenous - self-selected" than the economic crises and depend completely on the choice of individuals. The third factor is working conditions at the main job, which is somewhere in the middle of the scale. People might self-select into health damaging jobs, because they are better paid, but from the other side disadvantaged groups from lower socio-economic classes with lower level of education might have less occupational choices and are constrained to engage into physically of psychologically demanding jobs.

The policy maker might want to distinguish between self-selected versus exogenously predefined types of health-affecting events, and might want to protect at the first place the exogenously disadvantaged sub-groups of population rather than those who self-select into health-damaging activities.

The Chapter 1, co-authored with Tabea Bucher-Koenen from the Munich Center for the Economics of Ageing and Fabrizio Mazzonna from Università della Svizzera Italiana, studies the long-term effect of macroeconomic crises at the prime working age on health late in life using data from 11 European countries. We find that experiencing a severe crisis in which GDP dropped by at least 1% significantly reduces health later in life. Respondents hit by such a shock rate their subjective health as worse, are more likely to suffer from chronic diseases and mobility limitations, and have lower grip strength. The effects are larger among high-school dropouts.

The Chapter 2 considers the impact of decisions concerning rules of health behavior compared to the impact of medical care and takes a more structural approach to the problem. I modify the Grossman model of demand for health introducing in it an additional factor: disutility-inducing health rules. I derive an equation of demand for health, estimate it using Russian Longitudinal Monitoring Survey data and recover a linear relationship between parameters of health production function. Knowledge of this relationship allows me performing a back-of-the-envelope cost-benefit analysis that demonstrates that financing a health behavior promoting campaign can be way

more beneficial in terms of overall health gains than financing an additional medical care utilization.

The Chapter 3 studies the associations between working conditions at the main job and physical and mental health after retirement using SHARE data set. Further self-reported working conditions are instrumented with average across occupation and country, recognizing the potential common bias in self-reported health and working conditions at the main job. Instrumental variables analysis shows that there is a strong positive association between adverse working conditions and poor physical health, while no association with mental health measured as number of symptoms of depression is found.

Acknowledgements

I would like to thank my supervisor Giacomo Pasini for his kind guidance, support, help, constructive critics, for teaching me to apply the Stata commands in a conscious way and for encouragement. I am sure that in difficult or challenging moments in life I will recall Giacomo's words: "Do not underestimate yourself!.

I also thank Agar Brugiavini, my second supervisor, who's course of Empirical Economics actually inspired me to start a PhD at the first place. Besides very helpful discussions, I should also thank Agar for the opportunity of meeting in person and listening to the talks of famous economists from different countries and universities who were coming to conferences and seminars in Venice.

Remembering good times of my visiting period in Munich Center for the Economics of Ageing, I want to thank Tabea Bucher-Koenen and Fabrizio Mazzonna. Working on a joint paper for me was an exciting experience of how interesting and enriching it can be to do an academic research together. I hope our collaboration continues in the future.

Special gratitude to my colleagues Daria, Luis, Ludovico, Cristina and Enkelejda: thanks to them I was coming to my office in the mornings with pleasure.

I am grateful to my Mom, Dad and brother back to Russia. All these years they were very supportive and comforting, sometimes just expressing the wish that I would manage to graduate before Venice goes completely under water.

Special thanks goes to Patrick for his support, love and patience.

Last, but not the least I thank *un omo de oro* Sergio Barichello and the dragon boat team of universities Ca'Foscari and IUAV. The time I spent in the lagoon rowing helped me to keep the peace of mind in the last months of the PhD.

Contents

List of Tables

List of Figures

Chapter 1

Macroeconomic crunches during working years and health outcomes later in life

Liudmila Antonova [1], Tabea Bucher-Koenen [2] and Fabrizio Mazzonna [3]

Acknowledgement: We would like to thank Axel Börsch-Supan, Davide Cantoni, Paola Giuliano, Hendrik Jürges and Joachim Winter for helpful comments. We are grateful to seminar participants in Munich as well as conference participants at the PhD Workshop in Grindelwald (CH, January 2014) and at the annual conference of the German Association for Health Economics (Munich, March 2014). This paper uses data from SHARE release 2.5.0, as of May 24th 2011. The SHARE data collection has been primarily funded by the European Commission through the 5th framework programme (project QLK6-CT-2001- 00360 in the thematic programme Quality of Life), through the 6th framework programme (projects SHARE-I3, RII-CT- 2006-062193, COMPARE, CIT5-CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th framework programme (SHARE-PREP, 211909 and SHARE-LEAP, 227822). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04-064, IAG BSR06-11, R21 AG025169) as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions)

Keywords: SHARE; Health; Macroeconomic Conditions; Long-run Effects.

JEL: JEL: I15; J14; N14

[1] Advanced School of Economics, Ca'Foscari University, Venice, Italy. *antonova@unive.it*

[2] Max-Planck-Institute for Social Law and Social Policy; Amalienstrasse 33, 80799 Munich, Germany. Phone: +49-89-38602-332; *bucher-koenen@mea.mpisoc.mpg.de*

[3] Universita della Svizzera italiana , Lugano, Switzerland. *fabrizio.mazzonna@usi.ch*

1.1 Introduction

The recent economic crises and high unemployment rates especially among young Europeans have spiked a debate about the short- and long-term effects of macroeconomic conditions on population well being. In this paper we want to contribute to this debate by reporting evidence of negative long term effect of past economic crises on prime age workers.

Economic crises are seen as times of severe economic downturn, i.e., times of low economic growth, high inflation, and high unemployment. Generally, they are perceived to put a burden on population health. Shocks to wealth and income, less access to social protection and care such as health care, and an increase in stress due to job loss or job insecurity, are detrimental for health (see, e.g., Sullivan and von Wachter, 2009; Schwandt, 2014 for recent evidence on the negative effects of unemployment and stock market fluctuations on individual's health, respectively). The effects might be particularly severe among vulnerable subgroups. On the other hand aggregate mortality has been shown to to be positively related with business cycle fluctuations, i.e. in times of economic growth mortality increases and in recessions mortality declines (see, e.g., Ruhm, 2000, 2005). The explanations put forward are that during economic downturns opportunity costs of time decrease and individuals may follow healthier life-styles, i.e. smoke and drink less and spend more time exercising and eating healthy. Additionally there are fewer costs due to external effects like pollution and congestion that cause detrimental effects on health. However, there is still an ongoing debate about the cyclical movements of mortality.[4] Results, for example, differ by period of study, the selection of countries and their level of social protection (Stuckler et al., 2009), business cycle indicator, cause of mortality and the data used (individual outcomes vs. aggregate mortality). Most of the studies so far only consider immediate or very short term effects of economic crises (up to a four year lag, e.g., Ruhm (2000); Gerdtham and Ruhm (2006)).

Fewer studies have looked into the long-term effects of economic fluctuations because of important identification issues. In particular, it is very complicated to identify a proper control group, since when a crisis hits a country it might potentially affect all its citizens. One way to overcome the problem is to consider crises experienced during critical periods in life such as early childhood. Favorable economic conditions at the time of birth have been found to lower mortality and increase cog-

[4] See, e.g., Stevens et al., 2011 or Ruhm, 2013 for recent contributions and Ruhm, 2012 for a review of the literature.

nitive functioning later in life (see, e.g., van den Berg et al., 2006, 2009; Doblhammer et al., 2011). The authors argue, that adverse economic conditions around birth have a negative influence on the quality and quantity of nutrition, access to health care in early life, crowded housing conditions which cause a higher exposure to disease and increase the stress levels which have a long-term negative effect on health. The effect is particularly strong for individuals with low education.

Another critical period is early adulthood, when individuals make the transition from school to work. Unfavorable economic conditions during this period, may lead to worse labor market trajectories for those cohorts compared to cohorts graduating during a boom (see, e.g., Kahn, 2010; Oreopoulos et al., 2012). This in turn might lead to worse health outcomes later in life. Hessel and Avendano (2013) find that unemployment rates experienced in early adulthood (when leaving school) are associated with better physical health for men and worse physical health among women when old. Cutler et al. (2014) find that higher unemployment rates during graduation are related to lower life-time income, life-satisfaction and adverse health outcomes. Additionally they find that the level of education can play a protective role.

While it is important to know how macroeconomic conditions at critical periods in childhood and early adulthood influence outcomes later in life, little is known about the effect of crises during adulthood on health at older ages. Our objective is to study the effects of severe macroeconomic shocks which occur during years when a person is most likely to be active on the labor market (age 20 to 50) on health outcomes later in life. While there is some overlap between our study and studies investigating the effects of crises during early adulthood, we can add a layer of complexity by first looking into the effects of adverse conditions during adulthood in general. In a second step we split this rather large period into smaller age windows to look for critical periods. We run regressions controlling for crises experienced during different age windows simultaneously to see which periods are the most sensitive for invidivuals' health later in life. Leist et al., 2013 use a very similar approach: they analyze effect of business cycle fluctuations on cognitive functioning using data from 11 European countries. They find a negative effect of downward fluctuations of the business cycle on cognitive abilities for women experienced in early and mid adulthood and negative effects of recessions experienced after age 45 for men. The proposed mechanisms are that men have a higher probability to experience a lay-off, whereas women have a higher downward occupational mobility and are more likely to work part-time. Our approach differs in two central respects. First, we look into various different physical and mental health measures to get a broader picture of the effects of macroeconomic conditions on health later in life. Second, different from previous literature (e.g. van den Berg et al., 2009; Doblhammer et al., 2011), we are not interest in the effects of business cycle fluctuation but we want to focus only severe macroeconomic shocks. We define crises as the 5% worst years in terms of GDP growth experienced in Europe during the period 1954 to 2004. The effect of economic fluctuations is identified by comparing different cohorts across different European countries. As in van den Berg et al. (2009) and Doblhammer et al. (2011) the identification comes from the cohort specific deviation from their (country specific) long term health trend. However, different from these studies we do not focus on economic crises around birth, but on those crises that hit the cohorts of interest during their working life.

We use the first and second wave of the Survey of Health Aging and Retirement in Europe (SHARE) to shed some light on this relationship. Specifically we use individual level data from more than 20,000 individuals age 50+ from 11 European countries with different labor market and social policies. We match information about the number of country-specific macroeconomic crises during individuals' potentially active years (age 20 to 50) to the SHARE data. Our objective is to analyze the relation between macroeconomic conditions during individuals' working years and their health later in life. We focus on respondents between age 50 and 70 at the time of the SHARE interview. Differing crisis periods between European countries make our study particularly powerful. Furthermore, the SHARE data offers us a very rich set of health variables. While our main analysis focuses on self-reported health outcomes, we are also investigating the effects of crises on a larger set of objective health measures, specifically the number of chronic conditions, the number of symptoms, limitations in the instrumental activities of daily living (IADL), mobility limitations, depressive symptoms, grip strength, and recall abilities as a proxy for cognitive functioning.

We find a significant negative effect of the number of crises experienced between age 20 and 50 on self-reported health later in life. Respondents hit by such a shock rate their health as significantly worse compared to cohorts who experienced no severe macroeconomic shocks. The effect of experiencing one additional severe crisis is approximately equivalent in size to becoming two years older. Moreover, respondents who experienced a severe macroeconomic downturn suffer from more symptoms, such as fatigue or pain, report more chronic health problems and mobility limitations, have lower grip strength and lower recall abilities. The effects are substantially stronger for respondents with low education. For them one more crisis year leads to a decrease in self-reported health equivalent to about four more years of age. Moreover, we find a significantly higher number of symptoms and a decrease in grip strength more or less equivalent to 1.5 more years of age.

Only the worst 5% of the crisis have an effect on population health. If we relax the crisis definition and instead include the worst 10% of the crisis years, the effects on health become smaller and insignificant. Thus, only severe (enough) crisis seem to have a long-term impact on people's health. The results persist after performing many possible robustness checks. For instance, we account for selective mortality and for the effects of World War II. We also perform a sort of placebo test using childhood socio-economic and health conditions as dependent variables. We take this as evidence, that general cohort trends in health are not driving our results.

Our paper proceeds as follows. In section 1.2 we will introduce the SHARE data set, the variables we use, and our measure of macro-economic crises. Section 1.3 describes our empirical strategy. In section 1.4 we will present our results followed by some robustness checks in section 1.5 and a conclusion in 1.6.

1.2 Data

1.2.1 SHARE data

We use data from the Survey of Health, Ageing and Retirement in Europe (SHARE), a multidisciplinary and cross-national bi-annual household panel survey coordinated by the Munich Center for the Economics of Aging (MEA) with the technical support of CentERdata at Tilburg University.[5] The survey collects data on health, socio-economic status, and social and family networks for nationally representative samples of elderly people in the participating countries. The target population consists of individuals aged 50 and older who speak the official language of each country and do not live abroad or in an institution, plus their spouses or partners irrespective of age. Our data are from release 2 of the first two waves (2004 and 2006) of SHARE.

Our sample consists of the respondents of wave 1 (conducted in 2004) and the refreshment sample of wave 2 (conducted in 2006) residing in eleven European countries, namely, Sweden, Denmark, the Netherlands, Austria, Germany, France, Switzerland, Belgium, Greece, Spain, and Italy. The use of the refreshment sample from wave 2 is not only meant to increase the sample size but to help us to disentangle age effects

[5]For information on the data collection and methodology see Börsch-Supan et al. (2005) and Börsch-Supan et al. (2013).

from cohort effects.[6] We restrict our sample to those who were born in the period 1934-1954, so they were between the age of 20 and 50 in the period between 1954-2004 and they are between 50 and 70 years old at the point of the first wave data collection. We do not include individuals older than 70 in the sample since it could raise the selective mortality problem.[7] Our sample consists of 17,781 respondents from wave 1 and 5,099 respondents from wave 2.

1.2.2 Macroeconomic crises

While in general the period of 1954-2004 was characterized by post-war economic growth all over in Europe, still some serious macroeconomic recessions happened during this period. Among the most serious are the oil crisis 1973-1975, when the members of Organization of Arab Petroleum Exporting Countries proclaimed an oil embargo in response to the U.S. decision to re-supply the Israeli military forces. The price of oil quadrupled, followed by a stock market crash and economic recessions in many European countries. Another prominent crisis is the European Monetary System crisis of 1992-1993. The crisis was catalyzed by the reunification of Germany in 1990, an event unprecedented in history for the amalgamation of a large, rich economy with a smaller economy with a much lower standard of living. This event increased the interest rate in Germany, distorted the currency exchange balance between members of the European Monetary System and caused a severe economic crisis. Additionally there were a couple of more country specific crises in this period. We construct several measures that account for macroeconomic shocks a persons experiences during working life. For our main analysis we define crisis years based on the real GDP per capita data. In section 1.4.4 we also employ alternative definitions of crises based the mean GDP growth and mean unemployment rates experienced in a certain age windows.

GDP per capita is a widely used measure of macroeconomic conditions.[8] We take data on GDP per capita in 1990 PPP-adjusted dollars from the Agnus Maddison historical statistics on world population[9]. We use the data for the period 1954-2004 as this is the period when the SHARE respondents in our sample were age 20 to 50. Relative changes in GDP ($\frac{GDP_t - GDP_{t-1}}{GDP_{t-1}}$) are calculated. Table 1.1 provides some

[6]Each cohort has 2 more years of age in wave 2. For instance the 1954 cohort has 50 years old in 2004 and 52 in 2006.

[7]We comment further on the problem of selective mortality in the section on identification and the robustness checks.

[8]See, for example, Barro and Ursua (2012).

[9]Data can be downloaded on (http://www.ggdc.net/MADDISON/oriindex.htm)

critical values of the distribution of the relative change in GDP over the period from 1954 to 2004 for the countries included in our analysis.[10] Average GDP growth in this period was around 2.6%.

Table 1.1: Distribution of changes in real PPP-adjusted GDP in the period 1954-2004 in 11 European countries

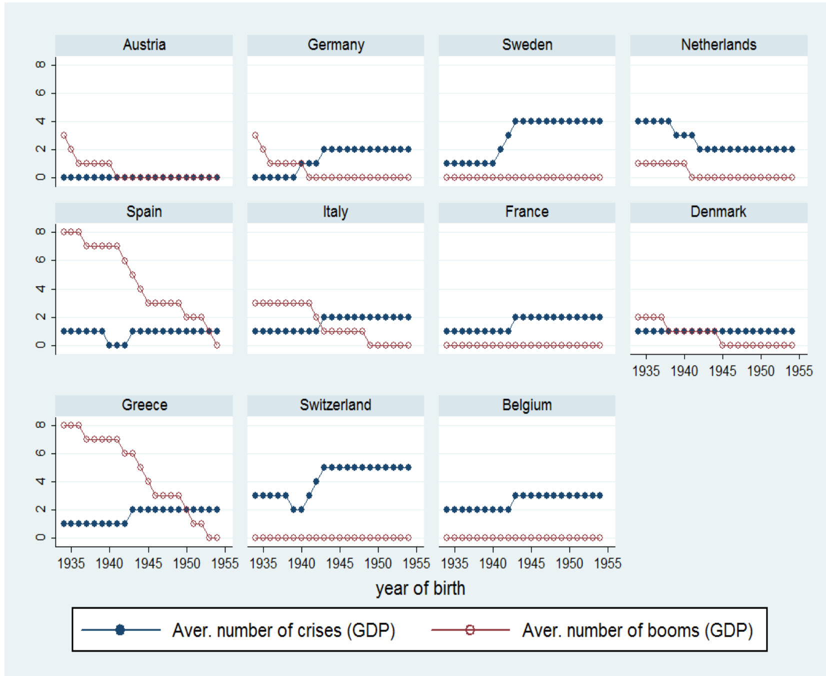
	1%	5%	10%	50%	90%	95%	99%
% Δ GDP	-2.75%	-0.95%	-0.09%	2.59%	5.71%	6.85%	10.10%

Note: The countries included are Austria, Belgium, Denmark, France, Germany, Greece, Italy, Netherlands, Spain, Sweden, and Switzerland.

We define several indicators for crises and booms in this period. Our standard measure is the following: we define the worst 5% of the years in terms of GDP dynamics as “crisis years”. The worst 5% of the country-years in terms of GDP growth correspond approximately to the years when GDP dropped by at least 0.95% compared to the previous year. Table A1 in the Appendix lists the years of crises by countries. Even though some crises were common for many countries, there is still a sufficient geographical and between-cohort variation in the number of crises. This can also be seen from Figure 1.1. We show the average number of crises per cohort of birth by country. The number of crises experienced between 20 and 50 varies from 0 (in Austria) to 5 (in Switzerland). In order to understand the sensitivity of our results to the crisis definition we change the threshold and redefine crises as worst decile of the country-years of the considered period (1954-2004). In this case years with a GDP drop higher than 0.09% are considered crises years. Moreover, in some specifications we analyze the effects of economic booms experienced between 20 and 50. “Boom years” are defined as the 5% best years in terms of GDP growth. In those years GDP per capita increased by at least 6.85%. The variation in the number of booms by country and cohort is also displayed in Figure 1.1. Finally, we split the period 20 to 50 into 5-year intervals to check if crises that hit individuals in certain critical periods within their working life have different impact. Thus, we construct indicators if individuals experienced a crisis at age 20–25, 26–30,..., 46–50 and include them as regressors simultaneously. Additionally, we create alternative period specific measures based on the mean GDP growth and the mean unemployment rate experienced in those periods.

[10] Germany is treated as one country in this data base even though it was separated into East and West between 1949 and 1990. We ran all our regressions dropping respondents from east Germany and also dropping Germany completely and our results are not changed. Therefore, we decided to keep the macro-economic data for Germany as it is despite the flaw.

Figure 1.1: GDP crises and booms over birth cohorts, by countries



1.2.3 Health measures and covariates

Our main dependent variable is a self rated health measure. Respondents are asked to evaluate their current health on a 5-point scale from "1-excellent" to "5-poor". Our indicator takes value 1 if the self-reported health was "1-excellent" or "2-very good" and 0 otherwise. Self-reported health status is among the most common measures used in public health surveys; it reflects various physical, emotional, social aspects of health and well-being and has been found to predict mortality (see, e.g. Idler and Benyamini, 1997; Jylha, 2009). About 37% of the respondents in our sample rate their health very good or better (see table A2) in the appendix).

In addition to that we provide analyses of a variety of other more objective health measures. The variable symptoms reports the number of symptoms that the respondent experienced in the last six months from a suggested list of 12 symptoms, such as fatigue, pain in the back, heart trouble, sleeping problems, etc. In addition to this we have a measure of the number of chronic conditions (including high blood pressure, heart attack, diabetes etc.), limitations in the instrumental activities of daily living (IADL), number of mobility limitations, number of depressive symptoms, a measure of cognitive functioning (i.e. memory), and grip strength. Grip strength reflects the overall muscle status of the respondent and has been linked to mortality in previous research (see, e.g., Gale et al., 2007). It is our most objective measure of health since the task is performed during the interview. In the appendix we list the exact definitions of all health variables and table A2 presents the summary statistics.[11]

Covariates. Besides the health and crisis measures we additionally use information on country of residence, gender, year of data collection, and birth year. Educational attainment is differentiated between low education—primary and lower secondary education (ISCED level 0-2), middle education—upper secondary and non-tertiary post secondary education (ISCED level 3-4), and high education—tertiary education (ISCED level 5-6). In a robustness check we are able to take account of childhood health and socio-economic circumstances, namely we have measures of height, self-reported health at age 10 and father's occupation. Table A2 also reports descriptive statistics for the covariates.

1.3 Empirical strategy

1.3.1 Empirical specification

The identification of the effect of interest is based on the deviations of the cohorts affected by a macro-economic shock from their country specific health trend. This empirical strategy has been already applied e.g. in the literature on the effects of macro-economic shocks around birth (Doblhammer et al., 2011). A very similar identification strategy can be also found in Giuliano and Spilimbergo (2013) who investigates how macroeconomic regional shocks during the "impressionable years" (18-25 years old) affect individual beliefs about success in life.

More specifically, we estimate the following equation:

$$Y_{isc} = \beta_0 + \beta_1 M_{sc} + \beta_2 X_{isc} + \delta_c + f(s, c) + u_{isc}$$

where Y_{isc} is the health outcome of individual i born in year s in country c; M_{sc} measures the macro-economic crises (or booms) experienced during ages 20 to 50; in our baseline specification this is a variable counting the number of crises and booms experienced. In the specifications presented in section 1.4.4 M_{sc} is a vector of crises experienced at the age of 20–25, 26–30.... 46–50. The crises measures used are an indicator if a GDP crises has been experienced in this period. In two alternative specifications we also use the average GDP growth and the mean unemployment rate experienced in the respective periods. X_{isc} contains other control variables, such as

[11] For an overview of all variables available in SHARE wave 1 and 2 see the questionnaires available on (www.share-project.org.)

gender and wave. The specification controls also for country fixed effects, δ_c, and country-specific polynomial trends (linear and quadratic) in birth cohort $f(s, c)$. We estimate the model separated by sex and educational level to evaluate the presence of heterogeneity in the effect of interest.

We present marginal effects derived after estimating probit models in all tables except table 5. Here linear models are estimated for the alternative health measures.

1.3.2 Identification issues

Since the identification of the effect of interest comes from variation between cohorts and countries the main concern is that we are picking up other country and cohort specific trends in health that are correlated with the number of crisis in the years. Looking at the trends in population health in Europe shows that health improved substantially over time, but the patterns were quite different among the countries. At the same time Figure 1.1 shows that the number of crises seems to have increased by cohort while the number of booms decreased for most of the counties considered. Thus, is is essential to control for country-specific trends in health. In this way, we are able to control for a wide variety of unobservable factors that might affect health, avoiding spurious correlations with our macro-economic indicator. In our main specification we include linear and quadratic country specific trends in health (which is the standard in this literature). However, in our robustness checks we play around with different trend specifications and show that results are robust to this variation.

Another way to check if we are confounding the effect of the crises with other cohort predetermined characteristics that affect health later in life is by looking at the effects of crises on childhood health. The idea is that crises experienced during working life should be uncorrelated with childhood health because of the time structure. In this sense our robustness check in section 1.5 should be interpreted as a sort of placebo test.

Another point of concern is selective mortality since our sample is composed of individuals age 50 and older. As already discussed in the introduction section, seminal work, e.g., from Ruhm, 2000 has shown that in the short run mortality shows pro-cyclical fluctuations. This might imply that at the time of observation those cohorts who experience worse macro-economic condition during their working years are positively selected by mortality. On the other hand, in the same papers Ruhm argues that these unfavorable health effects are partially or fully offset in the medium term if the economic growth is long-lasting. It means that in the presence of selective mortality the direction of the bias is not clear. We try to take this problem into account by including cohort-specific survival rates. Such a strategy should also solve the selection effect due to the second Word War (WWII). In this second case, the main concern is that our results might be driven by the fact that some of the cohorts involved in this study are born during the years of WWII. Previous research has shown that even today there are measurable effects of experiences during the war on health outcomes of those cohorts (Kesternich et al., 2014). As robustness check, we restrict our analysis only to the post-war cohort. The results from all the robustness checks so far described are shown in section 1.5 and they never cast doubt on our identification strategy.

1.4 Estimation results

1.4.1 Baseline results

In this section we present empirical evidence in order to reveal whether economic crises during working years have a causal effect on health outcomes later in life. In most tables we display marginal effects after running probit models where the dependent variable is being in "good health".[12] Our baseline specification includes controls for gender, wave, a full set of country dummies and country-specific linear and quadratic trends in age. We calculate robust standard errors clustered at the household level.

In the first column of table 1.2 we present our baseline results for self-reported health.[13] An increase in the number of severe macroeconomic crises significantly decreases the probability to report good health later in life by about 2.3%. Even though this might seem small the effect of one additional crisis year on self-reported health is equivalent to becoming almost 2 years older (See table A3). The baseline sample consists of all individuals who have lived in the country of current residence after age 20 and who have worked at least once in the period between 20 and 50. Including those who never worked reduces the effect of crisis on health (column 2). This seems plausible since individuals who never worked are less exposed to macroeconomic fluctuations. In our baseline regression we also exclude migrants who entered the country after age 20 because, first, migration can be related to macro-economic conditions both in the country of origin and the country of migration and, second,

[12]The exception are the results in Table 1.5. Here we use linear models for the other health variables.

[13]Marginal effects for the full set of controls are reported in table A3.

Table 1.2: Marginal effect of the number of crises on the probability to report good health

	(1) All	(2) All incl. immigrants	(3) All incl. immigr. and nvwrk
No. of crises	-0.0225**	-0.0189*	-0.0149
	(0.010)	(0.010)	(0.010)
Observations	22880	23981	25140
No. of crises	-0.0149	-0.0189*	-0.0225**
	(0.010)	(0.010)	(0.010)
Observations	25140	23981	22880

Note: Here and further: marginal effects are from probit regression model. Additional control variables are: gender, birth year, squared birth year, wave, country dummy and country-specific linear and quadratic trends in birth year. Standard errors clustered on household level are in parentheses. Number of stars denotes significance level: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Number of crises is the number of GDP crises experienced between age 20 and 50 where crisis year is defined as year in which real GDP in the country dropped by 1% or more with respect to the previous year. Excluded immigrants are those who migrated after the age of 20.

we cannot precisely estimate by which crises these persons were affected. Including the migrants (column 3) reduces the effect of crises on health further and renders it insignificant. Thus, for the following analyses we stick to our sample of resident population who did not migrate after age 20 and who worked at least once between age 20 and 50.

Table 1.3 presents the effect of the number of crises by gender and education levels. Separating the effects by education in column 1 reveals that the effect of macroeconomic crises on health is primarily driven by individuals with low levels of education. Among the low educated experiencing an additional crisis during working life decreases the probability to report good health by 4.2%. That is almost twice the effect in the overall sample. Also translated into years of aging: The effect of an additional crises year is equivalent to about 4 more years of age in terms of health among the low educated. The effect is substantially smaller and insignificant among those with higher educational levels.

The influence of macroeconomic shocks on health later in life is somewhat larger for women than for men (see column 2 and 3). However, analysing the interaction with education demonstrates that this is largely driven by differences in the level of education between men and women. Low educated men and women are about equally affected by experiencing crises during their working lives while the effects of crises for men and women with higher levels of education both are zero.

Table 1.3: Marginal effects of crises on the probability to be in a good health by levels of education and gender

	(1) All	(2) Males	(3) Females
All sample	-0.0225**	-0.0164	-0.0288**
	(0.010)	(0.014)	(0.014)
No. of obs.	22880	11286	11594
HS dropouts	-0.0424***	-0.0386**	-0.0444**
	(0.013)	(0.020)	(0.018)
No. of obs.	9778	4441	5337
HS graduates	-0.0149	-0.0040	-0.0235
	(0.019)	(0.026)	(0.026)
No. of obs.	7838	4011	3827
College graduates	0.0232	0.0394	-0.0099
	(0.025)	(0.034)	(0.039)
No. of obs.	4914	2667	2247

1.4.2 The effect of crises strength and booms

In this section we investigate whether the intensity of the macro-economic shocks are important. We already stress in the introduction that most of the literature does not distinguish severe macro-economic shocks from small business cycle fluctuations. One of the reasons why the effects of economic fluctuations on health are not consistent could be that crises are defined in different ways. For this reason we change the criterion to the worst decile of crises years—this corresponds to those years in which GDP dropped by at least 0.09%. Both the results of the regressions with the old and the new crises variable on self-reported health are provided in Table 1.4 for ease of comparison. When we relax the definition of the crisis years and consider also years with smaller drops in GDP the influence of macroeconomic crunches on health weakens. In the overall sample the effect is about half its original size and insignificant. For individuals with low education the effect is still significant even though it is reduced by almost half. These results suggest that only severe crises significantly harm population health in the long-run.

On the other hand, we would like to understand if booms, that is periods of exceptionally high economic growth, have the reverse effect on health. Booms are defined as the 5% country-years with the highest growth in GDP in our observation period. According to the distribution of the relative changes in GDP, we count a year as boom year if the relative GDP growth was at least 6.85% compared to the previous year (see table 1.1). The bottom raw of of table 1.4 shows the results for

Table 1.4: Marginal effect of number of crises of different strength on probability to be in good health

	All	HS dropouts	HS grad.	College grad.
No. of crises (5% worst years)	-0.0225**	-0.0424***	-0.0149	0.0232
	(0.010)	(0.013)	(0.019)	(0.025)
No. of crises (10% worst years)	-0.0107	-0.0230**	-0.0041	0.0135
	(0.009)	(0.012)	(0.015)	(0.020)
No. of booms (5% best years)	0.0318***	0.0300**	0.0286	0.0543*
	(0.011)	(0.014)	(0.022)	(0.029)
Observations	22880	9778	7838	4914

Note: A year in a given country is considered a crisis year on 5% criterion if in this year real GDP dropped by 1% or more with respect to the previous year. A year is considered a crisis on 10% criterion if in this year real GDP dropped by 0.09% or more that corresponds to the worst decile of country-years. Booms are defined as the years in which real GDP grew by 6.85% or more with respect to the previous year.

economic booms. We find that not only crises but also booms have a long-term impact on subjective health. Experiencing an additional economic boom increases the probability to report being in good health later in life by about 3.2% in the overall population. Effects are significant among the low educated and significant and slightly higher among those with high educational degrees. If we control for booms and busts simultaneously results are consistent.

Overall, our results suggest that people with a high level of education are less severely hit by economic crises and experience larger positive effects from economic booms. At the same time those with low levels of education are severely hit by economic crises and moderately profit from economics booms. We do not measure any effects of macro-economic fluctuations on health of those with medium levels of education later in life.

1.4.3 Other health outcomes

In addition to self-reported health SHARE provides a rich set of variables measuring health. In this section we provide evidence of the effects of crises on health outcomes later in life measuring health by the number of symptoms a respondent is suffering from, the number of chronic conditions, grip strength (which is measured during the interview), the number of mobility limitations, limitations in the instrumental activities of daily living (IADL), the number of depressive symptoms, and respondents' recall ability to recall (measured during the interview). The exact variable definitions are provided in the appendix. Linear models are estimated for all variables.

Overall our results presented in the table 1.5 appear to be consistent with previous findings using subjective health as an outcome. Respondents who experienced a larger number of crises during their potentially active labor market years report suffering from a larger number of symptoms and chronic conditions. They show lower grip strength, which is a measure of frailty and has been linked to mortality. They report a larger number of general mobility limitations but no higher probability to suffer from restrictions in the IADL. There are no effects on the probability to report depressive symptoms, but respondents who experienced severe crises have lower recall abilities later in life. As before the effects are stronger and significant among those with lower levels of educations. Exceptions are a strong effect of macro-economic shocks on mobility of those with a medium level of education and the adverse effect of crises on recall abilities of those with high levels of education.

Thus, overall we can conclude that the results of crises on health later in life are consistent for a variety of alternative subjective and objective health measures. Those with low education levels are particularly affected by dramatic macro-economic turbulence experienced over their life-course.

1.4.4 Timing of crises periods and alternative crisis indicators

In this section we would like to understand if crises experienced at different critical periods in the life cycle have different effects on health outcomes later in life. Such critical periods could for example be around labor market entry or exit were individuals might be particularly vulnerable to shocks.

For this purpose we construct different period specific crisis measures. First we employ a crisis measure similar to the crises measure used so far. Specifically, we split the period 20 to 50 in six 5-year intervals by creating six dummies that indicate whether the respondent experienced a crisis in a specific age window, from 20–25, to 46–50. We include all dummies simultaneously in the regression. Additionally, we measure the country-specific mean GDP growth rate and the mean unemployment in those intervals and add them as alternative measures of macro-economic conditions during these years.

Results of this exercise are presented in table 1.6 for the complete sample and separately for the low educated and the middle to high educated. Using the severe crisis

Table 1.5: Marginal effects of the number of crises on other health outcomes

	(1) All	(2) HS dropouts	(3) HS graduates	(4) College graduates
No. of symptoms	0.0523*	0.103**	-0.0404	0.0307
	(0.031)	(0.047)	(0.053)	(0.073)
	22879	9779	7837	4913
No. of chronic diseases	0.0592*	0.0919*	-0.0465	0.115
	(0.032)	(0.048)	(0.056)	(0.073)
	17772	7876	5977	3816
Grip strength	-0.401**	-0.529**	-0.327	-0.351
	(0.173)	(0.256)	(0.300)	(0.392)
	21245	9003	7302	4635
No. of mobility limitations	0.0357***	0.0448***	0.0093	0.0622***
	(0.010)	(0.015)	(0.018)	(0.022)
	22977	9796	7848	4923
IADL	0.0302	0.0594	-0.0160	0.0229
	(0.024)	(0.037)	(0.040)	(0.047)
	22977	9796	7848	4923
No. of depression symptoms	0.0060	0.0874	-0.140**	0.0238
	(0.041)	(0.063)	(0.068)	(0.089)
	22554	9615	7759	4851
Recall	-0.157**	-0.0869	-0.244**	-0.154
	(0.066)	(0.091)	(0.117)	(0.154)
	22638	9679	7778	4855

Standard errors in parentheses; number of observations below standard errors.

indicator shows that especially crises experienced between age 41 and 50 (namely the dummies 41–45 and 46–50) negatively affect health later in life. This effect is particularly severe for the low educated subsample. It is worth noticing that in particular for low educated also other age windows show negative and sizeable coefficients (e.g. 25–30; 36–40) but their standard errors are quite large. This might be due to the fact that controlling simultaneously for all age windows we are introducing some collinearity. More generally, with a cross-sectional approach is very complicated to establish whether the effect of crises experienced in the age window 41–50 had an impact on old age health because they were more recently experienced. Long panel data would be needed to answer this particular question.

Table 1.6: Marginal effect of 1) having at least 1 crisis in 5-years interval 2) mean GDP growth per capita in 5-year intervals 3) mean unemployment rate on probability to be in good health

	At least 1 GDP crisis			**Mean GDP growth**			**Mean unemployment**		
	All	HS dropouts	HS & college	All	HS dropouts	HS & college	All	HS dropouts	HS & college
20-25	0.0106	-0.0256	0.0412	0.0158*	0.0157	0.0138	-0.0364*	-0.0407	-0.0402
	(0.03)	(0.03)	(0.04)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)
26-30	0.0007	-0.0074	0.0115	0.0248**	0.0192	0.0299**	-0.0241*	-0.0043	-0.0295
	(0.02)	(0.03)	(0.03)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)
31-35	-0.0090	-0.0098	-0.0001	0.0200	0.0224	0.0158	-0.0191	-0.0150	-0.0132
	(0.02)	(0.03)	(0.03)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
36-40	-0.0247	-0.0118	-0.0335	0.0270	0.0370	0.0191	-0.0208	-0.0135	-0.0127
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
41-45	-0.0370***	-0.0400**	-0.0271	0.0275*	0.0574***	0.0024	-0.0159	-0.0092	-0.0097
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
46-50	-0.0323***	-0.0450***	-0.0165	0.0185*	0.0441***	-0.0057	-0.0027	-0.0045	0.0027
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N	22880	9778	12752	22880	9778	12752	21132	8747	12069

The importance of the age-windows 41–50 is confirmed also when using the mean GDP growth but not using mean unemployed. These alternative measures of the economic condition are of course correlated with our crises indicator. However, measuring the mean GDP or unemployed in a particular age window means analyzing the effect of the average economic condition in that age window and not specifically to the effect of severe negative macro-economic shocks. This might explain why these two alternative measures point out the importance of having favourable economic conditions during early adulthood (20–30). This result is consistent with the literature before mentioned (see, e.g., (Hessel and Avendano, 2013; Cutler et al., 2014)) which considers this period as critical because individuals make the transition from school to work. This is also confirmed by the fact that favourable economic conditions during the age windows 26–30 has positive effects only on the old age health of the high education group. Indeed, such age windows should corresponds with their early labor market years.

1.5 Robustness checks

In this section we address the potential concerns regarding our identification strategy.

Confounding trends. One possible concern with our results is that linear and quadratic trends might not be sufficient to describe the ageing dynamics of health. In this case our results might reflect some residual terms of ageing. For this reason we rerun the regressions first controlling for alternative trends and second we add cohort fixed effects. In the latter case we eliminate all possible within country-variation and capture only between-country variation. The results of the regressions only covering the high school drop outs are presented in table 1.7.

Column (1) presents results using country-specific linear trends in birth year; quadratic and cubic trends are added in columns (2) and (3). In column (4) cohort fixed effects and a linear trend are included; and finally in column (5) quadratic trends and fixed effects are taken into account. The bottom line in the table reports the variance inflation factor to show the degree of collinearity in the models.

Overall, our results do not seem overly sensitive to the inclusion of different linear, quadratic or cubic trends. However, the degree of collinearity increases substantially when using cubic trends. Adding cohort fixed effects reduces the size of the coefficients compared to the specifications without fixed effects. This, is not surprising since the identifying variation in the crises effect now only stems from variation between countries, standard errors increase too and the degree of collinearity is high.

Table 1.7: The effect of crises on health: controlling for country and cohort fixed effects

	Linear	Quad.	Cubic	FE+lin	FE+quad
	-0.0308 ***	-0.0424 ***	-0.0291 *	-0.0227	-0.0395 **
	(0.011)	(0.013)	(0.016)	(0.014)	(0.018)
N	9778	9778	9778	9778	9778
VIF	8.63	11.62	17.12	13.15	22.18

Childhood Outcomes. However, one might still be worried that we are only picking up cohort specific trends in health. Another way to strengthen our causal argument is to use outcomes that are established early in life, and thus should not be affected by crises experienced between age 20 and 50 and that are highly related to health outcomes later in life. Such outcomes are childhood health and socio-economic status (SES). There is a large literature showing that childhood health and SES are highly related to later-life health outcomes (see, e.g., Haas, 2008; Mazzonna, 2014). We use height, self-reported health at age 10 and fathers' occupation at age 10 as a dependent variables (see table 1.8). Height is measured in the standard SHARE questionnaire, while childhood health and father's occupation are only available for respondents to SHARELIFE. Thus, our samples in these specifications are smaller. Results in table 1.8 show that there is no relationship between the crises experienced between age 20–50 and health outcomes and SES measured earlier in life. Effects are close to zero and insignificant. We take this result as evidence that we are not picking up general cohort specific trends in health but truly measure the effects of severe macro-economic shocks on respondents' health.

Table 1.8: Marginal effect of number of crises on height, self-reported health at age 10 and fathers' occupation at age 10

	(1)	(2)	(3)
	Height	Childhood srh	Fathers' occupation
No. of crises	0.0032	0.0013	0.0112
	(0.137)	(0.002)	(0.007)
Observations	22659	13620	16507

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: the regression (1) is an OLS regression, regressions (2) and (3) are marginal effects of ordered probit models. Childhood srh and Fathers' occupation are known only for respondents who participated in wave 3. Childhood health is an ordinary variable taking values from 1 (excellent health) to 5 (poor health). Fathers' occupation at 10 is an ordered variable taking value from 1 most high-skilled to 3 most elementary.

World war II and selective mortality. Finally, two additional concerns with our results are that there could be a confounding influence of WWII, and that selective mortality could bias our estimation (see discussion in sections 1.3.2).

In column (2) of table 1.9 we replicate the analysis only using the sub-sample of individuals born after the end of the war (if the country of residence was affected by the war). The date is specific for each country: For Austria, Belgium, Germany, France, Italy, the Netherlands, and Denmark it is 1945; for Spain 1939; for Switzerland, Denmark, and Sweden we include all respondents, since these countries were not significantly affected by the war: Sweden and Switzerland were not under occupation, Denmark was under occupation, but experienced relatively less hardships than the other countries. The coefficient for the postwar subsample is almost identical to the coefficient for our baseline result. However, the effect becomes insignificant due to a substantially smaller sample.

In column (3) we address the issue of selective mortality and replicate our analysis adding the cohort- and country-specific survival rates as control variable. The results suggest that the selective mortality rate does not substantially affect our results. The coefficients in the regressions with or without survival rates are fairly similar to each other. Again significance is slightly smaller due to the smaller sample. We were only able to find detailed information on survival rates for 8 of our 11 countries.

Table 1.9: Effect of crises on health outcomes for all sample and post-war sub-sample

	All Sample	Postwar Sample	Sample with survival rate
No. of crises	-0.0225**	-0.0221	-0.0208*
	(0.010)	(0.014)	(0.011)
observations	22880	14306	16840

Note: Post-war sample are respondents from Austria, Germany, France, Italy, Netherlands, Greece, and Belgium born after 1945, from Spain after 1939, from Sweden, Switzerland, and Denmark - the whole sample. Survival rates for 8 countries (Sweden, Netherlands, Spain, Italy, France, Denmark, Switzerland and Belgium) were derived from the mortality rates. Data source: The Human Mortality Database, University of California, Berkley.

Overall, we are fairly confident that our results of crises experienced between age 20 and 50 are not driven by general cohort trends in health or other severe shocks like the experience of world war 2.

1.6 Conclusion

This paper investigates the causal effect of the number of macroeconomic crises experienced during prime working age (20 to 50) on different health outcomes in old age merging macro-economic data with micro-data from SHARE.

Our results indicate that crises measured as drop in GDP significantly affect the probability of being in good health, of having a larger number of symptoms, and reducing grip strength. We find some remarkable heterogeneity in the effects. The results are larger in magnitude and more significant for the low-educated sub-sample, while high-educated respondents seem to be mostly (positively) affected by positive macro-economic shocks. Effect size is similar to the effect size measured for cognitive function by Leist et al., 2013: One additional crises is equivalent to one additional year of cognitive decline after age 60.

Our results are not affected by the presence of war-related selective mortality or general trends in health. Finally, we show that the intensity of the crisis matters by showing that only the most severe macro-economic shocks had long lasting effects on the health of the affected cohort. Such a result is particularly relevant also in the light of the previous literature which never distinguished severe macroeconomic crises from business cycle fluctuations.

More generally, our study contributes to the existing literature on the long-term effect of the adverse conditions during the life-cycle on health outcomes later in life. To our knowledge, we are the first who considered the influence of adverse economic conditions during adulthood on old-age health. The fact that we find significant effects of macro-economic shocks, which affected the individuals later in life, suggests that not only early life circumstances (e.g. van den Berg et al. (2006)) have long lasting effects on health later in life. As a consequences there are other mechanisms, different from under-nutrition during the period of the body formation, through which macroeconomic turbulence affects health. Identifying such mechanisms is the task for future research.

More research is necessary to understand the mechanisms which are at work, to link macroeconomic cycles to individual career pathways, and ultimately to point to possible policy interventions.

Chapter 2

Non-monetary factor in the production of health: evidence from Russia Longitudinal Monitoring Survey

Acknowledgement: I thank the Russia Longitudinal Monitoring survey, RLMS-HSE, conducted by the National Research University Higher School of Economics and ZAO Demoscope together with Carolina Population Center, University of North Carolina at Chapel Hill and the Institute of Sociology RAS for making these data available.

Keywords: Grossman model; lifestyle choices; panel data; Hausman-Taylor model

JEL: I12, I18.

2.1 Introduction

20th century was the era of revolutionary improvements in health care technologies. Medicine has developed interventions that are very effective in treating acute health problems, especially infections diseases. The longevity increased substantially, which at the same time increased the number of years old people spend in frailty. The focus of medical care moved towards the treatment and prevention of non-infectious, chronic and degenerative diseases.

One of the property of these diseases, for example cardio-vascular conditions and cancers, is that they are to a large extent self-induced. In particular people who engage in unhealthy behavior like smoking or overeating have much higher probability of developing a chronic condition at a certain point of their lives. According to the British Heart Foundation's Cardiovascular Research Centre at the University of

Glasgow obese, middle-aged men have a 60% increased risk of dying from a heart attack than non-obese middle-aged men. The American Cancer Society reports 30% of all cancers to be caused by smoking.

Healthy and unhealthy behaviors are related to economic incentives, but market parameters are not the only factor of influence. For example, DellaVigna and Malmendier, 2006 study the statistics of fitness club attendances in New England and find that people who subscribe to the gym usually choose to pay a flat monthly or yearly rate rather than to pay for each visit, but then they go so rarely that ex-post they would have spent less money if they had chosen to pay for each visit. In this case behaving healthily was totally free (the fee was already paid), but still something prevented people from attending the gym regularly.

In this paper I modify the simplified version of Grossman model, assuming that health capital can be produced using medical care or following health rules, which induce disutility. The evolutionary origins of disutility of health behavior are discussed below.

I estimate the structural parameters of health production function using four waves of Russian Longitudinal Monitoring Survey (RLMS-HSE) and employing Hausman-Taylor panel data model which imposes minimum assumptions on the error term. I find that elasticity of health in respect to health rules is twice as large as elasticity in respect to medical care. I do not find the evidence of multiplicative effect of education on efficiency of health investment function usually assumed in theoretical models of health capital: according to the estimates education has no significant effect on health.

Using the estimates of input elasticities of health rules and medical expenses I do an exercise of cost-benefit analysis. I compare costs of an anti-smoking campaign "Quit and win" that took place in Finland in 1996-1997 and average medical expenses in OECD countries with potential gains in terms of health from quitting for those who actually quit an potential gains from medical care usage increase. I find that in terms of overall gains in health an anti-smoking campaign is more efficient than a cost-equivalent increase in medical care usage.

The paper is organized as follows. In the section 2.2 I discuss some considerations concerning disutility of health behavior coming from anthropological and evolutionary literature and reviews economic literature regarding health capital. Section 2.3 presents the modification of the simplified Grossman model augmented with health rules. Section 2.4 presents RLMS-HSE data set and provides variables' description and descriptive statistics. In section 2.5 the estimation strategy is discussed, while section 2.6 presents the results of the empirical analysis. Inferences about the structural parameters of the model recovered from the empirical estimates are presented in section 2.7; section 2.8 describes robustness checks and section 2.9 concludes.

2.2 Conceptual framework and related literature

In this paper I abstract from monetary and time costs of healthy and unhealthy behavior and assume that behaving healthily or abstaining from unhealthy behavior brings disutlility and at the same time produces health. The origins of disutlility of health behavior are related to the evolutionary history of human kind, in particular to the increase in brain size during evolution and the food shortage. For example, Leonard et al., 2007a explain that developing larger size of brains in humans compared to primates comes, in particular, to the cost of smaller muscle mass and larger fat "layer" compared to primates. Excess of fat allows to offset the high energy demand of the brain, since it is easier to "burn" than muscle tissue during the period of food shortage. As a consequence, development of food technologies, in particular availability of refined sugars, refined vegetable oils, salt, fatty meat and refined cereals together with propensity to accumulate a layer of fat in case of adverse environment conditions nowadays results in overeating, overweight and disbalanced diet in modern societies (Cordain et al., 2005). In developed countries, where the problem of food shortage is almost non-existent for the majority of population, individuals still have to adopt to the changed environment, in particular to learn not to create the protective fat reserves and to decrease the consumption of complex carbohydrates and refined sugar in favor of less caloric food containing more vitamins and minerals.

Another example is physical activity. The research shows that homo sapiens are unique among primates in their anatomical and metabolic capability for endurance running (Mattson, 2012). Endurance running allowed them to catch an animal or escape from the danger. In the era of technological progress and availability of the transport, a modern man has to learn to walk and do physical activity even if it is not necessary for procuring food supply.

I assume that this utility is the main reason of abstaining from health behavior, while the role of prices is negligible. At the end, many health behavior rules can be followed in a way that requires no money investment. For example, to perform a regular physical activity it is enough to go jogging in the park, and to eat healthily - to substitute junk food with cheap seasonal vegetables, legumes and fruits.

The model suggested in this paper follows the Grossman, 1972 framework of health production. In the Grossman model an individual inherits an initial stock of health capital. In each period the stock of health depreciates at some rate but can be increased by investment. Investment in health is a production function where the inputs are medical care and time. By choosing the quantities of inputs in the health production function, individuals determine their stock of health capital in each period.

Individuals invest in health for consumption benefits (health provides utility) and for production benefits (healthy individuals have less "sick days" and so earn more in a fixed period of time).

Analytical solutions for the Grossman model are usually based on the two sub-models: the pure investment model and the pure consumption model. In the pure investment model it is assumed that health does not provide utility per se, and people invest in health only because sick days decrease their earnings. In the pure consumption model, people invest in health because it enters their utility function, but does not affect the earnings.

The model predicts that in equilibrium the demand for health will positively depend on one's earnings and level of education and negatively depend on prices of medical care.

Further papers introduce modifications to the Grossman model, in particular embed in it lifestyle variables, which can be introduced into the model in different ways.

The simplest approach is to model lifestyle factors as affecting the depreciation rate at which the stock of health declines (Erbsland et al., 1995; Wagstaff, 1985). In this way health behavior is treated as an exogenous variable, together with the environmental variables such as air pollution and noise.

Another approach is discussed by Case and Deaton, 2005. In order to explain health disparities by occupational levels they develop a simplified version of Grossman model that I use as a starting point in my theory. Then they discuss several potential extensions of the model, in particular they suggest to embed in it the notion of "unhealthy consumption" such as smoking, consumption of junk food, sloth and engaging in cheap risk-taking activities. This type of consumption enters the felicity function bringing utility and is free in terms of price but has a cost in terms of higher rates of health deterioration.

Contoyannis and Jones, 2004 develop a structural model in which lifestyle factors are considered to make part of the consumption and enter the utility function with a positive sign. They estimate jointly reduced form equations for lifestyle and structural equation for health with Maximum Simulated Likelihood method, recognizing that lifestyle variables and health can be affected by common unobserved factors.

Di Novi, 2010 studies how the level of air pollution affects the amount of health investment. She considers healthy behavior to be an input into the health production function, together with preventive medical care, education and age, but does not introduce any relationship between health behavior and utility of individuals.

An important contribution to Grossman-type health capital models are works of Galama, 2011 and Van Kippersluis et al., 2013. In the theoretical work Galama, 2011 shows that the health investment function with decreasing returns to scale (DRTS) corresponds much better to the empirical data than the function with constant returns to scale. DRTS means that great improvements in health can be made with low level of health investment, whereas at high level of health investment, very expensive treatments often provide only a relatively small improvement in health.

In the subsequent work Van Kippersluis et al., 2013 the authors address an empirical observation that rich people on average engage much more often in healthy and much less in unhealthy behavior than poor people. They develop a theoretical model in which they distinguish healthy and unhealthy consumption. Both of them bring utility, but the unhealthy consumption speeds up the depreciation rate at which health stock declines, while the healthy one slows it down. The model implies that for richer people engaging in unhealthy consumption is costlier since for them the value of health is higher than for poor people.

2.3 A model of health rules and health production

The theoretical model described below is based on the paper by Case and Deaton, 2005 who build a simplified version of Grossman model of demand for health. Compared to the original Grossman model Case and Deaton, 2005 ignore the time aspect of health production, so that the health is produced just with medical care, and consider the length of the life-span to be exogenously defined. The novelty of my model is that it includes an additional input of health investment function: health rules (HR). Healthy rules are (infinitely many) rules of healthy behavior that everybody is assumed to know. Examples of such rules are: "do not smoke", "restrict the amount of sugar and simple carbohydrates in the diet", "do some physical activity regularly", etc. If an individual decides to follow such a rule, his health in the next period improves, but he experiences a disutility in the current period.

I assume variable HR to be continuous: the more rules and with higher intensity one follows, the larger is the value of HR.

It is assumed that following these rules is free in terms of time and money, in this case the "consumption" nature of some lifestyle variables, such as smoking, is neglected.

The individual maximizes the intertemporal utility function U:

$$U = \Sigma_{t=0}^{T} \frac{1}{(1+\rho)^t} u_t \tag{2.1}$$

where ρ is rate of time preferences and T is the exogenously defined length of the life span.

The utility function in each period is the function of consumption of generic good c, level of health H and number of health behavior rules followed HR:

$$u = u(H_t, c_t, HR_t) \tag{2.2}$$

The utility function is increasing in the consumption of good c_t ($\frac{\partial u(H_t,c_t,HR_t)}{\partial c_t} > 0$) and in the level of health ($\frac{\partial u(H_t,c_t,HR_t)}{\partial H_t} > 0$) and decreasing in the number of healthy rules ($\frac{\partial U(H_t,c_t,HR_t)}{\partial HR_t} < 0$). People experience utility from consumption and health, but disutility from obeying healthy rules.

The change in the stock of health capital over time is defined by gross investment in health I_t and age-specific depreciation rate of the existing stock δ_t.

$$\Delta H_{t+1} - H_t = I_t - \delta_t H_t \tag{2.3}$$

Health is assumed to be produced both purchasing medical care and following health rules. As in the previous models, health is increasing in education which is considered to be an exogenous time-invariant parameter.

$$I_t = I(m_t, HR_t; Educ) \tag{2.4}$$

The intertemporal budget constraint of the agent with a life span of T years and the initial level of wealth A_0 is defined by the equation:

$$\sum_{t=0}^{T} \frac{c_t}{(1+r)^t} + \sum_{t=0}^{T} \frac{P_t^M m_t}{(1+r)^t} = A_0 + \sum_{t=0}^{T} \frac{w_t h(H_t)}{(1+r)^t} \tag{2.5}$$

where r the rate of interest, w_t the wage rate, P_t^M the price of medical care and price of consumption c is normalized to 1. $h(H_t)$ is the amount of healthy days per period during which the individual works; healthy days is a positive function of health capital.

In every period, the total time available is divided into healthy time and sick time:

$$\Omega = h_t + TL_t \tag{2.6}$$

The solution of the problem is the level of factor inputs (m_t^*, c_t^*, HR_t^*) that maximizes the intertemporal utility function (2.2), subject to the budget constraint (2.5), the time constraint (2.6), the functional form of health investment function (2.4) and the dynamics of health capital (2.3). In the budget constraint it is useful to substitute in the equation above total costs of health investment $P^M m_t$ with the expression $\pi_t^I I_t$ where π_t^I is a marginal cost of health production.

The details of the solution of the optimization problem can be found in the B.A of Appendix. Below I provide directly three optimality conditions derived from the model.

Equations 2.7 to 2.9 define the optimal level of consumption, health capital and health rules:

$$\frac{\partial u}{\partial c_t} = \lambda \left(\frac{1+\rho}{1+r} \right)^t \tag{2.7}$$

$$\frac{\partial u}{\partial H_t} = \lambda \left(\frac{1+\rho}{1+r} \right)^t \left(\pi_{t-1}^I (1+r) - \pi_t^I (1-\delta_t) - w_t \frac{\partial h}{\partial H_t} \right) \tag{2.8}$$

$$\begin{aligned} &\frac{\partial u}{\partial HR_t} + \frac{1}{1+\rho} \frac{\partial U}{\partial H_{t+1}} \frac{\partial I}{\partial HR_t} = \lambda \left(\frac{1+\rho}{1+r} \right)^t * \\ &* \left(\frac{\pi_t^I (1+r) - \pi_{t+1}^I (1-\delta_t)}{(1+r)} \frac{\partial I}{\partial HR_t} - H_t \frac{\partial \pi^I}{\partial HR_t} (1+r) - \frac{\partial h}{\partial H_t} \frac{\partial I}{\partial HR_t} \frac{w_{t+1}}{(1+r)} \right) \end{aligned} \tag{2.9}$$

Equations 2.7 and 2.8 have a standard interpretation: marginal utility of consumption (health) is equal to net marginal cost of consumption (health stock). In case of health net cost is equal to the marginal cost of keeping an additional health stock minus gains in total wage obtained due to extra unit of health stock.

Equation 2.9 states that the optimal number of health rules will balance the net utility of HR (disutility of the health rules in period t plus utility of extra health derived from health rules in period t+1) with the net costs (cost of carrying extra units of health from period t to t+1 minus gains in wage in t+1 obtained thanks to following health rules).

As desribed in the Appendix B.A, to derive an equilibrium demand for health from optimality condition 2.8 a series of additional assumptions is made. First the Pure Investment sub-model is assumed, which implies that health does not have any direct utility and only indirect utility through wages: $\frac{u_t}{H_t} = 0$. Then the health investment function is assumed to be of a Cobb-Douglas type with decreasing returns to scale: $I_t(M_t, HR_t, Educ) = M_t^{\beta} HR_t^{\gamma} e^{\theta Educ}$. The health depreciation rate is assumed to depend on age and other socio-economic variables: $\delta_t = \delta_0 + \alpha_3 t + \alpha_4 Y_t$. The number of healthy days is assumed to be a positive function of stock of health: $h_t = \Omega - \alpha_1 H_t^{-\alpha_2}$. Applying these assumptions to the optimality condition 2.8 allows deriving the structural equation for the demand for health:

$$\mathrm{lnH}_t = \frac{1}{(\alpha_2+1)\beta}\left(\beta ln(\alpha_1\alpha_2) - \beta ln\delta_0 + \beta lnw_t - \beta lnP_t^M + \right. \\ \left. + \gamma lnHR_t + \theta Educ - \alpha_3\beta t - \alpha_4\beta Y_t\right) \quad (2.10)$$

The equation 2.10 suggests that level of health demanded in equilibrium positively depends on wage, educational level and number of health rules endogenously defined by equation 2.9 and negatively depends on price of medical care, age and other factors that speed up the health depreciation rate.

2.4 Data

The Russian Longitudinal Monitoring Survey (RLMS-HSE) is a survey run on a nationally representative sample of adults and children living in private households in Russia. It is carried out jointly by the Carolina Population Center at the University of North Carolina and the Higher School of Economics, Moscow. The first wave of the data was collected in 1992 and has been collected annually since then. In 1997 there was a complete change of the sample. The new sample contained 4,718 households and achieved more than 80% of response rate in the first wave. Since then, the initial sample was followed in each subsequent wave if respondents did not move away from the original dwelling and increased by refreshment samples.

The RLMS-HSE contains data on income, wealth, demographic characteristics, health status, medical care and health behavior. The data is collected on the individual as well as the household levels.

The analysis is run using 4 waves of the sample: wave 15 (2006) and waves 18-20 (2009-2011), since only in those waves the complete set of health behavior questions was asked. The sample is restricted to working individuals between the age 25 and 60 who did not change their educational level between waves and for whom the information on wages and health behavior is not missing. The final sample consists of 14303 observations over 8716 respondents.

2.4.1 Health and health behavior variables

Definitions and descriptive statistics of all the variables used in the analysis are provided in the table B1 of the Appendix B.B.

There are several health variables in the data set. The respondents are asked about self-reported health, the presence of chronic diseases in different organs, diabetes, high pressure, anemia and tuberculosis.

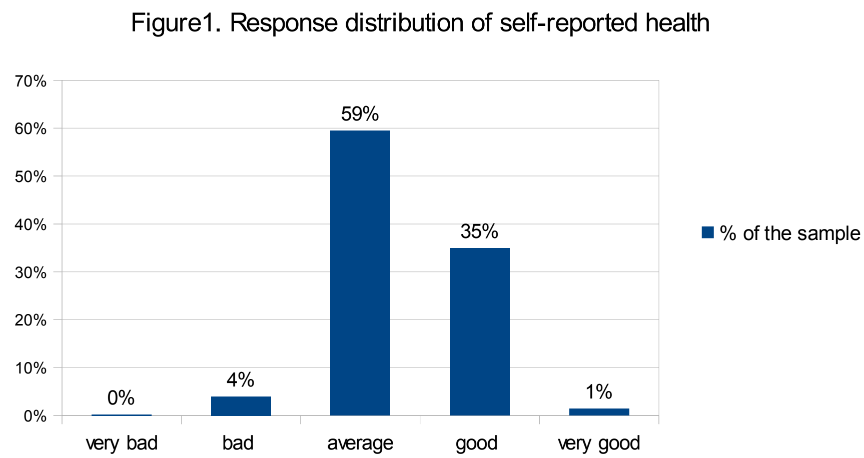
The main dependent variable used in the analysis is a binary indicator of self-reported health which takes value 1 if an individual rates his health as very good or good and 0 if average, bad or very bad. The use of a dummy instead of ordinal variable taking five possible values is justified by the fact that 94% of the sample respond either "average" or "good", while answers "bad", "very bad" and "very good" were chosen by less than 6%, so their presence almost does not add an additional information (see Figure 2.1). The distribution of answers differs a lot from Europe, even though the response categories used in the questionnaire are the same as in European studies like SHARE. For example, 15% of respondents of SHARE survey aged 50+ report to be in a very good health, and 60% either in very good or in a good health (Jürges et al., 2008), while for the current sample aged 25-60 these figures are 1% and 35%. From the other side 9.7% of SHARE respondents report to be in poor or very poor health, while only 4% of the Russian sample chooses these answers.

In the robustness checks some objective health variables are used: an aggregated health index and the dummy variables for chronic heart, stomach, spine and lung conditions, since those are more likely to be associated with unhealthy behavior.

For the reasons that are discussed below I mostly rely on the self-reported health measure of true health status.

Respondents of RLMS-HSE are asked a few questions about their health behavior. I mostly use 3 variables: current smoking (takes value 1 if currently smoking and 0 otherwise), drinking problem (takes value 1 if the female drinks more than 7 standard

Figure 2.1: Distribution of self-reported health responses



drinks a week or male respondent drinks more than 14 drinks a week and 0 otherwise[1]) and physical activity (1 if person reports to do at least some exercises and 0 if not). These variables are very similar to those used in "Alameda Seven" list of health habits which were shown to be associated with physical health status and mortality[2].

In order to obtain a single aggregated variable, the Principal Component Analysis is run over these variables and the first component is taken as a general measure of health behavior.

In the waves 19 and 20 an additional question related to health behavior was asked: "Do you manage to have meals regularly, at least 3 times a day?" This variable does not make part of the Alameda Seven factors. From one side having regular meals at least 3 times a day is likely to be positively correlated with having breakfast; from the other side it is likely to negatively correlate with avoiding snacks. Nevertheless some studies (for example, Jenkins et al., 1994) showed that increased frequency of meals reduces post-prandial insuline, glucose and cholesterol responses both in people

[1]the definition of "drinking problem" suggested by National Institute on Alcohol Abuse and Alcoholism

[2]"Alameda Seven" is a longitudinal study initiated in 1965 in Alameda Country, CA. A 9,5-year follow-up of 1965 Alameda cohort found that people who followed each of seven health habits: having never smoked, drinking less than 5 drinks in one sitting, sleeping 7-8 hours a night, exercising, maintaining normal weight for height, avoiding snacks and having breakfast tended to live longer than people who did not follow them, and that the effect is additive (see Schoenborn, 1986 for more details)

with non-insuline-dependent diabetes and without diabetes. Nutritionists also discuss so called "Staub-Traugott effect": individuals who have meals close to one another have better glucose response to the subsequent meals than individuals after the fast (Abraira and Lawrence, 1978). So the dummy variable for having regular meals is included as fourth variable in "Extended HR index" which serves as a robustness measure complementary to "HR index" based on 3 variables.

The composition of the first component of PCA run on three or four health behavior variables is reported in Table B4. The first component is able to explain 33% (for 4 variables) to 42% of the total variation in health behavior[3]

2.4.2 Other variables

The structural model includes, besides health rules, wages, price of medical care and years of education.

Monthly wages reported in each wave are adjusted to 2004 level of prices.

Retrieving price of medical care is a more challenging task. Medical care in state hospitals and polyclinics is officially free in Russia, but often people use medical services of private institutions, buy medicines, pay for additional tests not included in state coverage and give gifts to doctors.[4]

One of the advantages of the RLMS-HSE data set is that it contains detailed information about payments for medical services. The respondents are asked about the official and unofficial payments and gifts to the doctor during the last visit if it happened in the last 30 days prior to the interview; about the amount of money spent on medicines in the last month; about the amount of money spent in the hospital in the last 3 months. All the types of medical payments are summarized in the Table 2.1.

Payments for medical services are observed only for those who visited a doctor in the last month or stayed in a hospital in the last 3 months. In order to do so

[3]Since health as well as HR index enters the equation (2.10) in logarithmic form and variables aggregated Health Index and number of health rules can take zero values, taking logarithm of these variables would create a problem. To avoid it I use instead of logarithm the inverse hyperbolic sine (IHS) of HR index and of Aggregate health index. Inverse hyperbolic sine transforms variable x into $log(x + \sqrt{x^2 + 1})$. Burbidge et al., 1988 have shown that for large values of x IHS behaves like logarithm, but it has an advantage over logarithm: it maps 0 into 0. As for the variable goodhealth, since it is a dummy variable, I use it without applying any transformation.

[4]Gifts to doctors is a common practice in Russia. Since doctors officially do not receive payments from patients, gifts in the form of money or other goods allows expressing gratefulness to the doctor, counting on a better quality of medical services or receiving an access to the limited treatment. See the United Nations Research Institute for Social Development report "Commercialization of medical care and household behavior in transitional Russia" Blam and Kovalev (2003)for further details.

Table 2.1: All types of medical expenditures used to construct price of medical care

Nonemergency Expenses	Emergency expenses
Payment for the doctor visit - official	Payment for hospital stay - official
Payment for the doctor visit - unofficial	Payment for hospital stay - unofficial
Payment for the doctor visit - gifts	Payment for hospital stay - gifts
Payment for additional tests - official	
Payment for additional tests - unofficial	Expenditure on medicines in the hospital- official
Payment for additional tests - gifts	
Expenditure on medicines in the last month	Expenditure on medicines in the hospital - unofficial

each component of emergency and non-emergency medical expenses is regressed on the region dummies and the predicted values are assigned to all respondents on the basis of the region in which they reside. The reasoning behind this approach is that prices of medical services differ among regions on the basis of the level of development of medical system in the region, density of the hospital infrastructure and economic development of the region.

The results of regressions (presented in the Table B5) suggest that both price of emergency and non-emergency care and prices of medicines in some regions (in particular, in Moscow) are significantly higher than in others. Prices of medicines differ the most between regions, since expenses for medicines purchases are not reimbursed, they are likely to vary depending on income level in the region.

The educational system in Russia is quite similar to the European. After completing compulsory 9 years of school one receives a medium school diploma and can either graduate from the school or complete 2 more years to obtain the high school diploma. The medium school diploma makes a person eligible for entering professional courses of drivers, cooks etc., the professional technical schools ("PTU") or higher degree technical schools ("technicums"). The high school diploma provides eligibility to enter university ("institute").

In the survey respondents are asked about the highest educational degree obtained. Using ISCED-1997 codes developed by UNESCO, completed educational degrees are translated into years of schooling.[5]

I control for variables that are commonly used in studies about determinants of health: polynomial in age, gender, household size and composition. I also use wave dummies, allowing for non-linear trends in population health over time.

2.4.3 Descriptive analysis

Table 2.2 presents RLMS-HSE sample means for selected variables for sub-samples based on health behavior patterns. The first column gives means of the variables of interest for those who smoke, have drinking problem and do not do any physical exercise (HR index=0). The second column reports means for those who do not have a drinking problem, smoke and do not engage in physical activity. The last column are means for the subsample of respondents who do not smoke, do not have a drinking problem and perform some exercises (HR index=1).

Table 2.2: Selected variable means by sub-sample defined by patterns of a priori health behaviors

Variable	(1) Smoking drinking no physactiv N= 951	(2) Smoking not drinking no physactiv N=4,427	(3) No smoking not drinking physactiv N=1,952	(4) Diff. (2)-(1)	(5) Diff. (3)-(2)
Good health	0.324	0.386	0.398	0.062***	0.012
Aggr. HI	0.902	0.915	0.893	0.013***	-0.022***
Ln(real wage)	9.040	9.045	9.184	0.005	0.139***
Years of educ	12.165	12.294	14.361	0.129	2.067***
Male	0.773	0.725	0.379	-0.048***	-0.346***
Age	40.150	39.415	40.029	-0.735**	0.614**
Hh size	2.811	2.786	2.608	-0.025	-0.178***
Married	0.631	0.676	0.674	0.045***	-0.002

Stars (* $p<0.10$, ** $p<0.05$, *** $p<0.01$) denote the significance level of the differences in means between columns (2) and (1) and between columns (3) and (2)

Examination of Table 2.2 suggests a number of interesing observations. First observation is related to health measures. There is a significant difference in self-reported good health between columns (1) and (2), i.e. between those who break all the health rules and those who follow only the rule of not drinking. At the same time the difference between columns (2) and (3), i.e. between those who follow just the

[5]http://www.uis.unesco.org/Education/ISCEDMappings/Pages/default.aspx

rule of non-drinking and those who follow all the health rules is not significant. This observation gives the idea that drinking is the main factor affecting self-reported good health, while smoking and physical activity have less impact. If instead one makes the same comparisons regarding aggregate health index which summarizes all the objective health conditions[6], it increases significantly from columns (1) to (2) but then significantly drops from columns (2) to (3). As it will be shown later, this counterintuitive result might be due to the differences in medical care utilization between the subgroups.

The examination of the rest of the Table 2.2 brings to the conclusion that there is a clear health behavior - education gradient, especially in smoking and physical activity; that females follow on average more health rules than men, especially the rules concerning smoking and physical activity (the proportion of male drops from 72.5% in column (2) to 37.9% in column (3)). There is no significant difference between those who do not follow any health rules and those who follow the rule of not drinking in wages, while those who follow all the rules earn significantly more. Finally, there is some negative relationship between health behavior and household size and positive with being married.

2.5 Estimation strategy

The estimable counterpart of the structural equation 2.10 is:

$$lnH_{it} = a_1 + a_2 lnw_{it} - a_2 lnP_t^M + a_3 lnHR_{it} + a_4 Educ_i - a_5 \beta t - a_6 \beta Y_{it} + \epsilon_{it} \quad (2.11)$$

where $a_1 = \frac{ln(\alpha_1 \alpha_2}{(\alpha_2+1)} - \frac{ln\delta_0}{(\alpha_2+1)}$, etc; the subscript i denotes i-th individual and the subscript t the th-th year.

All the variables in 2.11 are of period t, so I do not have to take care of the dynamic aspects of the model. Estimating this equation with pooled OLS regression would create two kinds of concerns. One is that there might be omitted variables contained in the error term ϵ_t and potentially correlated with the explanatory variables. The second problem is endogeneity. As equation 2.9 shows, health behavior depends on the current state of health, while equation 2.11 contains health on the right-side and health behavior on the left-side.

[6]The technique of calculation of aggregate health index are discussed in the section 2.8

This requires the use of an econometric strategy which deals with both concerns, so I turn to the panel dimension of the data. Consider a model written as follows:

$$y_{it} = X_{it}\beta + Z_i\eta + \alpha_i + u_{it} \tag{2.12}$$

where i=1,...,N and t=1,...,T. The X_{it} are time-varying variables, Z_i are time-invariant variables, α_i is unobserved individual effects IID(0, δ_α^2), u_{it} is an idiosyncratic shock IID(0, σ_u^2).

The Random-effects model deals with omitted variables problem under the assumption that they are uncorrelated with explanatory variables: $cov(\alpha_i, X_{it}) = 0$ and $cov(\alpha_i, Z_i) = 0$. This strong assumption is not likely to hold in the equation (11). In particular, studies starting from Fuchs (1982) suggest that time preferences affect the level of health. At the same time equation 2.9 suggests that the optimal number $HR*$ will depend on time discounting rate ρ, so $cov(\rho, HR_t) \neq 0$.

A model that requires more parsimonious assumptions on unobservables is a Random-effects model with Mundlak, 1978 correction. Mundlak showed that under the assumption that time-invariant explanatory variables are uncorrelated with unobservables (cov$(\alpha_i, Z_i) = 0$), running the Random-effects model adding means of all time-varying variables as covariates is equivalent to running Fixed-effects model. Again, even this assumption might not hold: the time-invariant variable "years of education" is also likely to be correlated with time preferences.

The model that allows explanatory variables to correlate freely with the unobservables, is the Fixed-effects model. Its serious drawback is that it does not allow to estimate time-invariant variables. Estimating returns on education is an essential part of this study, since one of the crucial questions is how much is the contribution of health rules in respect to education, and if there is spillover effect of education on following health rules. In addition, Fixed-effects model does not solve the problem of endogenous explanatory variables.

A solution that I opt for is Hausman-Taylor model (Hausman and Taylor, 1981). It is an instrumental variables estimation which deals with endogeneity and unobserved heterogeneity under parsimonious assumptions on the correlation of unobservables with explanatory variables. It does not require external instruments for the endogenous variables, since it uses means and deviations from the means of both exogenous and endogenous variables as instruments.

The idea behind this model is to split k time-varying variables X into subsets of k_1 exogenous variables [X1] and k_2 endogenous variables [X2], and g time-invariant variables Z into g_1 exogenous variables [Z1] and g_2 endogenous variables [Z2]. X1 and

Z1 are not correlated with α_i and u_{it} while X2 and Z2 are correlated with α_i but not u_{it}.

Further define an orthogonal projection P_v which projects a T-dimentional vector of ones into $(I_N \bigotimes \frac{1}{T} i_T i_T^{')}$ and the one perpendicular to it $Q_V = I_{NT} - P_v$. Note that for any time-invariant variable Z $Q_v Z = 0$. Knowing this, it it possible to pre-multiply the equation 2.12 by Q_v to obtain unbiased and consistent estimates of time-varying variables: $\hat{\beta_w} = (\tilde{X_{it}}' \tilde{X_{it}}^{-1}) \tilde{X_{it}}' \tilde{Y_{it}}$, where $\tilde{X_{it}} = \bar{X} - X_{it}$.

It is also possible to obtain unbiased and consistent estimates of time-invariant variables if the number of exogenous time-varying variables k_1 is larger than the number of endogenous time-invariant variables g_2. If this condition holds, taking the residuals from within-groups regression $\hat{d_{it}} = Y_{it} - \hat{\beta_w} X_{it}$ and estimating the equation $\hat{d_{it}} = Z_i \gamma + \alpha_i + \epsilon_{it}$ with the 2SLS estimator where variables from X1 serve as instruments for Z2 leads to unbiased estimates of γ.

When it comes to estimating the equation 2.11, I assume that the set of time-varying exogenous variables X1 consists of [polynomial of age, price of medical care, married, household size, wave dummies], the set of time-varying endogenous variables X2 consists of [wage, health behavior]. Z1 is [male] and Z2 is [education]. The division into exogenous and endogenous variables is formally untestable, nevertheless it is reasonable to assume that price of medical care which is predicted based only on regional variation, polynomial of age and family circumstances do not have common unobservable factors with health status. Since Russia is the country with highest divorce rate and one of the highest marriages rates, and also the household size can vary a lot, depending of parents living together or separately from the adult children, the marital status and household size can be considered as an exogenously varying.

I program an ad-hoc Hausman-Taylor estimator in order to be able to put inside a linear constraint and to correct for unbalanced panel. To assure the correctness of standard errors, I bootstrap them.

To sum up, the advantages of using the Hausman-Taylor model are:

- it allows obtaining estimates of time-invariant endogenous variables under the minimum assumptions on the error term

- using an ad-hoc Hausman-Taylor estimator allows imposing a linear constraint on the equality of coefficients at two variables (as discussed further) while bootstrapping procedure assures the accuracy of standard errors.

- an ad-hoc estimator deals with the problem of an unbalanced panel.

2.6 Results

Table 2.3 presents the coefficient estimates of the econometric model 2.11 estimated with Hausman-Taylor with health measured as a dummy for being in a good health. The first column reports results of the regression when no additional linear constraint is imposed. The column (2) presents the results of the same analysis with an additional linear constraint imposed. This constraint is the equality of coefficients of wage and price of medical care predicted by the equation 2.11. In the column (2) these coefficients differ significantly (χ^2=16.89). This is likely to happen due to a very small within-subject variation of the price of medical care: only 1% of the variation is a within one. So from the column (2) in all the further regressions I impose the linear constraint on the equality of two coefficients.

The columns (3) an (4) present the results of the same model run on the older subsamples (35-60 and 45-60 years old). The idea behind running the analysis on older age subsamples is that young bodies might resist better to adverse effects of own unhealthy behavior, so the effect of health behavior might not show up in their self-reported health. In fact, the examination of columns (3) and (4) suggests that the coefficient at health behavior for the subgroup aged 35-60 (0.14) is almost twice as large compared to sub-sample 25-60 (0.086). The coefficients for the group aged 45-60 are very close to ones of 35-60. In all the next tables the regressions are run on the subsample aged 35-60.

Analysing the estimates from 2.3 suggests a number of observations. All the variables have signs predicted by the equation 2.11: HR index and wage positively affect demand for health while price of medical care and age have a negative effect. Being a male positively affects probability to be in a good health, which is likely to be related to the gender differences in self-reporting. For the sub-group of 35-60 being married negatively affects health while the household size has a positive effect. It is notable that years of education have no significant effect on health, meaning that all the returns on education come via differences in wages and in health behavior. While in Grossman model it is assumed that education has a multiplicative effect on health by increasing the efficiency of health investment, and in the European-based studies (for example, Erbsland et al., 1995) it is found that after controlling for wage and health behavior education per se has a positive effect on health, the insignificant results in the Russian data might be related to the country-specific circumstances. In the robustness check section I investigate further the link between health, health behavior and education.

Table 2.3: Hausman-Taylor coefficient estimates for structural variables of demand for health equation with linear constraint

	(1) 25-60 y.o. no constraint.	(2) 25-60 y.o. cons.	(3) 35-60 y.o. cons.	(4) 45-60 y.o. cons.
Ihs(HR index)	0.078	0.086*	0.14**	0.13*
	(0.05)	(0.05)	(0.06)	(0.08)
Ln(real wage)	0.041***	0.081***	0.068***	0.071***
	(0.01)	(0.01)	(0.01)	(0.02)
Ln(P^M)	-0.083***	-0.081***	-0.068***	-0.071***
	(0.01)	(0.01)	(0.01)	(0.02)
Age	-0.022***	-0.024***	-0.044***	-0.083**
	(0.00)	(0.00)	(0.01)	(0.03)
Years of educ.	-0.004	-0.010	-0.002	-0.001
	(0.01)	(0.01)	(0.01)	(0.01)
Male	0.12***	0.098***	0.12***	0.10***
	(0.01)	(0.01)	(0.02)	(0.02)
Married	-0.012	-0.009	-0.031***	-0.030**
	(0.01)	(0.01)	(0.01)	(0.02)
Hh size	0.005	0.004	0.010**	0.011*
	(0.00)	(0.00)	(0.00)	(0.01)
N	14303	14303	9486	5288

Here and in all the further regressions: bootstrapped standard errors in parentheses;
* p<0.10, ** p<0.05, *** p<0.01

2.7 Recovering structural estimates

In the previous section I estimated the model:

$lnH_{it} = a_1 + a_2 lnw_{it} - a_2 lnP_t^M + a_3 lnHR_{it} + a_4 Educ_i - a_5 \beta t - a_6 \beta Y_{it} + \epsilon_{it}$

which is an econometric counterpart of the equation

$$\text{lnH}_t = \frac{1}{(\alpha_2+1)\beta}\left(\beta ln(\alpha_1\alpha_2) - \beta ln\delta_0 + \beta lnw_t - \beta lnP_t^M + \gamma lnHR_t + \right. \\ \left. + \theta Educ - \alpha_3\beta t - \alpha_4\beta Y_t\right) \quad (2.13)$$

Assuming that Hausman-Taylor model run on the sub-sample of 35-60 years old is the correct specification, the following relations between the structural parameters can be found from the results in the column (3) of the Table 2.3:

$\frac{\beta_{log(wage)}}{\beta_{log(HR)}} = \frac{a_2}{a_3} = \frac{\beta}{\gamma} = \frac{0.068}{0.14} = 0.486$

$\frac{\beta_{educ}}{\beta_{log(HR)}} = \frac{a_4}{a_3} = \frac{\theta}{\gamma} = \frac{0.00}{0.14} \approx 0$

The health investment function $I(M_t, HR_t, X) = M^\beta HR^\gamma e^{\theta Educ}$ can be re-written as

$$I(M_t, HR_t) = M_t^\beta HR_{it}^{2.06\beta} \quad (2.14)$$

The equation 2.14 implies that input elasticities of health production function are:

$\epsilon_{H,M} = \epsilon_{I,M} = \beta$

$\epsilon_{H,HR} = \epsilon_{I,HR} = 2.06\beta$

$\epsilon_{H,Educ} = \epsilon_{I,Educ} = 0$

In other words, a 1% increase in health rules index results in a 14% increase in the probability of being in good health (marginal effect of HR_t on good health from the Table 2.3), while 1% increase in wages results in a 6.8% increase. For a person who does not drink and performs some physical activity, giving up smoking would increase the HR index by 45%, so for this person a 1% increase in health rules is equivalent to abstaining from smoking 1 of every 45 cigarettes.

The equation 2.14 suggests that there are only two ways to increase population's health: increase the consumption of medical care (for example by decreasing prices of medical services) or convince people to follow more health rules.

The following example can illustrate the magnitude of numbers when choosing between alternative policies. The government of Finland in 1996-1997 launched an anti-smoking campaign called "Quit and win" targeted at the adult smoking population of Finland (see Korhonen et al., 1999 for an overview). In the following-up surveys it was estimated that out of approximately 740,000 of adult smoking population 0.1% permanently stopped smoking as a result of "Quit and Win" program. The costs of the program in 1996-1997 were estimated around 130,000 dollars, that in prices of 2009 would be 178,000 dollars[7]. Giving up smoking for a person who does not have a drinking problem and does some physical activity means an increase in HR index from 0.69 to 1.00, i.e. a 45% increase. According to my estimates, the elasticity of health behavior in respect to health is twice as large as the elasticity of medical care, so in order to reach an improvement in health of 740 smokers equivalent to the effect produced by giving up smoking the policy maker would have to increase their medical care expenditures by 90%. In 2009 average medical expenses in Finland were approximately equal to the median of the OECD countries which was equal to more than 3200 dollars.[8] A 90% increase in medical expenses for 740 people with average medical expenses of 3200 euros would sum up to more than 2 mln dollars which is much more than what was spent on the anti-smoking campaign.

2.8 Robustness checks

First I question the correct choice of the econometric model to estimate the data. To compare the magnitude of coefficients obtained with Hausman-Taylor model with those from other models I run OLS (column (1) in the Table 2.4), Random-effects (RE) (column (2)), Mundlak (column (3)) and Fixed-Effects models (FE) (column (5)) and compare them with the estimates of the unconstrained Hausman-Taylor model (column (4)).

Coefficient estimates of Pooled OLS, Random-effects and Mundlak, as well as of Hausman-Taylor, have the signs predicted by theory: health level positively depends on the HR index, wages and years of education and negatively depends on level of prices of medical care. The coefficient on health behavior varies from 0.12 in OLS and RE models to 0.06 in FE models, with Hausman-Taylor coefficient of 0.08 lie in between. Mundlak and FE models give small and insignificant coefficients of wage while a coefficient from HT model is closer to the one from RE. Instead the FE coefficient on price of medical care is huge (-4.80), which is related to a low within-subjects variation in this variable.

[7] using historical Consumer Price index values

[8] http://www.oecd.org/els/health-systems/Briefing-Note-FINLAND-2013.pdf

The diagnostics tests comparing different models suggest that Random-effects model is preferred to Pooled OLS model (Breush-Pagan LM test χ^2 is equal to 1203.36, p-value=0.000); that Fixed-effects model are preferred to Random-effects and Random-effects with Mundlak correction (Hausman test χ^2 statistics are equal to 3238.83 and 3152.21, p-values in both cases are equal to 0.000). I am not able to run a Hausman test of Hausman-Taylor model versus FE model, because Hausman-Taylor estimator is programmed manually as 2SLS procedure.

Table 2.4: Coefficient estimates for pooled OLS, RE, RE with Mundlak correction, Hausman-Taylor and FE models

	(1) Pooled OLS	(2) RE	(3) Mundlak	(4) Hausman-Taylor	(5) FE
Ihs(HR index)	0.12***	0.12***	0.092***	0.078*	0.064**
	(0.02)	(0.02)	(0.03)	(0.05)	(0.03)
Ln(real wage)	0.063***	0.056***	0.015	0.041***	0.008
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ln(P^M)	-0.096***	-0.12***	-0.13***	-0.083***	-4.80***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.08)
Years of educ.	0.004**	0.004**	0.004*	-0.004	
	(0.00)	(0.00)	(0.00)	(0.01)	
Male	0.12***	0.12***	0.11***	0.12***	
	(0.01)	(0.01)	(0.01)	(0.01)	
Age	-0.023***	-0.020***	0.019	-0.022***	0.004
	(0.00)	(0.00)	(0.02)	(0.00)	(0.02)
Married	-0.018*	-0.021**	-0.022	-0.012	-0.005
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Hhd. size	0.007*	0.009**	0.005	0.005	-0.002
	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
N	14302	14302	14302	14303	14302

The coefficient estimates of the structural variables estimated with the Hausman-Taylor model are quantitatively in between of those estimated with Mundlak estimator and FE, which is an evidence in favour of this choice of the model.

Next I investigate how the definition of health rules index affects the magnitude of it's effect on health. Table 2.5 presents the results of the analysis with alternative measures of health behavior. First, health behavior index measured as a first component of Principal Component Analysis run on four variables (including having regular meals) (column (1)) and then as dummy variables for each factor of health behavior separately (columns (2) to (4)). The results of the table suggest that health behavior index built on four variables has a larger impact on health than the baseline index. Also it looks like the drinking problem has the strongest effect on the probability of being in good health: when health behavior is measured as single indicators, only the drinking problem has a significant impact on health, while smoking and physical activity have the expected sign but are not significant. Drinking problem in Russia is well-known to be one of the main causes of mortality and the main source of almost 10-years gap between male and female life expectancy (61 for males and 74 for females in 2007). Leon et al., 2009 estimate that 31-43% of deaths among working-age men are related to alcohol consumption. Alcohol excessive consumption is found to be negatively related to educational level and positively related to unemployment, so my estimates based only on working sub-sample might be downward biased.

Table 2.5: Hausman-Taylor coefficient estimates using alternative measures of health behavior with linear constraint, subsample of 35-60 y.o. respondents

	(1)	(2)	(3)	(4)
Ihs(HR index$_4$)	0.23**			
	(0.10)			
Drinking problem		-0.063**		
		(0.03)		
Physical activity			0.004	
			(0.02)	
Current smoking				-0.070
				(0.05)
Ln(real wage)	0.093***	0.069***	0.070***	0.066***
	(0.02)	(0.01)	(0.01)	(0.01)
Ln (P^M)	-0.093***	-0.069***	-0.070***	-0.066***
	(0.02)	(0.01)	(0.01)	(0.01)
Age	-0.051***	-0.044***	-0.045***	-0.045***
	(0.01)	(0.01)	(0.01)	(0.01)
Years of educ.	0.021	-0.001	-0.002	-0.002
	(0.02)	(0.01)	(0.01)	(0.01)
N	5217	9486	9486	9486

Further I employ alternative health outcomes instead of self-reported good health to see if the previous results hold also with objective health conditions. Several alternative measures of health are used: the aggregate health index and dummies for having chronic conditions of heart, stomach, spine and lungs.

The aggregated health index is calculated following the Jürges, 2007, except that I use the dummy for good health instead of ordinal variable, because of quasi-binary distribution of self-reported health variable in the RLMS-HSE data set. I run the probit model regressing good health on all the objective health conditions, take the predicted part and normalize it to take values between 0 (for the worst observed state of health) and 1 (for the best observed health). The regression coefficients of the probit regression are provided in the table B3 of the Appendix B.B. Note that this analysis is run on the full sample of respondents of waves 15 and 18-20 aged between 25 and 60 who do not report changes in educational levels between waves, while the sample for main regressions is constrained to those for whom the information on wages and on health behavior is not missing (see table B2 for details on sample selection).

All the objective health conditions except heart attack have negative and significant effect on good health. Obtained in this way health index represents the part of variation in self-reported health explained by existing objective conditions. I use it as an alternative measure of health capital and run the baseline Hausman-Taylor regression with the HR index based on three variables and aggregate health index as a dependent variable. The results are presented in the Table 2.6.

All the coefficients at health behavior index in Table 2.6 have an expected sign (except the cronic lungs disease) but are insignificant. The other variables of the estimated equation 2.11, wages and price of medical care, are significant with the correct sign for chronic spine and stomach conditions, but are not significant for the aggregate health index, heart and lungs diseases. Also Table 2.6 contains a counter-intuitive result: education has a negative effect on aggregate health index and positive effects on probability to have chronic heart or chronic spine condition.

The results of Table 2.6, even tough they partially contradict the model predictions and the vast empirical evidence, should be considered carefully. The difference between self-reported health and objective health conditions included in the aggregate health index and in the separate dummy variables is at least twofold. The first issue with chronic diseases, tuberculosis, anaemia etc. is that they should be diagnosed to be reported by respondents. The second issue that is related also to the first one is so called "latency period", a period of time when adverse health behavior already

Table 2.6: Hausman-Taylor coefficient estimates with aggregate health index and chronic conditions as dependent variables, subsample of 35-60 y.o. respondents

	(1) Aggr. health index	(2) Chronic heart condition	(3) Chronic stomach condition	(4) Chronic spine condition	(5) Chronic lungs condition
Ihs(HR index_3)	0.013	-0.021	-0.034	-0.031	0.013
	(0.01)	(0.03)	(0.05)	(0.05)	(0.03)
Ln(real wage)	0.003	-0.002	-0.033***	-0.031***	0.004
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Ln(P^M)	-0.003	0.002	0.033***	0.031***	-0.004
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Years of educ.	-0.005**	0.012*	0.007	0.017**	-0.003
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)
N	8840	9458	9446	9454	9473

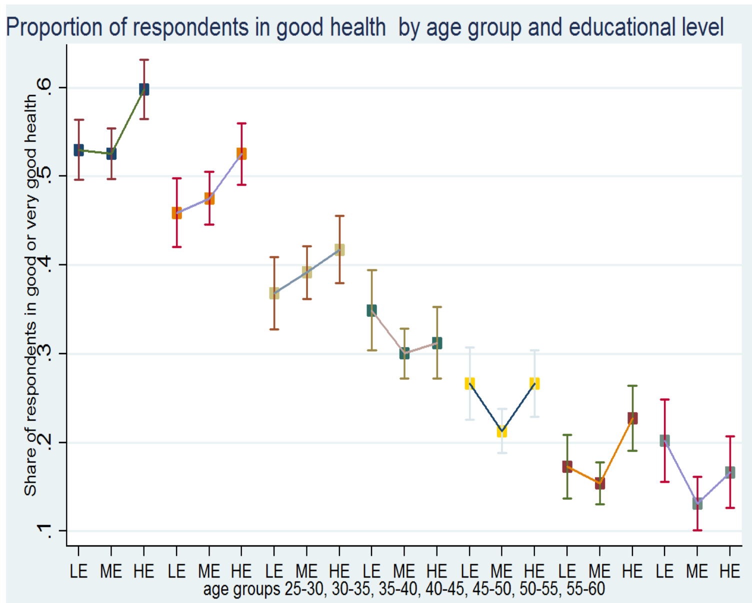
affected "true" health, but the changes in the "true" health did not yet show up in the measurable health outcomes. For example epidemiologists estimate the "latency period" between start of smoking and onset of the lung cancer to be about 20 years[9]. Since I have only a few waves of observations and high drop-out rates, it is likely that the analysis captures a short-term effect of health behavior on health which is best reflected in changes in self-reported health.

Insignificant or negative coefficient for education is another result contradicting the theoretical model. To have a better idea about the relation between education and health in the data set, I plot self-reported health (Figure 2.2) and aggregate health index (Figure 2.3) for sub-samples of low-, medium- and high-educated respondents. Low-educated are defined as those with 10 or less years of education (equivalent to high school dropouts), medium-educated are those with 12 to 14 years of education (equivalent to technical, pedagogical or artistic schools graduates) while high-educated are those with more than 14 years of education (university graduates or PhDs).

Figure 2.2 suggests that while there is some evidence of education-health gradient for younger cohorts (age group 25-40), there is no such evidence for older cohorts. The possible explanation for this result is that younger generations grew up in relatively stable socio-economic conditions, while older respondents received their education

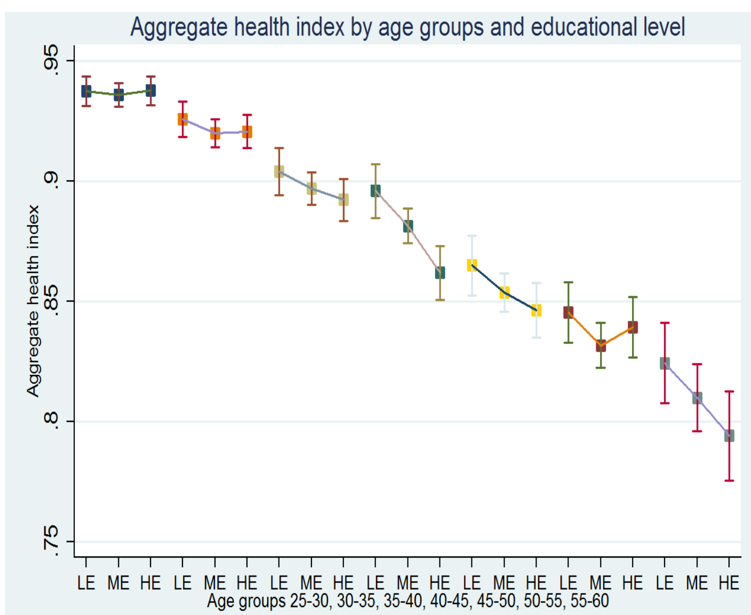
[9]http://www.epa.gov/ncea/ets/pdfs/etsch4.pdf

Figure 2.2: Proportion of respondents in good health by 5-years age groups for low-medium and high-educated subsamples



]

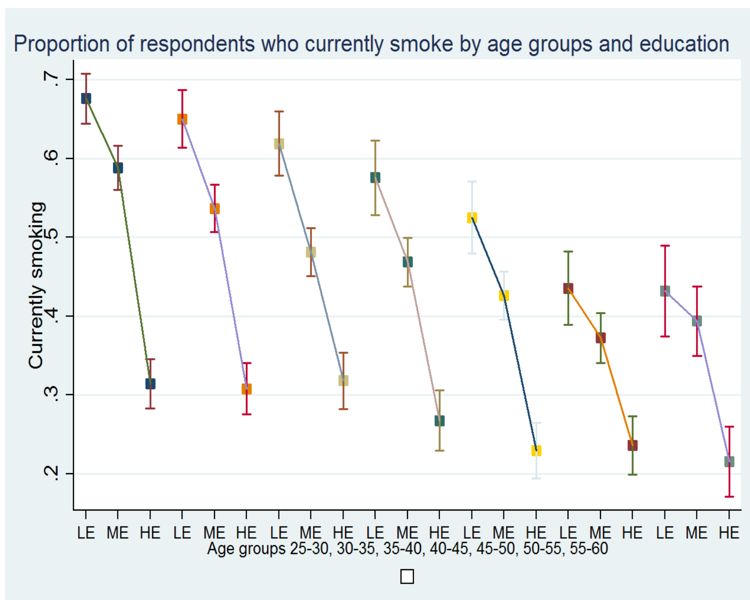
Figure 2.3: Aggregate health index by 5-years age groups for low-, medium- and high-educated subsamples



during the transition period from centralized to free-market economy, and many of them worked in occupations different from their educational formation. Figure 2.3 with health measured as aggregate health index contains no evidence that a higher education level results in a better health. If any, the evidence goes the other way: for several age cohorts (40-45, 45-50 and 55-60) low-educated are in somewhat better health than high-educated. This result contradicts the vast empirical evidence of the existence of education-health gradient in developed countries (see for example Cutler and Lleras-Muney, 2010).

At the same time there is a clear education-health behavior gradient (Figures 2.4-2.6): high-educated smoke less, drink less and do significantly more physical activity than low-educated.

Figure 2.4: Proportion of respondents who smoke at the moment of interview by 5-years age groups for low- medium- and high-educated subsamples



One of the possible explanation of the absence of education-health gradient when health is measured with objective measures is that high-educated have more probability to have their conditions to be diagnosed. In fact, the Figure 2.7 suggests that high-educated have higher frequency of doctor-visits per year, during which health conditions can be diagnosed.

Figure 2.5: Proportion of respondents with drinking problem by 5-years age groups for low- medium- and high-educated subsamples

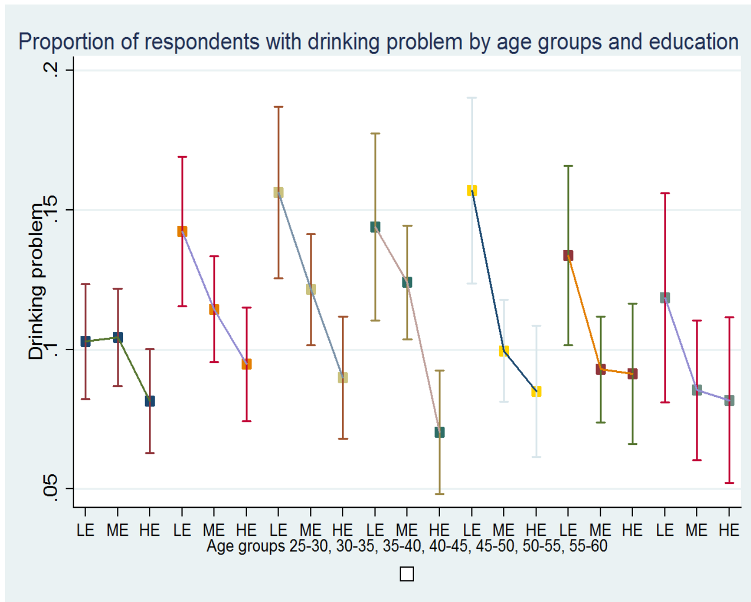


Figure 2.6: Proportion of respondents who do at least some physical exercises by 5-years age groups for low- medium- and high-educated subsamples

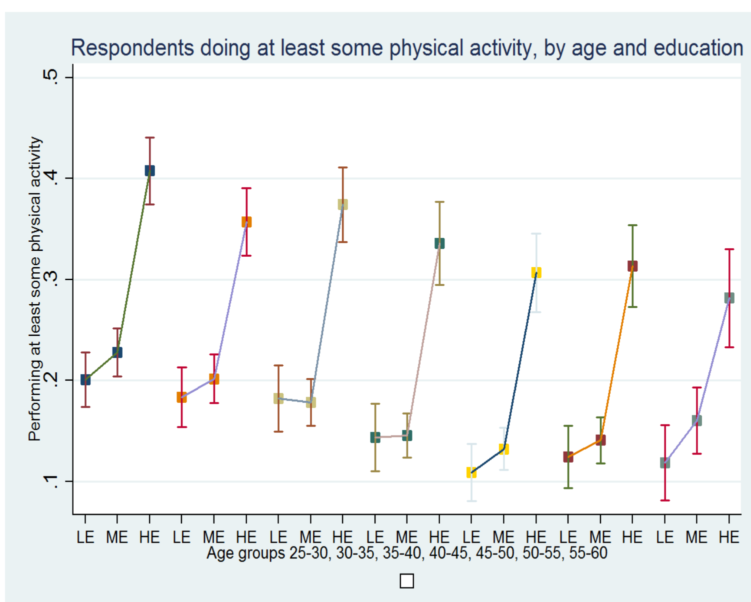
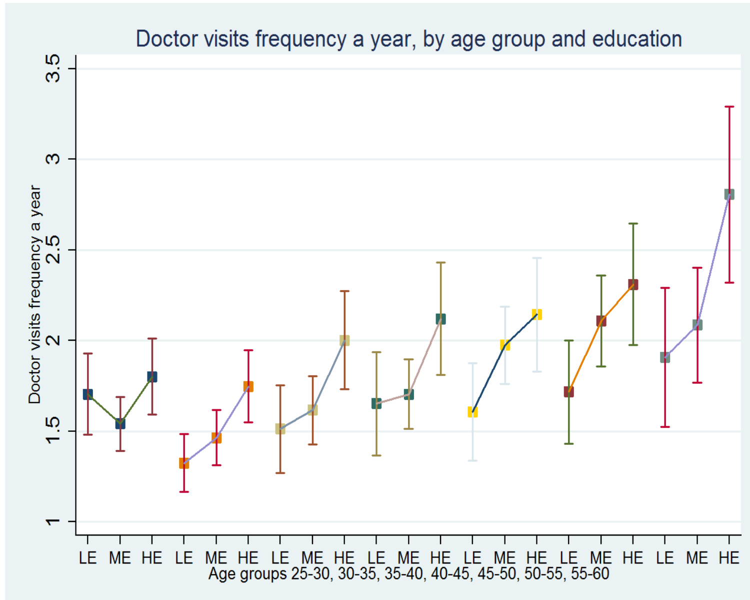


Figure 2.7: Frequency of visits to doctors during the year before the interview by 5-years age groups for low- medium- and high-educated subsamples



2.9 Conclusion

This paper presents a modification of Grossman model of health capital, in which health is produced either using medical care or following health rules. The optimal number of health rules is decided in each period based on future salary, marginal cost of health production, time discounting rate, psychological disutility of following health rules and other factors. I derive a structural equation for the demand for health that depends in particular on the optimal amount of health rules. In the empirical part I estimate the structural equation for the demand for health derived from the model using the Russian longitudinal data set RLMS-HSE.

Employing the Hausman-Taylor model, which under minimum assumptions on the error term provides unbiased estimates of the structural parameters, I find that the data partially confirms the model predictions: wages and optimal number of health rules positively affect the probability to be in good health, while price of medical care has a negative effect. The data does not provide evidence in favour of the model assumption concerning the amplifying effect of education: once controlled for wages, health rules, prices of medical care and demographical variables, years of education have no or, if any, negative effect on health. Among possible explanations of this phenomenon are: long "latency period" between the moment in which health behavior negatively affects true health and the moment when the damage "shows up" in the form of a chronic condition; late diagnostics of chronic conditions by low-educated related to their under-consumption of medical services; low returns on education for older cohorts related to living in the transition period with restricted occupational choices.

Recovering structural parameters from the reduced form estimates allows simplifying the health investment function into the function of just one parameter and comparing elasticities of health to medical care and to health rules. A simple enumerative example illustrates the reasoning that a policy-maker can use when choosing between different policies aiming to improve public health, such as anti-smoking campaigns versus medical care subsidies.

The study is a subject to serious limitations. In the theoretical part, assuming pure investment sub-model is a restrictive assumption: in reality people are likely to care also about their health per se, and not only because it produces healthy days. Also the model presented here completely ignores the whole literature on the myopic decision making in respect to health rules. In the example already discussed in introduction DellaVigna and Malmendier, 2006 show that people behave irrationally when choosing between different options of gym subscription. They suggest that the most probable reasons of irrationality are overconfidence about future self-control and attendance/overestimation and remark that, taking into account myopic behavior, subsidizing enrolment in health clubs is likely to have a low effect on the attendance. In the current paper agents are assumed to be perfectly rational.

In the empirical part, the model does not account for lagged effects: even if health behavior patterns have an immediate effect on the level of "true" health, it might take up to several years until the damage to health gets "revealed" in the form of objective health measure and even more until it gets diagnosed. I find the evidence that in a developing country such as Russia higher-educated follow more health rules than low-educated but also use more medical care. This finding underlines the importance of increasing the awareness of low-educated people both about the damaging effect of adverse health behavior and the necessity of regular medical check-ups.

Chapter 3

Job killing you? Characteristics of the main job and physical and mental health later in life

Keywords: SHARE; Health; Job-related stress; Long-run Effects

JEL: Classification: I14, J81, J14.

3.1 Introduction

In the paper of Muurinen and Le Grand, 1985 the authors develop a theory, according to which a human being has three types of capital: wealth, human capital (education) and health capital. While the first two are distributed quite unequally among individuals, the health capital is distributed more uniformly: the majority of us is born with approximately similar initial level of health. So for those groups of people whose endowment in terms of wealth and education is relatively low, the incentives to "sell" their health capital in exchange for consumption, engaging into hazardous manual jobs is higher than for advantaged groups. At the same time, selection into jobs is based on the level of education, health and possibly wealth, so poorer and less educated groups have less occupational choices.

Case and Deaton, 2005 develop this idea building a model in which blue-collar workers "sell" their health performing hazardous and heavy jobs, because these jobs require physical efforts, while health of white-collar workers whose jobs are based on human capital but not physical health expenditures, depreciates slower. Using the National Health Interview Survey (NHIS) the authors show that health of manual workers declines more rapidly during working years than does the health of nonmanual workers.

While nowadays physically demanding and hazardous jobs remain a socio-economic issue, there was a significant shift towards non-manual, computer-based jobs. Those types of jobs have no or little extent of direct physical depreciation, except maybe problems related to sedentary lifestyle, but might be characterised by a high degree of stress, routine and lack of private space. There is a vast strain of literature that studies the effect of psychologically stressful working conditions on health. In particular, two combinations of job characteristics are found to be harmful for health: high job demand combined with little freedom to decide (demand-control model) and high effort together with low reward (Effort-Reward imbalance model).

Demand-Control model was first introduced by Karasek Jr, 1979 who argues that high job demand creates stress, while the individual's job decision latitude ("control") modulates the release or transformation of stress (potential energy) into the energy of action. If the level of control is low, the unreleased energy may manifest itself internally as mental strain.

The Effort-Reward imbalance model (ERI) developed by Siegrist, 1996 refers to the principle of reciprocity of exchange in occupational life. It proposes that when the job requires high level of efforts but the reward in terms of renumeration, esteem and job perspectives is low, a worker experiences a state of emotional distress associated with the sense of injustice. Siegrist found that jobs characterized by high efforts and low reward are associated with incidence of cardiovascular conditions.

Wahrendorf et al., 2013 use the SHARE data set to study the effect of job characteristics, such as physical and psychosocial job demands, levels of social support, control and reward at the mid-life (at age 40-55 years old) jobs on the number of depression symptoms later in life. They find that men and women who experienced high physical or psychological demand, a low control, low reward or low social support at work during mid-life have significantly higher probability of high depressive symptoms. These associations remain significant after controlling for workers' health and social position in childhood.

In this paper I develop the work of Wahrendorf et al., 2013 in the following directions: first, I study the consequences of adverse working conditions on both physical and mental health, and the results suggest that there is an important difference between them. Second, while they consider the effect of the job in mid-age, I concentrate on the effect of the main job which lasted for at least 10 years, recognizing the importance of the cumulative effect of the adverse working conditions and assuming that working conditions at the main job can proxy those throughout the whole career. Third, I address the problem of potential bias in self-reported working conditions, which might be correlated with self-reported health and the number of depressive symptoms, so I instrument the self-reporting working conditions with the objective ones. Forth, I partially control for self-selection into jobs based on health, excluding respondents who were already in bad health prior to the beginning of the main job and using a rich set of control variables. Finally, I take into account the potential link from the poor physical health to depression by controlling for the physical health in regressions with the outcome number of depressive symptoms.

The remaining of the paper is organized as follows: section 3.2 discusses the theoretical relationship between working conditions and health. Section 3.3 describes the data set, variable construction and econometric strategy. Section 3.4 presents the results. Section 3.5 concludes.

3.2 Theoretical considerations

I follow Ravesteijn et al., 2014 who develop a Grossman-type theoretical model taking into account occupational stressors. I reproduce the model for the sake of discussing the assumptions I have to make in respect to the theoretical model in my econometric strategy.

In economic literature health is considered to be a durable capital stock which depreciates with age and can be increased by investment (Grossman, 1972). While age-related depreciation is exogenous, there are other socio-economic factors, both exogenous and endogenous, that increase the rate of depreciation, in particular adverse work conditions and stress at work.

An individual maximizes the expected value of life-time utility discounted with the rate β. Utility in each period is a positive function of consumption and health with diminishing marginal returns:

$$max_{\{c_{t+j},o_{t+j},m_{t+j}\}_{j=0}^{T-t}} E\Big[\sum_{j=0}^{T-t} \beta^j u(c_{t+j}, h_{t+j})\Big] \tag{3.1}$$

The choice variables are amount of medical care m, occupational stress o and consumption c.

The low of motion of health capital over time is described by equation:

$$h_{t+j} = h_0 + \sum_{k=2}^{t+j} (a_k + \phi^{t+j-k}(\gamma_o^{'} o_{k-1} + \gamma_m m_{k-1}^{\theta} + \eta_k)) \tag{3.2}$$

where h_{t+j} is health in the period t+j, h_0 is initial level of health, a_k is the rate of health deterioration due to age in the year k, ϕ is the rate of the influence of occupational stressors, health investments and shocks on health, which are assumed to decay at the same speed, $0 < \theta < 1$ represents decreasing returns to scale of health investment and η_k is an exogenous health shock.

A theory of "compensating wage differentials" assumes that individuals accept jobs with high level of stress because they are better paid than similar jobs with lower stress.

Total expenditure on consumption and health over all periods cannot exceed the total earnings:

$$\sum_{k=1}^{T}(p^c c_k + p^m m_k) \leq \sum_{k=1}^{T}(1+r)^{k-1} w(o_k, h_k) \tag{3.3}$$

where p_c and p_m are prices of consumption and medical care correspondingly, and r is the interest rate.

The three first-order conditions in respect to consumption, single occupational stress factor o_l and medical care are the following:

$$\frac{\partial u_t}{\partial c_t} = \lambda p_c \tag{3.4}$$

$$\lambda \frac{\partial w_t}{\partial o_{l,t}} = - \sum_{j=1}^{T-t-1} \frac{\partial h_{t+j}}{\partial o_{tl}} \Big[\beta^j \frac{\partial u_{t+j}}{\partial h_{t+j}} + \lambda (\frac{1}{1+r})^j \frac{\partial w_{t+j}}{\partial h_{t+j}} \Big] \tag{3.5}$$

$$\lambda p_m = \sum_{j=1}^{T-t-1} \frac{\partial h_{t+j}}{\partial m_t} \Big[\beta^j \frac{\partial u_{t+j}}{\partial h_{t+j}} + \lambda (\frac{1}{1+r})^j \frac{\partial w_{t+j}}{\partial h_{t+j}} \Big] \tag{3.6}$$

Equation 3.5 can be interpreted as follows: the marginal benefit of occupational stress represented by the product of λ and the instantaneous wage premium is equal to the marginal cost which includes the marginal deterioration of health in all future periods multiplied by the discounted marginal utility of future health, and the product of λ and the present value of the marginal wage returns to future health. This equation sheds the light on the pathways of endogeneity in the association "working conditions-health". It suggests that workers self-select into occupational choices choosing optimal level of stress on the basis of their preferences for health, marginal wage returns to health and the shadow price of income λ. In particular, the equation 3.5 raises concerns regarding the reversed causality and unobserved heterogeneity of occupational choices. Since utility function is concave, level of health in period t affects marginal utility of health, which in turn affects the optimal level of occupational stress. At the same time parameters of utility function might differ from subject to subject and be the source of heterogeneity.

In the subsection 3.3.4 I explain how I deal with these issues.

3.3 Data and methods

3.3.1 Data source and sample selection

In the analysis I use the data from the Survey of Health Ageing and Retirement in Europe (SHARE) for 12 countries: Austria (AT), Germany (DE), France (FR), Belgium (BG), Switzerland, Denmark (DK), Sweden (SW), Italy (IT), Greece (GR), Spain (ES), the Netherlands (NL), Poland (PL) and Czech Republic (CZ). I exclude observations from Switzerland since the data I am using to construct instrumental variables are missing for this country. I consider the respondents who participated both in the second wave of the survey (2006-2007) and in the third wave called SHARELIFE (2008-2009) in which respondents were asked retrospectively about the circumstances of their life, in particular about working career. Combining these two waves allows studying associations between job environment and health at the old age. An ideal case would be to have information on working conditions in all the jobs a respondent did in his career and construct a measure of accumulated stress throughout working life. But SHARELIFE contains information on the start and end date of the job, occupational title and reason of leaving for each job, while questions related to physical and psychological demands, control, reward and support are asked only about main job in the career and the last job, which might be the same job. For this reason I use only information on the main job and restrict the sample to individuals who spent in it at least 10 years. I assume that for them working conditions in the main job can well proxy those throughout the whole career.

I consider only retired people since retirement can be endogenous to working conditions and health. People with adverse working conditions are likely to retire earlier, both because the pension legislation offers this possibility for certain types of jobs and because they are more likely to exit from the labor force through disability retirement. At the same time retirement is found to have its own effect on health. For example, Coe and Zamarro, 2011 using statutory retirement ages as instruments find that retirement improves physical health.

Further I try to minimize health-related bias of selection into jobs with certain working conditions. Wave 3 contains information on childhood health at the age of 10 years old, but the adverse health events that could occur between this age and the beginning of the main job can both affect the choice of the jobs with certain working conditions and physical and mental health later in life.

Respondents of SHARELIFE are asked if they have had injury leading to disability and in which year they had it; they are also asked if they have had periods of ill-health that lasted more than a year. On the basis of this information I exclude respondents who had an injury leading to disability before the beginning of the main job, respondents who report to have had more than 3 periods of ill-health or those whose first period of ill health started before the beginning of the main job.

Details of sample selection are provided in the Table 3.1.

Table 3.1: Sample selection and some descriptive statistics on the selected sample

Sample selection	Males	Females
w2+w3-Switzerland	9607	12197
Ever done paid job	9478	12125
Report to have done main job	6592	7612
Responded to questions about phys. demand of the main job	5911	5911
Retired	4813	3966
Worked in the main job for at least 10 years	4542	3498
Excluding those with ill-health period or injury to disability started before the beginning of the main job	4252	3286

3.3.2 Measures

I use three outcome variables. One is self-reported health taking values from 0 to 4. Self-reported health status is among the most common measures used in public health surveys; it reflects various physical, emotional, social aspects of health and well-being and is found to be a good predictor of mortality (Burström and Fredlund, 2001)

Second one is an index of physical health. It is constructed in the following way: the Principal Component Analysis is run over the number of chronic diseases and mobility limitations and then the index is normalized to 0 for a person with highest number of limitations and chronic diseases and to 1 for a person with minimum number. This index represents the ”objective” health state presumably not affected by the reporting style.

Mental health is measured according to the EURO-D scale as the number of depressive symptoms from the list of 12 symptoms that a subject reports to have (Prince et al., 1999).

Measures of adverse conditions at the main job are derived from 13 questions covering different aspects of working environment. The list of questions used for constructing measures of working circumstances is presented in the Table 3.2. The answers are coded in the ordinal scale from 1 "strongly agree" to 4 "strongly disagree". I group these variables into five categories which are then used to construct measures of unfavourable working conditions, as described in details further. The five categories are: physical demand of the job, psychological demand, reward, control and support.

Table 3.2: Composition of generated work characteristics.

Job characteristics	**SHARE Variables**
Physical efforts	1) Work was physically demanding
	2) Work was uncomfortable
	3) The state took adequate measures
	to protect me from health hazards*
Psychological efforts	1) Heavy time pressure
	2) Work was emotionally demanding
	3) Work involves conflicts
Reward	1) Work gave recognition
	2) Work had adequate salary
	3) Employees were treated fairly
Control	1) Work had little freedom to decide*
	2) Work gave opportunity to develop new skills
Support	1) Work had adequate support
	2) Good work atmosphere

Two variables of interest are constructed in compliance with the literature concerning stressful effect of working conditions on health (Karasek Jr, 1979; Siegrist, 1996). The first one proposed by Siegrist is Effort-Reward ratio (ERI) defined as the ratio between applied efforts and reward from work in terms of monetary renumeration, esteem and status. I consider both physical and psychological efforts of doing the job and calculate the ERI ratio as follows:

$$\text{ERI ratio} = \frac{\text{physical efforts+psychological efforts}}{\text{2*reward}}$$

The numerator is sum of the scores of the variables listed in the Table 3.2 under categories "physical efforts" and "psychological efforts". The denominator is the sum of the scores of the variables in the category "reward". I multiply the denominator by 2 in order to normalize the ratio to take value 1 if the scores in the "efforts" and "reward" categories are equal. ERI ratio equal to one is treated in the literature as a threshold above which the job is considered to be stressful.

The second variable arising from the paper of Karasek Jr, 1979 and the followers is demand-control-support ratio. It is calculated as:

$$\text{Demand-control-support ratio} = \frac{\text{physical efforts+psychological efforts}}{\text{1.5*(control+support)}}$$

Again, denominator is multiplied by 1.5 to normalize the ratio and values higher than one are considered to be the evidence of stressful job conditions in terms of imbalance between high efforts and low control and support.

The control variables considered in the main specification are: country dummies, polynomial of age, childhood self-reported health at the age of 10, educational level and three variables that serve as a proxy for living arrangements and social status at childhood. These variables include a number of rooms per person in a residence of living at 10; a sum of scores for having each of the four amenities in the residence: fixed bath, cold running water, hot running water and inside toilet and a dummy variable for having enough books in the house to fill in at least one bookcase. Educational level is a generated variable in SHARE constructed according to ISCED classification.

Country dummies aim to capture differences in reporting of health status existing between different nationalities (see Jürges, 2007 for discussion of cross-cultural differences in response styles). Age polynomial captures depreciation of health over time. Childhood health and socio-economic situation represent initial health endowment and proxy educational and occupational opportunities that an individual had at the young age. Education is found in the literature to affect health through different mechanisms, in particular through income (see Cutler and Lleras-Muney, 2006 for the discussion on different pathways from education to health).

Table 3.3 presents descriptive statistics of the selected sample consisting of around 7,500 respondents. The last column of the table provides means of the main variables for the sample of all SHARE respondents excluding Switzerland who respond to have had a main job and the career and answered all the questions regarding physical demand of the job (around 12,000 respondents). Comparing first and last columns one can notice that the differences in main socio-demographic characteristics of two samples are small.

Another interesting observation arising from the Table 3.3 is that average number of years spent in the main job corresponds to more than 80% of total years spent at work. Also, main job is likely to be the last one in the career: ordinal number of the main job is very close to the total number of jobs an individual did in his life.

Table 3.3: Descriptive statistics

	mean	sd	min	max	count	mean overall sample
Self-reported health	1.626	1.035	0	4	7562	1.620
Ind. of obj. health	0.829	0.152	0.000	1.000	7556	0.827
No. of depress. symptoms	2.157	2.151	0	12	7565	2.296
Eff.-reward ratio	0.901	0.390	0.250	4	6380	0.917
Dem-cont-sup. ratio	0.876	0.319	0.250	4	7470	0.883
Age	71.78	7.694	50	100	7565	69.66
Male	0.564	0.496	0	1	7565	0.499
Srh at the age of 10	2.975	0.982	0	4	7538	2.929
Number of rooms p.p. when 10	0.674	0.417	0	7	7505	0.685
Accomod. facilities	1.529	1.529	0	5	7516	1.650
$\geq$ 1 bookcase of books	0.286	0.452	0	1	7516	0.297
Educ. level	2.502	1.480	0	6	7542	2.502
Average No. of jobs	2.500	1.690	1	17	7565	2.608
Average ordinal No. of the main job	2.459	1.662	1	15	7565	2.536
Average number of years worked	36.809	9.790	10	73	7565	33.552
Average number of years spent in the main job	29.483	10.810	10	67	7565	25.740

Note: The last column provides means of the variables for the overall sample of SHARE older than 50 and excluding Switzerland who respond to have had a main job and answered the questions about its characteristics (around 12,000 observations)

3.3.3 Associations between health and working conditions: prima-facie results

To have the first idea of the associations between working conditions at the main job and health later in life, I plot the three measures of health (Figure 3.1) and indices of physical and psychological demand of the main job (Figure 3.2) by occupational levels and by country. I distinguish between four occupational groups: high-skilled white collars (legislator, senior official or manager; professional), low-skilled white collars (technician, clerk or armed forces), high-skilled blue collars (sales worker, agriculture, craft worker or plant operator) and low-skilled blue collars (elementary occupation).

Figure 3.1 suggests that blue-collar workers report being in significantly worse health than white collars in almost all the countries except Belgium. They also have lower index of objective health and report higher numbers of depressive symptoms (except Sweden, Netherlands and Belgium, who interestingly belong to one Northern region). Health state is in most cases the best for high-skilled white collars, followed by low-skilled white collars, then by by high-skilled blue collars and low-skilled blue collars are in the lower end of health distribution.

At the same time Figure 3.2 suggests that there is a similar trend for physical demand of the job. Not surprisingly, blue-collar workers' main jobs required more physical efforts than white-collars workers'. As for psychological demand, there is not such a gradient. If anything, white-collar jobs are more psychologically stressful than blue-collar jobs, but these differences are in most cases not statistically significant.

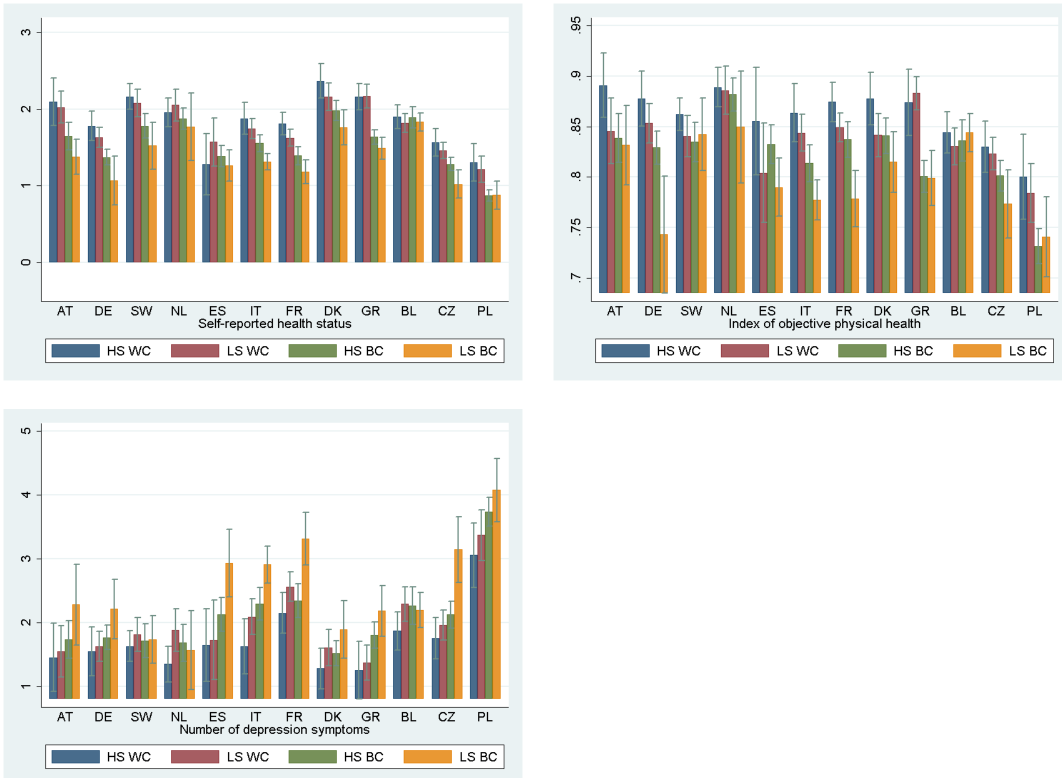
Overall, analysing the graphs brings to the conclusion that there is a negative association between physical and mental health and physical demand of the main job. The plots of raw data do not reveal association between health and psychological demand.

3.3.4 Estimation strategy

The model discussed in the section 3.2 suggests that subjects might choose the optimal levels of stress at work solving the optimization problem and taking into account future health decay and compensating wage differentials. In fact the common approach in the economic literature is to use several consecutive waves of the panel data and to see how health changes from one wave to another based on the occupational circumstances (for example, Robone et al., 2011). In this way occupation choice was already made in the past and current occupation can be treated as exogenous. The disadvantage of this approach is that it allows capturing only short-run effects of stressful working conditions, while I am interested in the long-run effects. So I run linear regressions associating working conditions of only the main job to health later in life, controlling for the rich set of childhood circumstances variables available in SHARELIFE. This solves the problem of endogeneity at least partially.

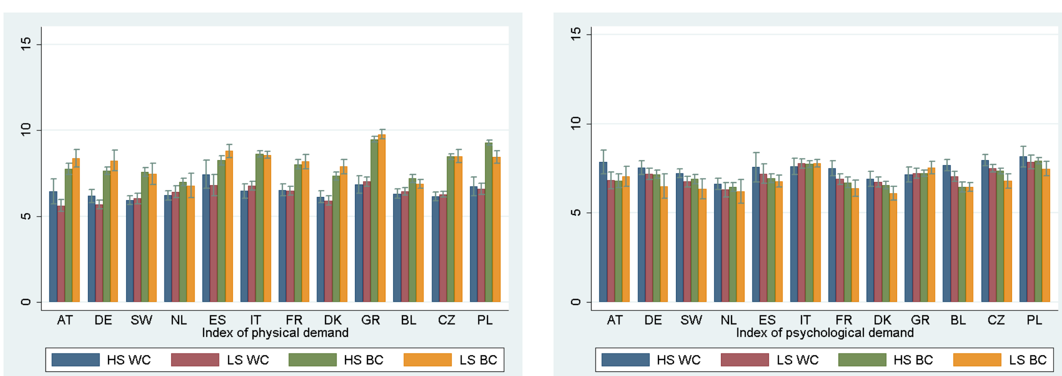
Besides the endogeneity of working conditions, there is one more potential problem which arises from the fact that both health and occupational circumstances are self-reported: a common bias (discussed, for example, by Griffin et al., 2007). For example, a respondent with more pessimistic views can report both his health and working conditions to be worse than they actually were.

Figure 3.1: Means of three health measures by country and occupational level



Note: The three health measures are: self-reported health on the left upper graph, index of objective health on the right upper graph and number of depression symptoms on the left lower graph. Here and further HS WC denotes high-skilled white collars, LS WC low-skilled white collars, HS BC high-skilled blue collars and LS BC low-skilled blue collars

Figure 3.2: Mean values of index of physical demand of the main job (left graph) and psychological demand (right graph) by country and occupation



To address this concern I use an instrumental variables approach, instrumenting subjective measures of working conditions with objective measures derived from the fifth wave of the European Working Conditions Survey.

Working Conditions Survey (WCS)[1] is a survey collected by European Union agency Eurofund since 1990 which provides an overview of working conditions in Europe. It reflects workers' views on a wide range of issues such as work organization, working time, equal opportunities, training, health, well-being and job satisfaction.

In the fifth wave of WCS collected in 2010 almost 44,000 workers (employees and self-employed) from EU27, Norway, Croatia, Macedonia, Turkey, Albania, Montenegro and Kosovo were interviewed face-to-face. The questions covered different aspects of working life, in particular the measures of interest of the current paper (physical and psychological efforts, reward, control and support). Table 3.4 lists the variables used for construction of the indices.

Differently from SHARE survey where possible responses to the working conditions questions ranged from "strongly agree" to "strongly disagree", in the Working Conditions Survey the possible answers were "yes" or "no". I use the information on the percentage of respondents by country and 4-level occupational category who answered "yes" to the questions concerning certain working conditions. Then I construct 5 measures of stress and two ratios of unfavourable working conditions exactly in the same manner as I did for the self-reported variables from SHARELIFE and merge them to the SHARE data by country and occupational level.

Correlations between ERI ratio and Demand-Control-Support ratio calculated from subjective SHARE variables and objective WCS variables are all around 27%.

3.4 Results

In the Table 3.5 I begin the baseline linear regression analysis linking two measures of working conditions with self-reported health. In columns (1) and (2) job conditions are measured as Effort-Reward (ERI) ratio, while in the columns (3) and (4) as Demand-Control-Support ratio.

I find evidence, consistent with prior studies, that health in the old age is positively related to health in childhood and to educational level. The results suggest that stressful conditions at the main job measured as any of the two constructed ratios negatively affect self-reported health. For example, one standard deviation in ERI index leads to the change in self-reported health equal to 0.340. This is an effect

[1] http://www.eurofound.europa.eu/surveys/smt/ewcs/results.htm

Table 3.4: Composition of generated work characteristics using WCS data

Physical demand	1) Exposure to chemical products or substances
	2) Job involves tiring or painful positions
	3) Job involves carrying or moving heavy loads
	4) Exposure to tobacco smoke
Psychological demand	1) Enough time to have job done
	2) Handling angry clients
	3) Job tasks conflict with personal values
	4) Knowledge of what is expected at work
Reward	Job is well paid
	2) Job motivates to give the best performance
Control	1) Able to choose or change the order of tasks
	2) Able to choose methods of work
	3) Able to choose or change speed of work
	4) Able to take breaks when wished
Support	1) Colleagues help
	2) Manager helps

comparable by its' magnitude with a difference in self-reported childhood health by two units (for example, "good" instead of "bad" assuming linearity of the scale of self-reported health).

In Table 3.6 I extend the analysis to other two health measures. I run the regression with dependent variable index of objective health (columns (1) and (2)) and number of depression symptoms (columns (3) and (4)). Finally, I take into account the fact that bad physical health induced by adverse working conditions might itself be a reason for depression. So the last two columns of Table 3.6 provide the coefficients of regressions with number of depression symptoms as a dependent variable where index of objective health is used as an additional control variable.

The results of the analysis with objective health index and number of depression symptoms confirm previous findings regarding self-reported health. Both measures of adverse working conditions negatively affect physical and mental health later in life. Controlling for objective health index reduces the coefficient at working conditions measures in regressions with number of depression symptoms, but it remains significant.

As the next step of the analysis I run Instrumental Variables regressions as described in the section 3.3.

Table 3.7 provides the results of the first stage regressions, where two measures of adverse working conditions are regressed on measures derived from the Working Conditions Survey and the set of controls used in the previous regressions. Since coefficients on the variables of interest are highly significant and F-statistics is higher when 10, the instruments can be considered valid.

The results of the regression with outcomes self-reported health and index of objective health are presented in the Table 3.8. Results in the columns (1) and (2) and (5) and (6) suggest that self-reported health is negatively affected by unfavourable working conditions. Coefficients in columns (3) and (4) and (7) and (8) indicate that when health is measured as the index of objective health, the working conditions ratios are significant only for females.

As for the magnitude of the effect, a change of one standard deviation in the ERI ratio for males results in 0.305 change in self-reported health, which corresponds to a change of around 19% in respect to the mean value of srh in the selected sample. This result does not differ much from the one in the OLS regression.

estimates suggest quite different results than OLS regressions. While adverse working conditions negatively affect self-reported health of both males and females, and moreover the magnitude of the effect does not change substantially, the effect on the depression symptoms disappears. For females the index of objective health measured as the absence of chronic diseases and mobility limitations is negatively affected by work, while for males the result is not significant. It might imply that males shift their estimates of the working conditions downwards in case they have chronic diseases or mobility limitations, while females are more objective.

Further I am interested to see if instrumented working conditions are significantly related also with mental health. The results of the IV regressions with outcome number of depressive symptoms are shown in the Table 3.10. They indicate that once controlling for the index of objective health, no significant effect of any of the two ratio on the number of depressive symptoms is found.

I refine the results splitting the measures of work stress on physical and psychological component. The idea behind is to see if it is purely physical load that negatively affects health or there is also an effect of psychological stress that transmits into physical health state. In Table 3.9 I present the results only for the ERI ratio, which I split into ERI ratio physical and ERI ratio psychological. The coefficients at ERI physical are highly significant for self-reported health and are significant on 10% level for the index of objective health for both males and females. No effect on depression is found, as well as no effect of ERI psychological on any health measure.

Another specification aiming to refine the previous results is using separate health conditions to detect the diseases affected by physical demand of the job. I use four types of conditions which are most likely to be induced by physical burden of the job: heart attack/chronic heart condition, chronic lung condition, high pressure and arthritis. The results provided in the Table 3.11 disclose differences between genders: while both for males and females there is an association between working conditions and heart problems, for males adverse working conditions are associated with the higher probability of arthritis, while for women with probability of chronic lungs disease. One of the possible pathways from adverse working conditions to arthritis is a physical demand of the job, in particular lifting heavy objects, doing repetitive tasks and being exposed to the risk of injuries. These jobs are more likely to be done by males. One of the possible mechanisms from adverse job conditions to lungs disease might be smoking, in particular passive.

To summarize, instrumental variables analysis confirms the results of linear regression analysis regarding self-reported health but puts a doubt on the results regarding mental health that were found in the linear analysis and in the previous literature (Wahrendorf et al., 2013).

3.5 Conclusion

I look at the associations between cumulative effect of adverse working conditions measured as high values of Effort-reward or Demand-control-support ratios and physical and mental health later in life. I partially control for self-selection by including in the analysis health and socio-economic circumstances at childhood and excluding respondents who were already sick before starting the main job.

The main contribution of the paper consists in addressing the problem of a reporting bias which affects both self-reported health and self-reported working conditions. I instrument self-reported working conditions with the mean working conditions by country and occupational level, and find that the results of the instrumental variables regressions differ substantially from linear analysis. While the results for self-reported health remain very similar in significance and sign, the effect of working conditions on mental health becomes insignificant.

The further stratification shows that physical health is affected only by physical component of Effort-Reward and Demand-Control-Support ratios, while no significant effect of psychological component is found. Using single health outcomes as dependent variables suggests that both males and females who experienced adverse working conditions in the main job have a higher probability of chronic heart diseases. In addition, exposed males have higher probability of arthritis, while females of chronic lungs disease.

A future research should clarify why in the instrumental variables regression no effect of adverse working conditions on mental health is found. Different possible explanations would imply different policy implications. If the reason is that subjects accurately report physical burden of health but reveal a substantial bias in reporting psychological stress, the policy makers should concentrate on revoking the adverse effect of physically demanding jobs on health. If instead the reason is that firms inside one occupational level are similar in terms of physical load but differ a lot in terms of psychological stress, so the instrument that varies only by country and by occupational group cannot capture these differences, a policy maker can consider mitigating the effect of psychologically stressful jobs on depression.

The results found so far underline the importance of policies that aim to mitigate the adverse effect of jobs with heavy physical loads. In particular, they provide an argument in support of preserving pension legislation that allows early retirement after a certain number of year in a physically demanding job.

Table 3.5: Linear regression for self-reported health

	(1) male	(2) fem	(3) male	(4) fem
Eff.-reward ratio	-0.376***	-0.354***		
	(0.05)	(0.02)		
Dem-cont-sup. ratio			-0.489***	-0.446***
			(0.07)	(0.04)
age	0.009	0.024	-0.007	0.007
	(0.03)	(0.02)	(0.03)	(0.01)
Srh at the age of 10	0.163***	0.171***	0.158***	0.169***
	(0.00)	(0.02)	(0.00)	(0.02)
No. of rooms p.p.	0.018	0.144***	0.024	0.158***
	(0.03)	(0.03)	(0.02)	(0.03)
Accomodation fac.	0.001	0.020***	-0.001	0.017***
	(0.01)	(0.01)	(0.01)	(0.01)
Number of books when 10	0.089	0.047	0.075	0.044
	(0.06)	(0.03)	(0.05)	(0.03)
Educ. level	0.073***	0.075***	0.073***	0.071***
	(0.01)	(0.02)	(0.01)	(0.01)
Constant	1.62	1.03	2.30	1.98
	0.92	0.58	0.86	0.33
Obs.	3477	2786	4153	3203
R^2	0.169	0.196	0.174	0.212

Note: Here and in all the following regressions standard errors (in parenthesis) are bootstrapped with 100 repetitions. * p<0.10, ** p<0.05, *** p<0.01. Coefficients on country dummies are omitted.

Table 3.6: Linear regression for 1) index of objective health (columns (1) and (2)) 2) number of depression symptoms (columns (3) and (4)) and 3) number of depression symptoms controlling for index of objective health (columns (5) and (6))

	(1) Ind. of obj. health males	(2) Ind. of obj. health fem	(3) No. of dep. symptoms males	(4) No. of dep. symptoms fem	(5) No. of dep. symptoms males	(6) No. of dep. symptoms fem
Eff.-reward ratio	-0.040***	-0.043***	0.672***	0.670***	0.475***	0.430***
	(0.01)	(0.00)	(0.06)	(0.11)	(0.08)	(0.10)
Index of obj. health					-4.945***	-5.560***
					(0.23)	(0.26)
Obs.	3470	2789	3437	2759	3460	2783
R^2	0.103	0.163	0.091	0.148	0.203	0.272
					(0.20)	(0.24)
Dem-cont-sup. ratio	-0.049***	-0.056***	0.883***	0.774***	0.637***	0.458***
	(0.01)	(0.00)	(0.08)	(0.17)	(0.10)	(0.11)
Index of obj. health					-5.007***	-5.828***
					(0.20)	(0.24)
Obs.	4144	3204	4100	3166	4127	3193
R^2	0.107	0.194	0.101	0.157	0.218	0.290

Table 3.7: First stage of instrumental variables regression

	(1) males	(2) fem.	(3) males	(4) fem
WCS Eff.-reward ratio	0.802***	1.385***		
	(0.11)	(0.14)		
WCS Dem-cont-sup ratio			0.441***	0.684***
			(0.05)	(0.06)
Age	-0.009	-0.001	-0.013	-0.001
	(0.01)	(0.01)	(0.01)	(0.01)
Age^2	0.000	-0.000	0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)
Srh at the age of 10	-0.020***	-0.019***	-0.021***	-0.016***
	(0.01)	(0.01)	(0.00)	(0.01)
No. of rooms p.p.	-0.027	-0.025	-0.009	-0.023
	(0.02)	(0.02)	(0.01)	(0.02)
Accomodation fac.	-0.003	-0.000	-0.003	-0.003
	(0.00)	(0.01)	(0.00)	(0.00)
$\geq$ 1 bookcase at 10	-0.020	-0.023	-0.024**	-0.039***
	(0.02)	(0.02)	(0.01)	(0.01)
Educ. level	-0.016***	0.006	-0.019***	-0.010**
	(0.01)	(0.01)	(0.00)	(0.01)
Constant	1.389***	0.920*	1.247***	0.760**
	(0.47)	(0.48)	(0.33)	(0.36)
Obs.	3477	2789	4153	3206
R^2	0.134	0.122	0.199	0.185
F-stats	28.07	20.28	54.14	32.94

*Note:*ERI ratio and demand-control-support ratio obtained from Working Conditions Survey serve as instruments for the ratios obtained from SHARE self-reported working conditions.

Table 3.8: Instrumental variables regressions with outcome variables self-reported health and index of objective health.

	(1) srh males	(2) srh fem	(3) Index of obj. health males	(4) Index of obj. health fem	(5) srh males	(6) srh fem	(7) Index of obj. health males	(8) Index of obj. health fem
ERI ratio	-1.175***	-0.838***	-0.074	-0.105***				
	(0.41)	(0.25)	(0.05)	(0.04)				
Dem-cont-sup. ratio					-1.111***	-1.088***	-0.066	-0.100**
					(0.38)	(0.27)	(0.04)	(0.05)
Age	0.003	0.025	-0.000	-0.002	-0.014	0.009	0.002	-0.004
	(0.04)	(0.03)	(0.00)	(0.01)	(0.03)	(0.03)	(0.00)	(0.01)
Age^2	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Srh at the age of 10	0.147***	0.162***	0.004	0.016***	0.145***	0.159***	0.004*	0.015***
	(0.02)	(0.02)	(0.00)	(0.00)	(0.02)	(0.02)	(0.00)	(0.00)
No. of rooms p.p.	-0.007	0.130***	0.004	0.019**	0.015	0.142***	0.007	0.022***
	(0.05)	(0.05)	(0.01)	(0.01)	(0.04)	(0.05)	(0.00)	(0.01)
Accomodation fac.	-0.002	0.018	-0.002	0.000	-0.004	0.013	-0.001	-0.000
	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
$\geq$1 bookcase when 10	0.066	0.027	0.002	-0.002	0.053	0.009	-0.003	0.002
	(0.04)	(0.04)	(0.01)	(0.01)	(0.05)	(0.05)	(0.01)	(0.01)
Educ. level	0.047***	0.065***	0.007***	0.008***	0.052***	0.049***	0.008***	0.007***
	(0.02)	(0.01)	(0.00)	(0.00)	(0.02)	(0.02)	(0.00)	(0.00)
Obs.	3477	2786	3470	2789	4153	3203	4144	3204
R^2	0.091	0.164	0.096	0.143	0.148	0.174	0.107	0.189

Table 3.9: Instrumental variables regression separately for physical and psychological components of job demand

	(1) srh males	(2) srh fem	(3) Index of obj. health males	(4) Index of obj. health fem	(5) No. of dep. symptoms males	(6) No. of dep. symptoms fem	(7) No. of dep. symptoms males	(8) No. of dep. symptoms fem
ERI ratio	-0.707***	-0.566***	-0.051*	-0.063**	0.283	0.704	0.058	0.354
phys.	(0.22)	(0.17)	(0.03)	(0.03)	(0.41)	(0.45)	(0.36)	(0.41)
Index of obj. health							X	X
Obs.	3478	2790	3471	2793	3437	2763	3430	2763
ERI ratio	8.241	-1.366	0.844	-0.090	0.462	1.437	3.580	0.983
psych.	(21.57)	(0.86)	(2.90)	(0.09)	(11.35)	(1.37)	(17.12)	(1.27)
Index of obj. health							X	X
Obs.	3548	2886	3541	2889	3507	2859	3500	2859

Note: Dependent variables are: 1) self-reported health (columns (1) and (2)) 2) index of objective health (columns (3) and (4)) 3) number of depression symptoms (columns (5) and (6)) and 4) number of depression symptoms controlling for index of objective health (columns (7) and (8)). The explanatory variable ERI ratio is split into physical demand-reward ratio and psychological demand-reward ratio.

Table 3.10: Regressions with outcome variable number of depressive symptoms

	(1) males	(2) fem	(3) males	(4) fem	(5) males	(6) fem	(7) males	(8) fem
Eff.-reward ratio	0.630	1.077**	0.323	0.504				
	(0.66)	(0.52)	(0.73)	(0.52)				
Dem-cont-sup. ratio					1.267**	1.244**	0.993	0.708
					(0.65)	(0.62)	(0.67)	(0.66)
Index of obj. health			-4.993***	-5.524***			-4.931***	-5.760***
			(0.36)	(0.30)			(0.27)	(0.33)
Age	-0.069	-0.160**	-0.077	-0.168**	-0.049	-0.126*	-0.046	-0.144**
	(0.07)	(0.08)	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)	(0.06)
Srh at the age of 10	-0.122***	-0.204***	-0.100***	-0.116***	-0.108***	-0.205***	-0.087**	-0.121***
	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)
No. of rooms p.p.	-0.067	-0.058	-0.051	0.047	-0.091	-0.124	-0.048	0.010
	(0.08)	(0.13)	(0.07)	(0.13)	(0.07)	(0.13)	(0.06)	(0.13)
Number of books when 10	-0.232***	-0.287***	-0.241***	-0.298***	-0.188**	-0.279**	-0.197***	-0.268***
	(0.07)	(0.10)	(0.07)	(0.09)	(0.08)	(0.11)	(0.07)	(0.09)
Educ. level	-0.065*	-0.107***	-0.032	-0.061**	-0.055*	-0.078**	-0.017	-0.037
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)
Accomodation fac.	0.024	0.024	0.011	0.026	0.031	0.027	0.021	0.025
		(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Obs.	3437	2759	3430	2759	4100	3166	4092	3165
R^2	0.091	0.144	0.203	0.272	0.098	0.153	0.217	0.289

Note: In columns (3) and (4) and (7) and (8) I additionally control for index of objective health.

Table 3.11: Probit IV two-step regression for alternative health measures

	(1) heart males	(2) heart fem	(3) lungs males	(4) lungs fem	(5) pressure males	(6) pressure fem	(7) arthritis males	(8) arthritis fem
ERI ratio phys.	0.786**	0.790**	0.574	1.160***	-0.150	-0.008	0.723**	-0.253
	(0.33)	(0.32)	(0.44)	(0.42)	(0.28)	(0.24)	(0.34)	(0.25)
Age	0.123**	0.207***	0.140*	0.0290	0.259***	0.144***	0.0955*	0.120**
	(0.06)	(0.06)	(0.08)	(0.08)	(0.05)	(0.05)	(0.06)	(0.05)
Age^2	-0.001	-0.001***	-0.001	-0.000	-0.002***	-0.001***	-0.001	-0.001**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Srh at 10	-0.004	-0.074**	-0.077**	-0.143***	-0.010	-0.034	-0.012	-0.071***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)	(0.03)
No. of rooms p.p.	-0.004	0.058	0.093	-0.071	0.020	-0.151**	-0.019	-0.168**
	(0.07)	(0.09)	(0.09)	(0.13)	(0.06)	(0.07)	(0.08)	(0.08)
Accomod. fac.	0.052**	-0.008	0.022	0.045	-0.019	-0.043**	0.022	-0.008
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
≥1 bookcase when 10	0.052	0.092	-0.050	-0.028	0.045	0.035	-0.000	0.044
	(0.07)	(0.08)	(0.10)	(0.11)	(0.06)	(0.06)	(0.08)	(0.07)
Educ. level	0.017	-0.006	-0.061	-0.012	-0.040	-0.035	-0.037	-0.057**
	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)	(0.02)	(0.03)	(0.02)
Constant	-7.15***	-10.34***	-7.16**	-3.68	-9.72***	-5.60***	-5.47**	-4.47**
	(2.14)	(2.38)	(2.92)	(3.07)	(1.83)	(1.69)	(2.17)	(1.82)
Obs.	3473	2793	3473	2793	3473	2793	3473	2793

:*Note:* Dependent variables are: 1) heart attack/chronic heart disease (columns (1) and (2)) 2) chronic lungs disease (columns (3) and (4)) 3) high pressure or hypertension (columns (5) and (6)) and 4) arthritis (columns (7) and (8)). Coefficients reported are marginal effects

Appendix A

A.A Tables

Table A1: Years of GDP crises by countries for the period 1954-2004

Country	Periods of GDP crises
Austria	-
Belgium	1975, 1981, 1993
Denmark	1974
France	1975, 1993
Germany	1990, 1993
Greece	1974, 1993
Italy	1975, 1993
Netherlands	1958, 1961, 1981, 1982
Spain	1959, 1993
Sweden	1977, 1991-1993
Switzerland	1958, 1975, 1982, 1991, 1992, 1993

Health measure	**Definition**	**Set of possible answers**	**Model used**
Good health	Reporting very good/good health in wave 1 or excellent/very good health in wave 2	W1: very good, good, fair, bed, very bed; w2: excellent, very good, good, fair, poor	Probit regression

Symptoms	No. of Symptoms experienced in the last 6 months from the list of 12 symptoms	Pain in the back/other joint, Heart trouble/angina, breathlessness, persistent cough, swollen legs, sleeping problems, falling down, fear of falling down, dizziness/faints/blackout, stomach/intestine problems, incontinence, other (wave1) or fatigue (wave2)	Linear regression
Grip strength	Grip strength shown in the test	0-100	Linear regression
No. of mobility limitations	No. of activities with which respondent reports to have difficulty, from the list of 10 activities	Walking 100m, sitting for about 2 hours, getting up from a chair, climbing several flights of stairs, climbing 1 flight of stairs, stooping, reaching or extending your arms above shoulder level, pulling or pushing large objects, lifting or carrying weights over 10 pounds, picking up a small coin from the table	Linear regression

Chronic	No. of chronic conditions ever diagnosed, wave 1: list of 14 chronic diseases, wave2: list of 17 chronic diseases	Heart attack, high blood pressure, stroke, cerebral vascular disease, diabetes, chronic lung disease, asthma, arthritis, osteoporosis, cancer, ulcer, Parkinson disease, cataracts, hip fracture, other fractures, Alzheimer disease, benign tumor	Linear regression
IADL	No. of instrumental activities of daily living (IADLs) from the list of 13 activities with which respondent reports to have difficulties	Dressing, walking across room, bathing or showering, eating, such as cutting up your food, getting in or out of bed, using the toilet, using a map, preparing hot meal, shopping for groceries, making phone calls, doing work around the house or garden, managing money	Linear regression
Depression symptoms	No. of depressions symptoms experienced in the last month out of 12 symptoms forming EuroD scale	Depressed mood, pessimism, suicidality, guilt, sleeping problems, lack of interest, irritability, lack of appetite, fatigue, problems with concentration, lack of enjoyment, tearfulness	Linear regression
Recall	Sum of no. of words remembered at the immediate recall and delayed recall	Max. 20 words	Linear regression

Table A2: Summary statistics

	All sample			
	Mean	Std.dev	Min. value	Max. value
Health measures:				
Good health	0.37	0.48	0	1
Symptoms	1.28	1.47	0	11
Grip strength	36.66	12.25	0	100
IADL	0.25	1.01	0	13
High pressure	0.36	0.48	0	1
Physical inactivity	0.06	0.23	0	1
High cholesterol	0.20	0.40	0	1
Chronic diseases	1.28	1.29	0	12
Mobility limitations	0.99	1.73	0	10
Depression symptoms	2.03	2.08	0	12
Recall	9.05	3.36	0	20
PCA: Health	0	1.74	-2.79	13.58
Explanatory variables:				
No. of crises - GDP, 5%	1.85	1.17	0	5
No. of booms - GDP, 5%	1.01	1.93	0	8
No. of crises - GDP, 10%	3.59	1.77	1	9
No. of crises - Consumption, 5%	1.95	1.95	0	7
No. of crises - Unemployment, 10%	4.11	2.94	0	12
No. of crises - Unemp., 10% county-specific	3.37	0.79	0	4
Females	0.51	0.50	0	1
Age	59.80	5.94	0	1
MED	0.27	0.44	0	1
CONT	0.43	0.50	0	1
NORTH	0.30	0.46	0	1
High school drop-outs	0.43	0.50	0	1
High school graduates	0.35	0.48	0	1
College graduates	0.22	0.41	0	1

Note: Summary statistics for the sample including 22977 respondents from waves 1 and refreshment sample of wave 2 of SHARE data set who satisfy all the following conditions: 1) were born between 1934 and 1954; 2) reside in one of the following countries: Austria Belgium Denmark France Germany Greece Italy Netherlands Spain Sweden Switzerland; 3) did not immigrate in the country after the age of 20; 4) report having worked at least once in life.

Table A3: Marginal effects of all the explanatory variables in the baseline regression

	(1) All	(2) HS dropouts	(3) HS graduates	(4) College graduates
No. of crises	-0.0225**	-0.0424***	-0.0149	0.0232
	(0.010)	(0.013)	(0.019)	(0.025)
Birth year	0.0122***	0.0109***	0.0103***	0.0083***
	(0.001)	(0.001)	(0.001)	(0.002)
Female	-0.0436***	-0.0452***	-0.0234**	-0.0233*
	(0.006)	(0.009)	(0.010)	(0.014)
Wave	-0.0244***	-0.0243**	-0.0052	-0.0495***
	(0.008)	(0.012)	(0.013)	(0.017)
AT	0.0885	-0.141	0.0728	0.252
	(0.095)	(0.164)	(0.159)	(0.205)
DE	-0.324***	-0.304*	-0.445***	-0.130
	(0.089)	(0.181)	(0.147)	(0.186)
SW	0.0250	0.0121	0.0193	0.245
	(0.085)	(0.111)	(0.172)	(0.202)
ES	-0.179**	-0.107	-0.370	-0.0382
	(0.090)	(0.107)	(0.324)	(0.254)
IT	-0.181**	-0.146	-0.233	-0.0184
	(0.083)	(0.103)	(0.176)	(0.293)
NL	0.150	0.233*	0.0228	0.226
	(0.096)	(0.125)	(0.189)	(0.223)
FR	-0.272***	-0.168	-0.404**	-0.204
	(0.086)	(0.114)	(0.167)	(0.193)
DK	0.147*	0.0891	0.0702	0.219
	(0.085)	(0.139)	(0.145)	(0.174)
GR	-0.115	-0.0444	-0.0392	-0.0641
	(0.082)	(0.106)	(0.168)	(0.215)
CH	0.124	0.0684	0.0582	0.891**
	(0.103)	(0.149)	(0.161)	(0.366)
Observations	22880	9778	7838	4914

Note: All coefficients are marginal effects of a probit model. In addition we control for country-specific linear and quadratic trends in birth year.

Table A4: Mean and Minimum GDP growth by age group, mean and maximum unemployment by age group

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
dummy(≥ 1 gdp crisis) 20-25	0.231	0.421	0	1	22880
dummy(≥ 1 gdp crisis) 25-30	0.237	0.425	0	1	22880
dummy(≥ 1 gdp crisis) 30-35	0.230	0.420	0	1	22880
dummy(≥ 1 gdp crisis) 35-40	0.303	0.460	0	1	22880
dummy(≥ 1 gdp crisis) 40-45	0.314	0.464	0	1	22880
dummy(≥ 1 gdp crisis) 45-50	0.240	0.968	0	1	22880
mean(gdp) 20-25	3.892	1.375	-0.094	8.696	22880
mean(gdp) 25-30	3.139	1.69	-0.552	9.184	22880
mean(gdp) 30-35	2.505	1.469	-0.552	8.388	22880
mean(gdp) 35-40	1.99	1.082	-0.806	6.209	22880
mean(gdp) 40-45	1.729	0.903	-0.829	4.399	22880
mean(gdp) 45-50	1.872	0.968	-0.829	4.399	22880
min(gdp) 20-25	0.009	0.023	-0.065	0.055	22880
min(gdp) 25-30	0.006	0.022	-0.065	0.064	22880
min(gdp) 30-35	0.002	0.019	-0.065	0.064	22880
min(gdp) 35-40	-0.003	0.016	-0.065	0.047	22880
min(gdp) 40-45	-0.002	0.014	-0.065	0.028	22880
min(gdp) 45-50	0.001	0.015	-0.038	0.037	22880
dummy(≥ 1 unemp. crisis) 20-25	0.319	0.466	0	1	20327
dummy(≥ 1 unemp. crisis) 25-30	0.448	0.497	0	1	20327
dummy(≥ 1 unemp crisis) 30-35	0.499	0.500	0	1	20327
dummy(≥ 1 unemp crisis) 35-40	0.459	0.498	0	1	20327
dummy(≥ 1 unemp crisis) 40-45	0.433	0.496	0	1	20327
dummy(≥ 1 unemp crisis) 45-50	0.274	0.446	0	1	20327
mean(unemp) 20-25	2.539	1.668	0.004	7.349	21132
mean(unemp) 25-30	3.735	2.748	0.003	15.865	21132
mean(unemp) 30-35	5.617	3.948	0.003	20.523	21132
mean(unemp) 35-40	7.062	4.051	0.205	20.523	21132
mean(unemp) 40-45	8.392	4.18	0.283	22.547	21132
mean(unemp) 45-50	8.615	4.025	0.598	22.547	21132
max(unemp) 20-25	3.417	2.084	0.007	9.91	21132
max(unemp) 25-30	4.794	3.448	0.006	20.212	21132
max(unemp) 30-35	6.815	4.441	0.006	21.602	21132
max(unemp) 35-40	8.374	4.497	0.394	24.171	21132
max(unemp) 40-45	9.702	4.535	0.394	24.171	21132
max(unemp) 45-50	9.85	4.382	0.752	24.171	21132

Appendix B

B.A Derivation of equilibrium demand for health

As described in the Section 2.3, an individual solves the optimization problem finding the optimal levels of (m_t*, c_t* and HR_t*) that maximize the intertemporal utility function 2.2 subject to the low of motion of health capital 2.3 and to the intertemporal budget constraint 2.5.

In the original Grossman, 1972 model length of life is defined endogenously in the model: individual dies when his health level drops below a threshold H_{min}. I follow the simplified setting of Case and Deaton, 2005 and assume that length of life is defined exogenously and is equal to T years.

After substituting medical expenses $P_t^M m_t$ with $\pi_t^I I_t$, where π_t^I is the marginal cost of health and doing some algebraical rearrangement, the budget constraint can be written as:

$$\sum_{t=0}^{T} \frac{c_t}{(1+r)^t} + \sum_{t=0}^{T} \frac{H_t(\pi_{t-1}^I(1+r) - \pi_t^I(1-\delta_t))}{(1+r)^t} = A_0 - \sum_{t=0}^{T} \frac{w_t h(H_t)}{(1+r)^t} - (1+r)(\frac{H_{T+1\pi_{T+1}^I}}{(1+r)^{T+1}} - H_0\pi_0^I) \tag{B.1}$$

This is a discrete time optimal control problem, and under assumptions that functions u and I are differentiable and function u is concave, it can be solved using the Lagrange multipliers method. I assume $\pi_{T+1}^I = \pi_0$=0.

The Lagrangian function is:

$L(c, H, HR, \lambda) = \sum_{t=0}^{T} \frac{u(c_t, H_t(HR_t), HR_t)}{(1+\rho)^t} - \lambda(\sum_{t=0}^{T} \frac{c_t}{(1+r)^t} + \sum_{t=0}^{T} \frac{H_t(\pi_{t-1}^I(1+r) - \pi_t^I(1-\delta_t))}{(1+r)^t} - A_0 - \sum_{t=0}^{T} \frac{w_t h(H_t)}{(1+r)^t})$

Optimality conditions are: $\frac{\partial L}{\partial c_t} = 0$, $\frac{\partial L}{\partial H_t} = 0$ and $\frac{\partial L}{\partial HR_t} = 0$.

These conditions imply:

$$\frac{\partial u}{\partial c_t} = \lambda \left(\frac{1+\rho}{1+r}\right)^t \tag{B.2}$$

$$\frac{\partial u}{\partial H_t} = \lambda \left(\frac{1+\rho}{1+r}\right)^t \left(\pi^I_{t-1}(1+r) - \pi^I_t(1-\delta_t) - w_t\frac{\partial h}{\partial H_t}\right) \tag{B.3}$$

And finally:
$\frac{\partial u}{\partial HR_t} + \frac{1}{1+\rho}\frac{\partial U}{\partial H_{t+1}}\frac{\partial I}{\partial HR_t} =$

$$= \lambda \left(\frac{1+\rho}{1+r}\right)^t \left(\frac{\pi^I_t(1+r) - \pi^I_{t+1}(1-\delta_t)}{(1+r)}\frac{\partial I}{\partial HR_t} - H_t\frac{\partial \pi^I}{\partial HR_t}(1+r) - \frac{\partial h}{\partial H_t}\frac{\partial I}{\partial HR_t}\frac{w_{t+1}}{(1+r)}\right) \tag{B.4}$$

These equations define the optimal level of consumption, health rules and level of health that partially depends on the number of health rules.

I further work with equation B.3 and make the following assumptions to be able to derive the explicit functional form of demand for health function:

- Pure Investment Submodel: $\frac{\partial U}{\partial H_t} = 0$
- Health investment function is of Cobb-Douglas type $I_t = M_t^\beta HR_t^\gamma e^{\theta Educ}$, where θ is an efficiency of health investment proxied by years of education. As shown in Galama, 2011, this function is of decreasing returns to scale, so that $\beta+\gamma < 1$, even though this fact does not affect further derivations. Following Case and Deaton, 2005, I assume that education increases the efficiency of input factors in health production.
- Parameters of the function translating health into healthy days: $h = \Omega - \alpha_1 H_t^{-\alpha_2}$
- Parameters of health depreciation function: $ln\delta_t = \delta_0 + \alpha_3 t + \alpha_4 Y_t$ where Y_t is the vector of parameters different from age factor and number of health rules.
- Following Erbsland et al., 1995 I assume that $\ln(\pi^I_{t-1}(1+r) - \pi^I_t(1-\delta_t)) \approx ln\pi^I_t\delta_t$

Equation B.3 after making this assumptions and taking logarithms simplifies into:
$ln\pi^I_t + ln\delta_t = lnw_t + ln\frac{\partial h}{\partial H_t}$
π^I_t can be found from the condition $I_t = M_t^\beta HR_t^\gamma e^{\theta Educ} = 1$:

$$\pi^I_t = P^M_t e^{-\frac{\theta Educ}{\beta}} HR_t^{-\frac{\gamma}{\beta}} \tag{B.5}$$

Plugging this expression into the equation above and performing some simple algebraical operations results into the following equation for the demand for health:

$$lnH_t = \frac{\beta}{(\alpha_2+1)}(\beta ln(\alpha_1\alpha_2) + \beta lnw_t - \beta lnP^M_t + \gamma lnHR_t + \theta Educ - \alpha_3 t - \alpha_4 Y_t') \tag{B.6}$$

B.B Tables

Variable definitions and summary statistics

Table B1: Variable definitions and summary statistics

Variable	Variable definition	Mean	Std. Dev.	N
Health				
Good health	1 if self-reported health is very good or good, 0 if average, bad or very bad	0.363	0.481	14303
Aggr. HI	predicted part of ordered probit regression of self-reported health on all the objective health condition and normalizing it to be between 0 and 1	0.887	0.120	13392
Chronic cond. - heart	1 if reports to have a chronic heart disease, 0 if not	0.075	0.264	14251
Chronic cond. - stomach	1 if reports to have a chronic stomach disease, 0 if not	0.144	0.350	14228
Chronic cond. - spine	1 if reports to have a chronic spine disease, 0 if not	0.136	0.342	14245
Chronic cond. - lungs	1 if reports to have a chronic lung disease, 0 if not	0.039	0.194	14267
Health behavior				
Currently smokes	1 if current smoker, 0 if not	0.446	0.497	14302
Drinking problem	1 if drinks more than 14 drinks a week for men and more than 7 drinks a week for women where a standard drink is defined as containing 14 grams of alcohol	0.108	0.311	14302
At least some phys.activ.	1 if reports to do al least light physical exercises, 0 if not	0.220	0.414	14302
Regular meals	1 if reports to eat regularly, at least 3 times a day, 0 otherwise	0.440	0.496	7689

Table B1: Variable definitions and summary statistics

Variable	Variable definition	Mean	Std. Dev.	N
Health				
HR index$_3$	1st component of PCA based on 3 health behaviors (curr. smoking, drinking and physical activity)	0.622	0.256	14302
HR index ext.$_4$	1st component of PCA based on 4 health behaviors (original plus regular meals)	0.607	0.237	7689
Other variables defined in the model				
log (wage)	Logarithm of monthly wage PPP-adjusted to the level of 2004	9.026	0.773	14302
log(price of med.care)	Predicted price of all possible out of pocket medical expenses PPP-adjusted to the level of prices of 2004 (see the text for details)	7.681	0.526	14302
Years of education	Total years of education	13.138	2.759	14302
Demographic variables				
Age	Age in years	40.480	9.849	14302
Male	1 if male, 0 otherwise	0.504	0.500	14302
Married	1 if married, 0 otherwise	0.676	0.468	14302
No. of people in the hh	total number of people in the house	2.731	1.134	14302

Table B2: Sample selection

Sample	No. of obs.	In which regressions used
Overall sample waves 15, 18-20	48,359	
Overall sample aged 25-60	37,254	Calculating P^M
Aged 25-60, with completed education	32,795	Calculating aggr. Health Ind.
Aged 25-60, non-missing wage and HR_3 ind.	14,527	Initial analysis (Table 2.3)
Aged 35-60, non-missing wage and HR_3 ind.	9,646	Baseline regressions (Table 2.4), robustness checks (Tables 2.5- 2.6)
Aged 35-60, nonmissing wage and HR_4 ind.	5,312	Robustness checks (Table 2.5)

s

Table B3: Coefficient estimates from probit regression of self-reported good health on objective health conditions

	Probit regression	
Good health		
Chronic cond. - heart	-0.22***	(0.00)
Chronic cond. - stomach	-0.21***	(0.00)
Chronic cond. - lungs	-0.25***	(0.00)
Chronic cond. - kidney	-0.15***	(0.00)
High pressure	-0.14***	(0.00)
Chronic cond. - liver	-0.17***	(0.00)
Chronic spinal disease	-0.20***	(0.00)
Chronic cond. - other	-0.23***	(0.00)
Obese	-0.086***	(0.00)
Anemia last 12m	-0.084***	(0.00)
Diabetis	-0.14***	(0.00)
Ever tuberculosis	-0.14***	(0.00)
Ever hepatitis	-0.058***	(0.00)
Ever stroke	-0.16***	(0.00)
Ever heart attack	-0.031	(0.53)
Observations	30157	
Ln Likelihood	-15943.366	

Table B4: First component of Principal Component Analysis run on 3 and 4 measures of health behavior

	Comp1
Drinking problem	0.6044
Current smoking	0.6622
Physical activity	-0.4429
Variance explained:	42.07%

	Comp1
Drinking problem	0.5497
Curr smoking	0.6466
Physical activity	-0.4221
Regular meals	-0.3187
Variance explained:	33.08%

Table B5: Coefficient estimates of pooled OLS model for aggregate medical expenses

	(1) Payments to doctor official		(2) Paym. to the hospital official		(3) Medicines in the hosp. official		(4) Medicines		(5) Payments to doctors unofficial		(6) Payments to hosp.l unofficial		(7) Medicines in the hosp. unofficial	
Moscow	999.3***	(254.29)	1303.0***	(441.81)	261.0	(435.58)	284.2***	(69.70)	174.1*	(94.27)	1825.1***	(405.63)	88.6	(72.99)
St.Petersb.	195.6	(336.61)	1008.1	(730.59)	-53.3	(612.64)	177.6**	(82.76)	-12.5	(124.23)	283.3	(704.77)	1.0e-11	(103.03)
Moscow region	88.9	(277.26)	108.3	(499.84)	856.9*	(457.57)	221.4***	(76.37)	13.6	(102.52)	304.9	(459.98)	19.6	(79.96)
Komi	325.3	(266.87)	880.8**	(441.81)	1055.4**	(412.70)	277.9***	(76.68)	-28.0	(98.75)	-123.7	(405.63)	23.4	(71.06)
Spb region	44.2	(332.28)	42.4	(558.24)	-67.6	(534.76)	136.0	(84.01)	-27.6	(122.66)	-99.3	(509.48)	1.0e-11	(89.37)
Smolensk	18.0	(343.25)	895.2	(635.53)	-149.2	(661.21)	51.7	(88.52)	558.7***	(127.39)	73.1	(593.25)	9.9e-12	(108.52)
Kalining. region	145.7	(351.83)	957.5	(604.10)	-96.1	(541.98)	-146.0	(90.51)	15.1	(130.62)	345.5	(549.67)	1.0e-11	(91.96)
Tula region	265.1	(344.26)	1704.2***	(558.24)	91.8	(515.86)	239.1***	(85.79)	131.2	(128.16)	117.4	(522.80)	1.0e-11	(87.14)
Kaluga region	-51.9	(374.79)	-2.24	(563.06)	-31.5	(541.98)	-11.6	(88.22)	-15.9	(138.11)	-70.5	(513.72)	1.1e-11	(90.62)
Gorkiy region	122.9	(334.84)	552.6	(590.86)	303.9	(541.98)	-71.5	(85.26)	-3.00	(124.23)	1508.0***	(543.74)	117.8	(89.37)
Chuvashiya	15.7	(339.36)	202.6	(584.74)	270.0	(541.98)	-169.9*	(89.81)	-11.8	(125.24)	-133.5	(549.67)	10.6	(98.62)
Pensa region	-63.7	(542.57)	13.1	(654.41)	1263.4**	(558.12)	-22.1	(99.51)	40.3	(200.78)	-105.7	(584.85)	31.8	(91.96)
Lipetsk region	-71.9	(333.12)	-2.24	(618.89)	-220.0	(558.12)	-276.3***	(91.44)	-29.2	(123.27)	385.9	(549.67)	1.1e-11	(91.96)
Tambov region	-34.0	(335.71)	161.1	(544.98)	174.6	(495.59)	-223.6**	(89.07)	-0.11	(124.23)	127.5	(505.41)	49.2	(82.69)
Tatarstan	298.8	(329.05)	32.7	(563.06)	-181.0	(534.76)	89.6	(81.93)	19.9	(121.77)	-133.5	(518.16)	8.34	(88.22)
Saratov	92.5	(374.79)	710.9	(604.10)	-161.8	(577.06)	-48.8	(86.68)	-15.4	(138.11)	294.4	(555.93)	233.4**	(93.41)
Saratov region	-3.77	(350.69)	183.9	(537.01)	-90.6	(558.12)	-198.9**	(90.97)	-2.38	(129.36)	-63.2	(497.78)	1.0e-11	(95.00)
Volgogr. region	-46.6	(368.68)	82.6	(563.06)	399.2	(558.12)	-148.4*	(90.16)	-9.88	(136.43)	-121.3	(518.16)	1.0e-11	(91.96)
Kabardino-balk.	-83.3	(558.66)	-2.24	(808.91)	-97.7	(681.69)	203.2	(214.00)	79.2	(217.04)	137.5	(645.20)	1.2e-11	(115.59)
Rostov	85.3	(359.10)	365.3	(626.96)	2095.9***	(521.75)	102.7	(85.92)	-3.00	(133.36)	451.2	(555.93)	145.0	(89.37)
Krasnodar	68.7	(270.75)	52.3	(446.13)	-58.9	(439.35)	28.4	(72.72)	45.7	(100.60)	634.4	(412.67)	111.5	(72.39)
Stavropol	34.0	(356.59)	667.1	(604.10)	117.7	(528.03)	-69.6	(85.92)	-15.1	(132.42)	1041.5*	(543.74)	19.6	(89.37)
Cheliabinsk	30.0	(171.13)	78.8	(309.45)	-77.5	(283.60)	-3.23	(42.72)	20.3	(63.33)	-66.7	(288.48)	5.1e-12	(47.50)
Kurgan	-6.82	(342.25)	133.1	(730.59)	-195.4	(906.73)	-198.7**	(91.53)	-14.0	(127.02)	-133.5	(672.32)	1.1e-11	(160.37)
Udmurtiya	-27.8	(306.03)	60.7	(505.03)	-56.6	(510.33)	-146.2*	(84.32)	-28.2	(113.06)	148.2	(464.75)	37.0	(85.19)
Orenburg	43.4	(363.05)	133.9	(644.66)	-77.1	(643.14)	-101.5	(94.41)	-22.4	(133.85)	-133.5	(611.82)	1.0e-11	(100.71)
Perm	12.4	(342.25)	653.2	(584.74)	108.7	(567.20)	-71.2	(89.90)	-26.5	(127.02)	-133.5	(549.67)	1.1e-11	(93.41)
Cheliab. region	10.0	(359.10)	283.8	(715.22)	-42.7	(627.06)	-97.6	(92.12)	5.88	(133.36)	-133.5	(672.32)	9.4e-12	(108.52)
Tomsk	120.1	(299.62)	221.7	(558.24)	-87.6	(534.76)	-33.3	(83.27)	-25.2	(110.87)	424.7	(513.72)	1.1e-11	(88.22)
Berdsk	219.3	(367.23)	145.4	(654.41)	411.8	(627.06)	50.9	(93.17)	108.4	(135.90)	392.2	(593.25)	295.2***	(103.03)
Altai	-25.8	(289.87)	46.5	(434.11)	202.8	(415.01)	-125.3	(76.59)	-25.5	(107.15)	-115.5	(400.32)	1.1e-11	(69.92)
Krasnojarsk	73.5	(275.71)	372.1	(490.38)	-128.5	(479.37)	-89.3	(76.30)	22.1	(102.03)	22.6	(445.39)	1.0e-11	(80.59)
Vladivostok	425.8	(388.70)	1809.9***	(688.09)	151.4	(612.64)	354.6***	(90.42)	35.0	(144.55)	-120.9	(672.32)	1.1e-11	(115.59)
Constant	83.3	(215.03)	2.24	(334.24)	220.0	(327.47)	583.5***	(63.21)	29.2	(79.57)	133.5	(307.59)	-1.1e-11	(55.33)
Observations	4479		1372		904		12326		4464		1376		847	
R^2	0.011		0.031		0.057		0.025		0.010		0.045		0.034	

Note: Amounts are in roubles normalized to 2004 level of prices using official inflation index

Bibliography

Abraira, C. and Lawrence, A. (1978). The staub-traugott phenomenon. iii. effects of starvation., *The American journal of clinical nutrition* **31**(2): 213–221.

Barro, R. J. and Ursua, J. F. (2012). Rare macroeconomic disasters, *Annual Review of Economics* **4**(1): 83–109.

Blam, I. and Kovalev, S. (2003). Commercialization of medical care and household behavior in transitional russia, *RUIG/UNRISD project* **1**.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S. and Zuber, S. (2013). Data resource profile: The survey of health, ageing and retirement in europe (share), *International Journal of Epidemiology* **42**(4): 992–1001.

Börsch-Supan, A., Brugiavini, A., Jürges, H., Mackenbach, J., Siegrist, J. and Weber, G. (2005). *Health, ageing and retirement in Europe First results from the Survey of Health, Ageing and Retirement in Europe*, Mannheim Research Institute for the Economics of Aging (MEA)., Mannheim.

Burbidge, J. B., Magee, L. and Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable, *Journal of the American Statistical Association* **83**(401): 123–127.

Burström, B. and Fredlund, P. (2001). Self rated health: Is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes?, *Journal of Epidemiology and Community Health* **55**(11): 836–840.

Case, A. and Deaton, A. S. (2005). Broken down by work and sex: How our health declines, pp. 185–212.

Coe, N. B. and Zamarro, G. (2011). Retirement effects on health in europe, *Journal of health economics* **30**(1): 77–86.

Contoyannis, P. and Jones, A. M. (2004). Socio-economic status, health and lifestyle, *Journal of health economics* **23**(5): 965–995.

Cordain, L., Eaton, S. B., Sebastian, A., Mann, N., Lindeberg, S., Watkins, B. A., OKeefe, J. H. and Brand-Miller, J. (2005). Origins and evolution of the western diet: health implications for the 21st century, *The American journal of clinical nutrition* **81**(2): 341–354.

Cutler, D., Huang, W. and Lleras-Muney, A. (2014). When does education matter? the protective effect of education for cohorts graduating in bad times, *NBER Working Paper* **20156**.

Cutler, D. M., Knaul, F., Lozano, R., Mndez, O. and Zurita, B. (2002). Financial Crisis, Health Ourtcomes and Aging: Mexico in the 1980s and 1990s, *Journal of Public Economics* **No. 774684**: 279–303.

Cutler, D. M. and Lleras-Muney, A. (2006). Education and health: evaluating theories and evidence.

Cutler, D. M. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education, *Journal of health economics* **29**(1): 1–28.

DellaVigna, S. and Malmendier, U. (2006). Paying not to go to the gym, *The American Economic Review* pp. 694–719.

Di Novi, C. (2010). The influence of traffic-related pollution on individuals' life-style: results from the brfss, *Health economics* **19**(11): 1318–1344.

Doblhammer, G., van den Berg, G. J. and Fritze, T. (2011). Economic Conditions at the Time of Birth and Cognitive Abilities Late in Life: Evidence from Eleven European Countries, *IZA Discussion Paper* **No. 5940**.

Erbsland, M., Ried, W. and Ulrich, V. (1995). Health, health care, and the environment. econometric evidence from german micro data, *Health Economics* **4**(3): 169–182.

Fuchs, V. R. (1982). Time preference and health: an exploratory study.

Galama, T. (2011). A contribution to health capital theory.

Gale, C., Martyn, C., C., C. and AA, S. (2007). Grip strength, body composition, and mortality, *International Journal of Epidemiology* **36**(1): 228–235.

Gerdtham, U. and Ruhm, C. (2006). Deaths rise in good economic times: evidence from the oecd, *Economics and Human Biology* **4**(3): 298–316.

Giuliano, P. and Spilimbergo, A. (2013). Growing up in a recession: Beliefs and the macroeconomy, *Review of Economic Studies* **forthcoming**.

Griffin, J. M., Greiner, B. A., Stansfeld, S. A. and Marmot, M. (2007). The effect of self-reported and observed job conditions on depression and anxiety symptoms: a comparison of theoretical models., *Journal of occupational health psychology* **12**(4): 334.

Grossman, M. (1972). On the concept of health capital and the demand for health, *The Journal of Political Economy* pp. 223–255.

Haas, S. (2008). Trajectories of functional health: The long arm of childhood health and socioeconomic factors, *Social Science & Medicine* **66**(4): 849861.

Hausman, J. A. and Taylor, W. E. (1981). Panel data and unobservable individual effects, *Econometrica: Journal of the Econometric Society* pp. 1377–1398.

Hessel, P. and Avendano, M. (2013). Are economic recessions at the time of leaving school associated with worse physical functioning in later life?, *Annals of Epidemiology* **23**(11): 708–715.

Humphreys, B. R., McLeod, L. and Ruseski, J. E. (2014). Physical activity and health outcomes: evidence from canada, *Health economics* **23**(1): 33–54.

Idler, E. and Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies, *Journal of Health and Social Behaviour* **38**: 2137.

Jenkins, D., Jenkins, A. L., Wolever, T., Vuksan, V., Rao, A. V., Thompson, L. U. and Josse, R. G. (1994). Low glycemic index: lente carbohydrates and physiological effects of altered food frequency., *The American journal of clinical nutrition* **59**(3): 706S–709S.

Jürges, H. (2007). True health vs response styles: exploring cross-country differences in self-reported health, *Health economics* **16**(2): 163–178.

Jürges, H., Avendano, M. and Mackenbach, J. P. (2008). Are different measures of self-rated health comparable? an assessment in five european countries, *European journal of epidemiology* **23**(12): 773–781.

Jylha, M. (2009). What is self-rated health and why does it predict mortality? towards a unified conceptual model, *Social Science and Medicine* **69**(3): 307–316.

Kahn, L. (2010). The long-term labor market consequences of graduating from college in a bad economy, *Labour Economics* **17**(2): 303–316.

Karasek Jr, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign, *Administrative science quarterly* pp. 285–308.

Kesternich, I., Siflinger, B., Smith, J. and Winter, J. (2014). The effects of world war ii on economic and health outcomes across europe, *Review of Economics and Statistics* **96**(1): 103118.

Korhonen, T., Urjanheimo, E.-L., Mannonen, P., Korhonen, H. J., Uutela, A. and Puska, P. (1999). Quit and win campaigns as a long-term anti-smoking intervention in north karelia and other parts of finland, *Tobacco Control* **8**(2): 175–181.

Leist, A., Hessel, P. and Avendano, M. (2013). Do economic recessions during early and mid-adulthood influence cognitive function in older age?, *Journal Epidemiology Community Health* **68**(2): 151–158.

Leon, D. A., Shkolnikov, V. M. and McKee, M. (2009). Alcohol and russian mortality: a continuing crisis, *Addiction* **104**(10): 1630–1636.

Leonard, W. R., Snodgrass, J. J. and Robertson, M. L. (2007a). Effects of brain evolution on human nutrition and metabolism, *Annu. Rev. Nutr.* **27**: 311–327.

Leonard, W. R., Snodgrass, J. J. and Robertson, M. L. (2007b). Effects of brain evolution on human nutrition and metabolism, *Annu. Rev. Nutr.* **27**: 311–327.

Mattson, M. P. (2012). Evolutionary aspects of human exerciseborn to run purposefully, *Ageing research reviews* **11**(3): 347–352.

Mazzonna, F. (2014). The long-lasting effects of family background: A european cross-country comparison, *Economics of Education Review* **40**(C): 25–42.

Miller, D. L., Page, M. E., Huff Stevens, A. and Mateusz, F. (2009). Why Are Recessions Good for Your Health?, *American Economic Review: Papers and Proceedings* **99**(2): 122–127.

Mundlak, Y. (1978). On the pooling of time series and cross section data, *Econometrica: journal of the Econometric Society* pp. 69–85.

Muurinen, J.-M. and Le Grand, J. (1985). The economic analysis of inequalities in health, *Social Science & Medicine* **20**(10): 1029–1035.

Oreopoulos, P., von Wachter, T. and Heisz, A. (2012). Short- and long-term career effects of graduating in a recession, *American Economic Journal: Applied Economics* **4**(1): 157–196.

Prince, M. J., Beekman, A., Deeg, D., Fuhrer, R., Kivela, S.-L., Lawlor, B., Lobo, A., Magnusson, H., Meller, I., Van Oyen, H. et al. (1999). Depression symptoms in late life assessed using the euro-d scale. effect of age, gender and marital status in 14 european centres., *The British Journal of Psychiatry* **174**(4): 339–345.

Ravesteijn, B., Van Kippersluis, H. and Van Doorslaer, E. (2014). The wear and tear on health: What is the role of occupation?

Robone, S., Jones, A. M. and Rice, N. (2011). Contractual conditions, working conditions and their impact on health and well-being, *The European Journal of Health Economics* **12**(5): 429–444.

Ruhm, C. (2000). Are recessions good for your health?, *Quarterly Journal of Economics* **115**(2): 617650.

Ruhm, C. (2005). Healthy living in hard times, *Journal of Health Economics* **24**(2): 341363.

Ruhm, C. (2012). *Elgar companion on health economics*, Cheltenham, UK: Edward Elgar, chapter Understanding the relationship between macroeconomic conditions and health, pp. 5–14.

Ruhm, C. (2013). Rescessions, healthy no more?, *NB* **19287**.

Schoenborn, C. A. (1986). Health habits of us adults, 1985: the" alameda 7" revisited., *Public Health Reports* **101**(6): 571.

Schröder, M. (2012). Jobless Now, Sick Later? Investigating the Long-term Consequences of Unemployment on Health, *unpublished Working Paper* .

Schwandt, H. (2014). Wealth shocks and health outcomes: Evidence from stock market fluctuations, *CEP Discussion Paper* **1281**.

Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions., *Journal of occupational health psychology* **1**(1): 27.

Smith, K. R., Mineau, G. P., Garibotti, G. and Kerber, R. (2009). Effects of childhood and middle-adulthood family conditions on late-life mortality: Evidence from the Utah Population Database 1850-2002, *Social Science and Medicine* **68**: 1649–1658.

Stevens, A. H., Miller, D. L., Page, M. E. and Filipski, M. (2011). The best of times, the worst of times: Understanding pro-cyclical mortality, *NBER Working Paper* **No. 17657**.

Stowasser, T., Heiss, F., McFadden, D. and Winter, J. (2011). "Healthy, Wealthy and Wise?" Revisited: An Analysis of the Causal Pathways from Socio-Economic Status to Health, *NBER Working Paper* **No. 17273**.

Stuckler, D., Basu, S., Suhrcke, M., Adam, C. and McKee, M. (2009). The public health effect of economic crises and alternative policy responses in Europe: an empirical analysis, *The Lancet* **374**(July 25): 315–323.

Sullivan, D. and von Wachter, T. (2009). Job Displacement and Mortality: An Analysis Using Administrative Data, *Quarterly Journal of Economics* **124**(3): 12651306.

van den Berg, G. J., Doblhammer, G. and Christensen, K. (2009). Exogenous determinants of early-life conditions, and mortality later in life, *Social Science and Medicine* **68**: 1591–1598.

van den Berg, G. J., Lindeboom, M. and Portrait, F. (2006). Economic Conditions Early in Life and Individual Mortality, *American Economic Review* **96**(1): 290–302.

Van Kippersluis, H., Galama, T. et al. (2013). Why the rich drink more but smoke less: The impact of wealth on health behaviors.

Wagstaff, A. (1985). Time Series Analysis of the Relationship Between Unemployment and Mortality—A Survey of Econometric Critiques and Replications of Brenner's Studies, *Social Science and Medicine* **21**(9): 985–996.

Wahrendorf, M., Blane, D., Bartley, M., Dragano, N. and Siegrist, J. (2013). Working conditions in mid-life and mental health in older ages, *Advances in life course research* **18**(1): 16–25.

Estratto della Tesi di Dottorato

Studente: Liudmila Antonova
Matricola: 955824
Dottorato: Economia
Ciclo: 26

Titolo della tesi: ”Tre saggi sul effetto di fattori economico-sociali sulla salute: approcio *”life-course”*.

Estratto: Questa tesi adotta l’approccio ”life-course” per studiare gli effetti dei diversi fattori economico-sociali sulla salute individuale. Questo approccio si basa sulla idea che la salute è un capitale simile a capitale finanziario o umano. L’effetto di un evento avverso in un periodo può rivelarsi vari anni se non è stato eliminato dagli supplementari investizioni nella salute o politiche favorevoli. Nel Capitolo 1 si studia l’effetto di lunga termine della crisi negli anni lavorativi (20-50) sulla salute nell’età avanzata mediante i dati su 11 paesi Europei. La crise macroeconomica è un evento perfettamente esogeneo che colpisce tutta la popolazione del paese. I risultati mostrano che averne sperimentata una profonda (almeno 1% declino nel PIL) riduce significativamente la salute nell’età avanzata con una incidenza piu larga nei gruppi con meno anni di istruzione scolastica (i.e. che non hanno finito la scuola superiore). Al contempo, la popolazione con alto livello di istruzione non risulta colpita dalle crisi: inoltre, si riscontra un effetto positivo degli anni di crescita economica sulla salute dei laureati. Nel Capitolo 2 viene modificato il classico modello della capitale della salute di Grossman, 1972: viene infatti introdotta la nuova variabile ”regole della salute”. Tali regole, che sono assunte come la conoscenza universale, partecipano nella produzione di salute, non hanno alcun costo monetario ma vengono seguiti al costo di indurre la disutilità e rappresentano una scelta completamente esogenea. La nuova equazione strutturale di domanda per la salute è stata sviluppata e stimata avvalendosi di dati sulle famiglie russe. I parametri strutturali sono stati recuperati dalle stime e assunti come cardini per condurre un’analisi di costi ed efficienza. Il Capitolo 3 studia l’effetto delle caratteristiche del lavoro principale nella carriera lavorativa sulla salute fisica e mentale nell’ età avanzata. La scelta effettuata automamente di lavorare nel settore caratterizato di effetti aversi sulla salute fisica e mentale representa uno spartiacque fra fattori esogenei ed endogenei. Da una parte le persone possono scegliere questi lavori a fronte di una renumerazione monetaria per gli rischi elevati, dall’altra i gruppi sociali svantaggiati (ad esempio quelli con scarsa istruzione) imbattendosi in un ventaglio di scelta occupazionale assai ridotto debbono coattivamente accettare lavori a elevato rischio per la salute. L’analisi lineare mostra l’esistenza di una la forte correlazione fra le avverse charatteristiche di lavoro e salute sia fisica che mentale. Sucessivamente le caratteristiche di lavoro vengono instrumentati con quelle medie per occupazione e paese, risolvendo il propblema di potenziale errore nelle risposte sulle domande sul charatteristiche di lavoro e la salute. L’analisi a variabile strumentali conferma i resultati precedenti sulla la salute fisica, mentre non risulta alcuna associazione fra le avverse condizioni di lavoro e la salute mentale.

La politica pubblica può cercare di distinguire fra gli eventi esogenei che non dipendono dalle scelte degli individui, e quelli in cui loro si auto-scelgono a base di loro preferenze per la salute e altri beni.