Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Life expectancy and its determining elements: A study for the Czech Republic at the beginning of the 21th century

Author: **Vojtěch Korbelius**

Supervisor: **Mgr. Michal Paulus**

Academic Year: **2013/2014**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 13, 2014

Signature

Acknowledgments

The author is grateful especially to his supervisor Mgr. Michal Paulus for his unending help and expert advice.

Abstract

In our thesis we are modelling a life expectancy function for the Czech Republic at the beginning of the 21th century. Our model is using three types of explanatory variables - environmental (socio-economic), health care and environmental pollution.
Our study is the first study not only for the Czech Republic but it is a first study to include environmental pollution variables into a complex life expectancy model.
As a result, we found two different life expectancy functions where one is applicable for male ta the age of 45 and 65 and female at the age of 45 and the second which is the best model for female ta the age of 65.
General outcome points out three things – only one health care factor is significant at a time, environmental pollution is category that has to be considered and the percentage of people over the age of 65 is significant variable determining the life expectancy the most.

JEL Classification	A13, C51, E24, H51, H53, I10, I12, J10, Q25
Keywords	life expectancy, health infrastructure, socioeconomic variable, environmental influences

Author's e-mail	v.korbelius@gmail.com
Supervisor's e-mail	mpaulus0@gmail.com

Abstrakt

V této práci modelujeme funkci délky dožití pro Českou republiku na začátku 21. století. V modelu používáme tři kategorie vysvětlujících proměnných – socioekonomické prostředí, zdravotní péče a znečištění životního prostředí.
Tato práce je nejen první prací tohoto druhu pro Českou republiku, ale je i první studií, která do komplexního modelu pro délku dožití zahrnuje i proměnné vyjadřující znečištění životního prostředí.
Ve výsledku jsme nalezli dvě různé funkce délky dožití, kdy jedna je nejlepším možným modelem pro muže ve věku 45 a 65 let a pro ženy ve věku 45 a druhou, která je nejlepším modelem pro ženy ve věku 65 let.
Naše studie přišla se třemi hlavními závěry – vždy je signifikantní pouze jeden faktor z kategorie zdravotní péče, znečištění životního prostředí je kategorie

proměnných, která nesmí být opomíjena a procentuální vyjádření počtu lidí ve věku 65+ je signifikantní proměnou dokonce s největším vlivem na celkovou délku dožití.

Klasifikace JEL	A13, C51, E24, H51, H53, I10, I12, J10, Q25
Klíčová slova	dožití, očekávaná délka života, zdravotní infratruktura, socioekonomické faktory, vlivy životního prostředí

E-mail autora	`v.korbelius@gmail.com`
E-mail vedoucího práce	`mpaulus0@gmail.com`

Contents

List of Tables

List of Figures

Acronyms

LF Life expectancy

SLF Subjective life expectancy

CSO Czech Statistical Office

Bachelor Thesis Proposal

Author	Vojtěch Korbelius
Supervisor	Mgr. Michal Paulus
Proposed topic	Life expectancy and its determining elements: A study for the Czech Republic at the beginning of the 21th century

Topic characteristics The goal of bachelor thesis: *"Life expectancy and its determining elements: A study for the Czech Republic at the beginning of the 21th century"* is to determine the dependence of life expectancy on various factors. Life expectancy is understood as the probability of living up to a certain age, for which data are available through Czech Statistical Institute. First of all, we will consider factors connected with health infrastructure in particular division/area. Secondly, we will take into account environmental (socio-economic) variables and at last but not least we will focus on environmental pollution influences as well. We will try to determine which of those factors are the most influential and how do they influence life expectancy.

Aim The aim of this thesis is to find factors determining the life expectancy within the Czech Republic. We want to find a model which would in the future help to optimalize the health care and pension system within the Czech Republic. Moreover the practical use of this thesis could be a stimulus in the future for similar thesis in other countries.

Research question Our main research question is:
What factors are influencing life expectancy in the Czech Republic nowadays at the beginning of the 21th century, and how we can improve them.

Methodology In our study we are going to look for model for life expectancy for male and female at the age of 45 and 65 respectively. We are going to use

the method of Ordinary least squares socalled OLS due to its simplicity and convenience.

Outline

1. Introduction
2. Literature review
3. Life expectancy
4. Hygiene and democracy
5. Subjective life expectancy
6. Effect of increasing life expectancy
7. Model
8. Explanatory variables
9. Methodology
10. Comments on results
11. Discussion
12. Robustness test
13. Conclusion

In the first part we will analyze the previous results in this field and state by which factors is the life expectancy effected. In the second part we will focus on setting up a robust econometrical model. In the third, analytical part, our goal will be to explain the effects of discrete explanatory variables including their significancy. In the conclusion we will summarize the information about the development of discrete explanatory variables and their total significancy.

Core bibliography

1. Shaw *et al.* (2005)
2. Walker (2013)
3. Eggleston & Fuchs (2013)
4. Retzlaff-Roberts *et al.* (2004)
5. Miller Jr & Frech (2000)
6. Lin *et al.* (2012)

Author Supervisor

Chapter 1

Introduction

Life expectancy or in other words longevity has become topic of increasing importance in recent years thanks to the fact that during the 20th century life expectancy has increased according to researches between 40 and 50 percent (Eggleston & Fuchs, 2013 ; Walker, 2013). We could say that it may not have been such a shocking fact unless we realize that based on some theories life expectancy has increased during the past two million years from the very beginning of human civilization until the end of the 19th century only to 48 years. (Walker, 2013) In that light 30 years, by which life expectancy increased during the 20th century can seem as a miracle. This rapid and sudden increase is the reason for life expectancy to be in the center of interest from various points of views.

The goal of this paper is to determine factors influencing life expectancy in the Czech Republic on district level. To achieve this goal we have collected various factors in the three main categories environmental (socioeconomical) (Shaw et al., 2005) , health care (Retzlaff-Roberts, Chang, & Rubin, 2004) and a newly considere category envrionmental pollution. Based on those factors we are going to determine relationship between those explanatory variables and life expectancy as the dependent variable.

Life expectancy as a measure is not important only from the point of view of health science but it has direct economic consequences in form of pension incentives and public health investments. Even though that those facts do not have to be evident on the first sight they are inherent part of phenomenon of increasing life expectancy.

From economic point of view we have to distinguish if the increase in life expectancy is increase in productive age or only in post-productive age. So if an

individual has or does not have more time to work and save money for longer period in retirement. (Yakita, 2005) .
The problem appear if work force or state do not realize the increase in life expectancy. In this case people will spend more time in pension, so they will get more money from pension funds even thought they did not pay them there during the time they worked. It is the same scenario with health care system. Simply said people will take more from the system on pension and health care than they paid into the system on taxes. It means that unrealized increase in life expectancy would not cause economic growth but rather decrease. (Echevarr, 2003) From that follows that if we do not realize change in longevity correctly it can cause more harm than good. Walker (p. 36, 2013) states a very interesting question in his paper: *"Are clients* (clients of insurance companies A/N) *ready-mentally and financially-for potentially longer life? Reality and statistics suggests that they are not."* For this simple reason many authors mention that action that should go along with increasing life expectancy is enlightenment about this increase and education that can assure not only longer but better life too.
In the Czech Republic the situation during the 20th century has been the same as in the developed parts of the world. Thanks to the introduction of basic hygiene and democratization at the end of the century we got to the almost same level as the rest of the western world so we could start evolving in other areas, mainly in health care system (improvement and increase of outpatient facilities, increase in consumption of pharmaceutical etc.) and in socioeconomical field with economic growth and connection to healthier way of living. And exactly this is going to be the topic of our study. The goal of this paper is to examine factors influencing life expectancy in the Czech Republic on district level and comment on possible improvements. From this statement we can easily define our research question:

"What factors are influencing life expectancy in the Czech Republic nowadays at the beginning of the 21th century and how we can improve them"

For answering our research question we have to find a satisfactory model based on our own data and on adapting previous works done in this particular field. We will base model on mainly Shaw et al., (2005) but we will look for methodological support eleswhere too, primarily in Miller Jr & Frech (2000) whose work was focused on consumption of pharmaceutical and in Chan & Kamala

Devi, (2012) who modeled life expectancy on district level as it is our intension. Thanks to the work previously done in this field we can state few hypothesis which we want to either confirm or reject.

H.1. The consumption of pharmaceuticals have a significant effect on the life expectancy.

H.2. The enviromental pollution factors - polution of water and air - are significant variables influencing the life expectancy.

H.3. The amount of people over 65 in the population is important factor determining the life expectancy.

The thesis is organized as follows. We start with Chapter 2 Literature review. We are will begin with a deeper explanation what life expectancy is and why it is important, we introduce a new term subjective live expectancy that is bringing more light into the problem of life expectancy and its influence on economy and why longer life expectancy without knowing about it can cause more harm than good. At the end of this chapter we summarize previously done research in this field.
In Chapter 3 Methodology we will present model of Shaw et al., (2005) and what are its advantages and disadvantages. We have chosen this model as the basis for our own model which we will present next.
After defining the model and running the simulations we will be able to interpret the results in the Chapter 4 Interpretation.
Based on our findings we will able to have a Discussion in Chapter 5 where we will comment on our findings and explain the phenomenon that will appear in our results.
Before concluding the whole study we will test whether the model is able to stand face to face to changes of initial conditions, in form of explanatory variables, in Chapter 6 Robustness test.
At the end we will highlight our findings and summarize them in a simple and clear way in Chapter 7 Conclusion. We will make remarks about limitations of our study, which we expect to be to some extent the same as in (Chan & Kamala Devi, 2012) and we will recommend what should be the next step to research in this topic and areas connected to it.

Chapter 2

Literature review

2.1 Life expectancy

So let us now proceed with a chapter focused on life expectancy and the development it went through in the 20th century. According to Walker (2003) the 20th century was the time of the biggest change in life expectancy during the whole human history. And just this significant change is the reason to focuse what increase in life expectancy means from a broader point of view.
Life expectancy is defined as *"the average period that a person may expect to live"* (Definition of life expectancy in English; 2014) . As stated earlier it is one of the key factors for measuring and comparing the health, mortality and morbidity in various countries, regions or parts of the world (Egidi & Spizzichino, 2008) .
Life expectancy can be measured at any point in life. In this paper we will focus on life expectancy at the age 45 and 65. We do not consider sometimes-mentioned life expectancy at birth for simple reason, which will be explained in later chapter (3.2) about explanatory variables.
Many studies have focused on change of life expectancy during the past more than 100 years (Ashraf et al., 2013; Eggleston & Fuchs, 2013; Lin, Chen, Chien, & Chan, 2012) . During the 20th century conditions in Europe and the whole world have significantly improved. We can highlight the fact that life expectancy at birth has doubled during those more than 100 years and the reason for that is very simple. Low life expectancy at birth at the end of the 19th century was mainly caused by high infant mortality (Eggleston & Fuchs, 2013) , which was mainly caused by poor hygienic life conditions. On the other hand we feel the need to clarify that life expectancy has increased generally,

at any point in life but not as much as at birth. Life expectancy has generally increased during the past 160 years by 3 month every year (Seifarth, McGowan, & Milne, 2012) . In total it is 40 years.

2.2 Hygiene, democracy and their role

As we mentioned earlier the 20th century was time of great changes in the field of life expectancy. A lot of authors agree on the two very important factors that helped life expectancy to increase in such a significant way. They were introduction of basic hygiene conditions of life and democratization. (Lin et al., 2012; Mackenbach, Hu, & Looman, 2013; Miller Jr & Frech, 2000; Shaw et al., 2005) .

Let us take democratization per se. In Europe we had two waves of democratization during the past 100 years (Mackenbach et al., 2013) , which have helped to create life expectancy gap between Middle-Eastern Europe and Western Europe, which has not closed fully yet. Lin et al. (2012) and Mackenbach, Hu, & Looman (2013) agreed that democracy has certainly some influence on life expectancy. Both of those researches found out that life expectancy is positively correlated with democracy. Lin et al. (2012) point out that democracy does not have an immediate effect but its impact can be observed with some delay. This can be simply explained by the argument of Mackenbach et al., (2013) who is stating that democracy means less bribing, clearer financing and in consequence more resources that would be otherwise used inefficiently and can be better invested in health care. This argument is based on the research of Sung (2004) that found out that democracy in long run reduces bribing. The fact of more free resources for investment in health care system is an opinion of Mackenbach et. al., (2013) .

Other reason for the major increase in life expectancy, and probably the biggest, has been improvement in sanitization of cities and villages by introduction of sewerage and water supply. This factor has been pointed out by many authors (Miller Jr & Frech, 2000; Shaw et al., 2005) and has been considered as the core foundation for increasing life expectancy.

We can confidently say that we were able to observe influence of both of those factors during the past more than a century and see its consequences on increase in life expectancy in the Czech Republic.

This leads us to one phenomenon, life expectancy has increased but do people

know about it? And if yes do they know about the particular length? Is it possible that people overestimate or underestimate life expectancy and if yes is it important? Those questions are very important for our research question and the answer for them can be found in so called "Subjective life expectancy". Chan & Kamala Devi (2012) who are in their paper focusing on life expectancy in the South-East Asia are coming to the conclusion that the most important thing to do for increase in life expectancy on the national level is to blur the differences between various areas, distinct cities and villages, not only in the basic life conditions but in the health care system as well. This is aming for creating one standard level of health care system as a safest net across the whole country.

2.3 Subjective life expectancy

Life expectancy at first glance may seem as, more or less, health phenomenon that do not make any sense to research from an economic point of view. But if we think about consequences of life expectancy we realize that expected length of life is influencing length of retirement which is closely linked with amount of money paid into pension funds during work-years and money paid out of those funds back to the people.

Nation as a workforce is the source of a wealth in each state. From this simple relationship we can anticipate that life expectancy, health of the nation, is influencing GDP, wealth of the nation. (Ashraf et al., 2013; Eggleston & Fuchs, 2013; Miller Jr & Frech, 2000; J Mirowsky, 1999) .

As we said in the previous chapter, life expectancy is the amount of years a person is expected to live at the certain point in time. Life expectancy is calculated as follows:"*They are the results of completed life tables prepared for each administrative districts of municipalities with extended powers separately for men and women using the indirect method of calculating the probability of dying, one-year age interval, and for the reason to eliminate random fluctuations they are calculated for five-year calendar period.*" (Life expectancy in districts (LAU 1) in 2008-2012 (data extra regular outputs), 2013) But this amount is usually unknown for average people so we are getting to a different phenomenon called subjective life expectancy."*Subjective (or self-estimated) life expectancy (SLE) is a measure that quantifies the perceived extent of one's remaining years, providing a personalized timeframe that can act as a guide for apportioning work, leisure, and finances.*" (Griffin, Loh & Hesketh, p. 79, 2013)

Government is setting up retirement age based on estimated life expectancy of an average person, but average person is setting up amount of savings and is planning its future based on its own subjective life expectancy. (Miller Jr & Frech, 2000; J Mirowsky, 1999; John Mirowsky & Ross, 2013) Person is estimating its subjective life expectancy based on many factors but mainly on longevity of people they knew and comparing their healthy and unhealthy habits with their own. This leads them to certain numbers. Those number may be different from the calculated life expectancy.
Eggleston & Fuchs (2013) point out that people early in life tend to underestimate this number; this may cause wrong decision early in life that can be hardly corrected later on. From this fact we can easily see that education in this field is important to use longer life expectancy to the fullest. Eggleston & Fuchs (2013) found out that people realize their more accurate life expectancy late in life when they are already in the retirement. The effect is that they become more dependent on governmental transfers after they run out of their savings, because they have expected shorter life.
The other possible effect on economy presented by Mirowsky (1999) comes from the fact that people feel that with every year lived they are not closer to death by the whole year which means that they are stretching their longevity with every year of their life. That means that they are not spending all their saved money on the contrary they are keeping the money away from the cycle which is from Keynesian's point of view slowing the economy.
How we have pointed out, longevity has increased significantly during the past more than a century (Eggleston & Fuchs, 2013; Seifarth et al., 2012) but people did not realize it in the whole picture (Miller Jr & Frech, 2000; Mirowsky, 1999; Mirowsky & Ross, 2013) . We can easily see the consequences of people and government not realising the increasement in the life expectancy. If people are saving money for their retirement that will be longer than they expect or if government is letting people go to the pension too early it will mean in a long-run lack of resources, which will display in form of inbalanced state financing. But what does it mean? What bad can it cause?

2.4 Effect of increasing life expectancy

There are many possibilities what can longer longevity cause in the long run. Ashraf et al., (2013) are pointing out that longer life expectancy can cause overpopulation which may lead to lack of resources that will eventually con-

duct in economic decline and ultimately again shorter life expectancy.
Other issue connected with overpopulation is mentioned by Eggleston & Fuchs (2013) who are saying that longer life expectancy can cause over-saturated job market which will sooner or later emerge in the situation in which young people without experiences will be unemployable or that older people with a lot of experience demanding higher salaries will be fired and nobody is going to employ them again even though they would like to work.
Those problems are just the simplest possible scenarios caused by not realized life expectancy. This means that it is necessary not only to realize longer longevity on governmental level but it is important to ensure that people are aware of the fact of rising life expectancy. Many of the authors are aware of this matter and are paying some attention to a solution. One of the solutions for this problem is almost identic across different papers and it is: better education. (Ashraf et al., 2013; Asiskovitch, 2010; Bulled & Sosis, 2010; Eggleston & Fuchs, 2013; Chan & Kamala Devi, 2012; Lin et al., 2012; Miller Jr & Frech, 2000)
Education as a structural solution is preferred by Ashraf et al. (2013) because more schools and teachers are creating conditions to educate more people to become better part of society and it helps them to gain higher status in our civilization. For those people it is easier to find use for their knowledge and skills and can become independent more easily. On the contrary we have to be aware of the phenomenon of overeducating people.The number of jobs with low education requirements has recently decreased same as the number of low educated who thanks to the new options gained higher education. Unfortunately because of the lack of positions with suitable educational requirements for the higher educated people they are competing with the lower educated for their position. In this competitions the better educated are most likely to get those low-paying jobs meaning the low educated people do not have assertion anymore (Aberg, 2014) . On the other hand thanks to the fact those people have more education than necessary, they are often counterproductive in their position (Tsang & Levin, 1990) . So we have to realize that even education can be double-barrelled.
Other problem with education is that better education is increasing health, which is mapped in higher life expectancy. At the same time higher life expectancy usually causes lower fertility because people are postponing their parental function with belief of longer life (Bulled & Sosis, 2010) and this can once again create an effect of aging population.

We can see that effects of life expectancy do not have to be clearly positive (Ashraf et al., 2013; Eggleston & Fuchs, 2013) and solution in form of education (Ashraf et al., 2013; Asiskovitch, 2010; Bulled & Sosis, 2010; Eggleston & Fuchs, 2013; Chan & Kamala Devi, 2012; Lin et al., 2012; Miller Jr & Frech, 2000) does not have to have the required effect. We must proceed with caution in structurally utilizing increasing longevity.

2.5 Previous research

As it was said, life expectancy has been in the spotlight of the research for quite some time. But it is necessary to understand that not everybody looks at this topic from the same perspective.

The first study that was aiming for finding the determinants of life expectancy was the work of Auster, Leveson & Sarachek (1969), who thought of two categories determining the life expectancy-environmental (socio-economic) and health care factors. Their main conclusion was that the environment is more important for increasing or decreasing the life expectancy. They have marked household income as one of the most important variables. The interesting fact is that the life expectancy is increasing with increasing income until a certain level where people become too rich to keep the healthy habits and the life expectancy starts dropping.

Other impressive study that was aiming for finding the main determinants of life expectancy was done by Mazudmar (2001). His work is exceptional thanks to the extensive dataset and general factors he has been studying for example the level of sanitation, availability of drinkable water or literacy, which he considered the most important factor together with the number of health care facilities. Unfortunately he did not go into more depth to find out which kinds of health care facilities are the most important ones.

Except the environmental factors in the socio-economic sense in our study we have been working with the variables expressing the environmental pollution and the level of fighting with it. Even though there have been studies focusing on the way how pollution affects the life expectancy, our study is the first to include such factors in more complex models and trying to find out the best variables possible to describe the pollution. We can mention one study for all. Pope, Ezzati & Dockery (2009) found out that in regions with lower air pollution in the US the increase in life expectancy have been in recent year higher than in the regions with high air pollution.

As the last cluster of studies focusing on life expectancy we can mark the category aiming at health care.

The special focus has been paid to the pharmaceuticals consumption that has been increasing in recent years (Berndt, 2002) together along with the life expectancy. Most of those studies *"Peltzman (1987), Babazono and Hillman (1994), Lichtenberg (1996, 1998), Frech and Miller (1999), and Miller and Frech (2000)"* (Shaw et. al., p. 769, 2005) found a link between the increasing pharmaceuticals consumption and life expectancy. But there are still studies that are pointing out the problem with this conclusion.

There are so many illnesses and so many different drugs that to be able to find a connection between increasing consumption and increasing life expectancy we would have to study consumption of some particular type of drugs on a precise sickness. Until that is done we are not able to make a general conclusion. (Grootendorst, Piérard & Shim; 2009)

The last study I would like to highlight is the study of Shaw et. al. (2005) . This study served as a core of our study and we are going to talk about it more in the next chapter but we have to at least introduce his study here. Shaw et. al. were modeling the situation of life expectancy in the 19 out of 30 OECD countries and came to a final model where the special importance was found in inclusion of variables of pharmaceuticals consumption and the percentage of people over 65 in the population.Nevertheless, we are going to present more detailed information in the following chapter.

.

Chapter 3

Model & Methodology

3.1 Theoretical introduction

Let us repeat our research question:

What factors are influencing life expectancy in the Czech Republic nowadays at the beginning of the 21th century and how we can improve them.

Preceding topics were necessary to understand the topic we are going to research in our work, but they were not the backbone of our work. The core of our study is looking for factors influencing life expectancy in the 21th century and trying to find out possible ways how to increase longevity in the Czech Republic even more. Retzlaff-Roberts et al., (2004) point out that there are two ways we can follow, either we can increase life expectancy with present investment by streamlining health care system or we can keep the level of the present health care and decrease present investment. In that scenario we will be able to use free financial funds to increase the life expectancy through different ways. His theory is based on the fact that money is not invested in the most efficient way possible so there is place to improve status quo.
During the 20th century we have witnessed the biggest increase in life expectancy (Eggleston & Fuchs, 2013; Mackenbach, 2013) thanks to the change in the core of the society-democratization, hygiene, economic growth-but in recent years it has become more and more difficult to improve health standards and life expectancy. Retzlaff-Roberts et al., (2004) state that the biggest improvement in life expectancy has been, in recent years, caused by improvement in health care system, which is exactly in the center of our interest.
We can suppose that improvement in health care system means increase in

the number of doctors or more precisely said less people for one doctor, more money (adjusted for inflation) for pharmaceutical and etc. Those factors are the ones that play the main role in the quality of health care system in the developed world and which we are going to examine.
So we are getting to the core of this work, the model. Let us look at how we have built up our model. First of all we are going to present the model by Shaw et al., (2005), which was our alpha and omega. Even though his model is very complex we were able to go deeper. Thanks to the fact we are modeling the situation in one country we are getting more reliable results because our sample is more homogeneous. This allowed us to drop some factors that have to be controlled while comparing different countries but not so much in one homogenous area. Lowering the number of explanatory variables is the reason why we could include other different variables that showed to be significant and important explanatory variables. Not only that our study is the first work on the topic of the life expectancy for the Czech Republic, it is the first one to look closer at the environmental polution as a part of a complex model of life expectancy.

3.2 Original model

Shaw et al. (2005) start their work with explanation why life expectancy is such an important topic to study, their trace of thought goes in the same directions as ours although they do not respect the problem in the whole range as we did in the beginning of our work. Their model is based on comparison of life expectancy among 19 OECD countries that have approximately the same level of well-being. By collecting many different variables in the end they come up with a model as follows:

$$\begin{aligned} lnLE97_{ij} = {} & \beta_0 + \beta_1 lnGDP85_i + \beta_2 lnPHARM85_i + \beta_3 lnHEALTH85_i \\ & + \beta_4 lnAGEDIST85_i + \beta_5 lnALCOHOL80_i + \beta_6 lnSMOKE80_{6i} \\ & + \beta_7 lnBUTTER80_i + \beta_8 lnVEG80_i + \beta_9 SPAIN_i + u_i \end{aligned}$$

Where GDP85 is the GDP in each of the countries, PHARM85 is the consumption of pharmaceutical in the 1985 in each of the countries, HEALTH85 is the money invested in the health care systenm except the pharmaceuticals in the year 1985 in each of the countries, AGEDIST1985 the percentage of people

over the age of 65 in each of the countries, ALCOHOL80, SMOKE80, BUTTER80 and VEG80 are consumption of alcohol, tabacco, butter and vegetable respectively in the year of 1980 in each of the country and SPAIN is the dummy variable for SPAIN where some of the variables were missing.
Let us look at different types of variables Shaw et al. (2005) use. In their paper they uses various categories of variables. We could characterize their variables in few groups: factors connected to health care system through expenditure and financing, demographic factors characterizing the country and then lifestyle factors describing the particular society living within one country. All of those categories are pretty straightforward with simple reasoning why they are included in the equation. They are expecting different factors to have a cumulative effect over time."*That is, the consumption of factors over time by an individual will have either positive or negative effects on that individual's longevity.*" (Shaw et. al., p. 771, 2005)
Shaw et. al. got two main findings. They found a possitive relationship between consumption of pahrmaceuticals and life expectancy but at the same time they are pointing out that it is sensitive to the age distribution of a given country, which is a reason for inclusion of the variable lnAGEDIST85. The seoncd finding is an importance of the life-style factors - lowering the amount of bad habits (such as smoking) means increase of life expectancy and increase of healthy habits (such as fruit and veetable consumption) means increase in life expectacny. There are few points I would like to higlight about the model of Shaw et. al. (2005) because they will be necessary for our own model or have a deeper meaning for understanding the differences in his and our model.

- First point I would like to stretch from this model is the pick of dependent variable. They have decided not to include life expectancy at birth, which has been used by some of the authors, for simple reason. Child right after birth is mainly influenced by the health conditions of the mother and her healthy or unhealthy habits. That way it would be very complicated to find proper explanatory variable for life expectancy of a child because we would have to collect data from the mother. Because he is not interested in this he has decided not to include life expectancy at birth but only at 40, 60 and 65 years.

- Second issue is connected to lagging variables that are not expected to have immediate results but rather cumulative results. Same as Lin et. al. (2012) are mentioning that democracy does not have an immediate result

on contrary after 3 years it has very little effect but it is growing, Shaw et al. (2005) expected some of the variables to show some effect later on. They have decided lag some variables by 12 and others by 17 years. They are mentioning that there is little academic evidence for the right amount of years and the main reason for the precise amount is usually the dataset we possess.

- Other interesting phenomenon highlighted in their work is so-called Sisyphus syndrome. We can understand this phenomenon as existing dual correlation between life expectancy and consumption of pharmaceuticals. Because higher consumption of pharmaceutical may mean longer life expectancy in the same manner as longer life expectancy can mean higher pharmaceutical consumption because of problems connected with age. Where problem of endogenity may appear. For this purpose Shaw et. al. (2005) introduced variable representing percentage of people over 65 in the population which is commonly agreed as end of productive phase of life.

- Last point I would like to go through before talking about concrete factors influencing the life expectancy is the influence of different genders. Many authors are pointing out the fact that women live on average longer than men (Asiskovitch, 2010; Seifarth et al., 2012) , this phenomenon has been studied by many academicians and even though this gap has been closing in recent years (Asiskovitch, 2010) it is still important to take care of this issue. When we look at the Shaw et al. (2005) model we can see how they cope with this issue by dummy variable, which allowed him to pool the data in the end across age and gender. So he came up in the end with more complicated model which meaning is not different from the model above.

3.3 Our model

So let us introduce our model:

$$lnLE_{ij} = intercept + SocioEconomicalFactors + HealthCareFactors + EnvironmentalPollutionFacotrs + disturbances$$

The main highlight about our model that has to be made in contrast to the previous model is that our model is expressing situation within one country. The same as in the previous model we are not going to model situation for the life expectancy at birth. The reasoning is the same as in the original work.
Lagging of variables is a painful issue in our case. The reason for that is short existence of the Czech Republic and even shorter existence of districts, as we know them now. Which means short period of consistent data. Other reason for complication is calculation of life expectancy for four years intervals instead for each year individually. The reason for longer period over which the life expectancy is calculated lies in the aim of exclusion incidental phenomenon that could have influenced only small area in one year. The rule is: smaller the region longer the period over which life expectancy is calculated. In our model we use two sources of data - Czech statistical office and Institute of Health Information and Statistics of the Czech Republic - which databases unfortunately do not start in the same year, so for the best interpretation of our results we have chosen one static lag of 9 years. Reasoning behind this decision is simple, as we saw in case of Shaw et. al. (2005) he has been laging by 12 or 17 years, because the literature sudgests to lag by 15 to 20 years. Unfortunatelly thanks to the short history of the Czech Republic and its institutions we were not able to get data older than 12 for some variables and 9 years for other variables. However we did not find any reasoning for lagging by two different periods so we have decided to lag every explanatory variable only by 9 years, which is the common longest lag possible.
We have adopted the solution to the Sisyphus syndrome in the full range because it is the most logical and the simplest key to the problem and the statistic of 65+ is monitored.
The last point highlighted in the work of Shaw et al. (2005) was the gender inequality as for length of life. Many authors focusing on the life expectancy has come across the phenomenon of differences in life expectancy among male and female (Asiskovitch, 2010; Okamoto, 2006; Seifarth et al., 2012) . Even though all of them found out the proof that women have on average longer lives men still think that they have longer life (J Mirowsky, 1999) . This is caused by the fact that man seemed to be usually healthier until an older age and they have on average higher socio-economic status, which is giving them confidence. Nevertheless the gap between longevity of male and female is still there but it is shrinking. How Asiskovitch (2010) points out in recent years women have been closing the gap in longevity among men and women because

Table 3.1: Explanatory variables

variable	variable shortcut	variable measurement
Unemployment rate	U	percentage
Average wage	AW	Czech Koruna
Amount of poisonous oxides in the environment	PoiOx	ton issued per year
Solid	Solid	ton issued per year
Investment into the environment	InvEnv	thousands of Czech Koruna
Number of eople per doctor	NPpDoc	number
Number of hospitals	NH	number
Number of recipes per person	NRepP	number
Number of people per pharmacy	NPpPhr	number
Revenue from pharmaceuticals	RPha	Czech Koruna
A percentage of male at the age and over the age of 65	M65+	percentage
A percentage of female at the age and over the age of 65	F65+	percentage

Source: www.czsu.cz; www.uzis.cz)

they have begun to gamble with their health more and more.

3.4 Explanatory variables

We can define three categories of factors that can be seen in our work. They are: environmental (socioe-conomic), health care and enviromental pollution factors. Now let us look at each of this category separately and explain the intention behind it.

Demographic factors in which I am including socio-economic factors is a category containing indicator of present situation in the society. There are many possible factors. The most often mentioned indicator from this category is GDP (Ashraf et al., 2013; Eggleston & Fuchs, 2013; Leung & Wang, 2010; Walker, 2013) . Unfortunately there does not exist any variation of GDP on district level at least not in the Czech Republic so we had to look at what GDP is supposed to control in this model. Shortly said GDP describes the wealth of a nation and there is an implication that richer the country the more they can invest in the health care system which would positively influence the life expectancy.

Blackman (2011) points out that GDP is not necessarily the best measure of wealth of a country and wellbeing of people living there. The reason for poorness of GDP as an indicator of a wellbeing is the lack of equality in distribution of wealth among the whole population. That is the reason why we have decided to use the unemployment rate and average wage instead of some form of district GDP in our case. We know that average wage is again likely to be influenced by not equal distribution of wealth among the whole population but lower income group at least dilute the exceptionally high wages of the few at the top. Because only average wage did not seem strong enough we have decided to try to add the unemployment rate as well. Support for substitution of a GDP by unemployment can be found in texts of Bray (2013) and Sogner (2001) where the authors comment on close links between those two indicators.
Other group of factors that I have called health-care factors in the beginning of this part is the center of our interest. We understand that factors from other categories can and have been influencing life expectancy but the goal of our study is mainly to find factors connected to health care system that have an influence on life expectancy. Reason for special interest in health-care factors can be justified by the fact that it might be the easiest to change redistribution within existing health care system. Changing economic position of a region or environmental pollution may be harder than increasing the number of doctors or encouraging consumption of pharmaceuticals.
Because this is core of our study we had to pick those factors very carefully. Fortunately we had a very extended data set which did not place any constraints on us in the first place but the limited number of districts gave us reason to choose factors wisely so we would not loose so many degrees of freedom.
In connection to the health-care factors I would like to stretch one other fact. The original intension was to include investment in hospitals or even in outpatient care. Unfortunately access to those data is restricted and we did not get the chance to create as complex model. This is our main recommendation for a future research but let us get to it in the conclusion.
For now we will present you a quick overview of factors used, the complete list with explanations is included in Appendix A.0.5. First of all we use "number of people per one doctor" (Retzlaff-Roberts et al., 2004) as one of the factors, because less people mean more time for their each patient. Other indicator mentioned in the work of Miller Jr & Frech (2000) is consumption of pharmaceuticals. They put an equal sign between higher pharmaceutical consumption and longer life expectancy and we will try to find out if the same can be said

about the Czech Republic. Of course people need medical care when something happens to them but a really good way to prevent any illness or injury is to adjust one's lifestyle to healhier one.Other factors that we tried to include were "number of people per hospitals, number of people per prescriptions, number of people per one pharmacists or number of hospital's bed per 10000 people"
Tamakoshi et al. (2010) and Miller Jr & Frech (2000) point out that life expectancy is often influenced by habits that people have. The worst of them is smoking. Other factors that can characterize life style can be found in Shaw et al. (2005) where they introduce consumption of fruit and vegetable (source of vitamins) and consumption of fat. Unfortunately even though those indicators are well-founded we do not have such statistics. We have put an email from the head of information services of the Czech statistical office explaining the lack of those indicators on district level in the Appebdix (Appendix A.0.6). For the reason above we have dropped the whole category of lifestyle factors and consider the life-style among the inhabitants of the Czech Republic to be the same on average.
However there is one factor that had been usually omitted and we see importance in including it and it is environmental pollution. Our study is a first one that examine the influence of the environmental pollution onto the life expectrancy in such a depth. Every one of us is a part of our society and is shaping his or her surrounding in some way, just like the whole society does. People are aware of the fact that air or water pollution is serious issues that can be fatal. This is the reason for societies, especially the developed, to invest in keeping the environment as it is or even cleaning the pollution from the past.
There are two main indicators that can be observed in connection with environmental pollution. First of all we can measure the amount of chemicals in the air or water or we can calculate how much money has been invested in preserving nature being destroyed. As Pautrel (2009) points out in his work that environmental pollution does shorten the average longevity of people but at the same time focus on environmental pollution can cause decline of economic growth by sloving down the industry by different restrictions. This fact implies that environmental factors can be the key element to consider. On one hand we can increase the economic growth through preserving the nature and increasing the longevity on the other hand exaggerated attention to the environment can be fogyish for the simple reason of slowing the economic expansion.
Because we found out environmental pollution factors to be so crucial we have decided to add them in our model even though we have not found a precedent

in form of life expectancy model including environmental pollution factors.
So we have gone through what life expectancy is. We know how has it has changed and what consequences it has. We have realized that even though life expectancy is very important in making life-changing decision people are not aware of its function. There has been quite few academicians observing life expectancy and trying to find the best possible model. Due to the fact that Shaw et al. (2005) were aiming for the same goal as we do, we have accepted his model as our bases and on top of that made important changes. The biggest differences are exclusion of lifestyle, which was possible thanks to the homogenity of our data due to the fact that data were collected within one country, and inclusion of environmental pollution factors. Those changes should give us the favourable edge to find factors influencing life expectancy in more precise way.

3.5 OLS

Now we are going to introduce our method of modeling we have used, reason why we picked this method and what are its benefits and disadvantages.
Even though we had the possibility to choose from variety of explanatory variables we had to bear in mind the limiting fact of finite sample, which forced us to limit ourselves in choosing the explanatory variables.
Now I would like to focus on the statistical/econometrical method we have used and it is Ordinary least squares. OLS is the most frequently used method due to its simplicity and apprehensibility. The logic behind this method is easily understandable and interpretable, which assure easy explanation of results.
On the other hand clarity of the whole OLS method is drawback at the same time. It can be misused in a situation when it is not the most suitable option. There are some limitations to which we should pay attention before applying this method on our data.
The first problem appears at the beginning with collecting the data. Thanks to the nature of the OLS and how it works (minimizing sum of squared errors) we have to look at all of our data and think about the possibility that some of our observations may be outliers. In our case this outlier can be Prague in many different points of views.
Other attribute of the OLS is linearity. OLS is a linear regression which means that can model linear relationships but in case of non-linear relationship we will get very poor results. Very common way of dealing with this issue is trans-

forming the collected data into linear form. The most common way is to use the logarithm.
The second benefit of using logarithms is the simplicity of interpretation. Which allows to interpret the difference of two values as change in percentage.
Those two drawbacks of the OLS are not the only thing we have to pay attention to when applying this widely used method. OLS in its basics is standing on few prerequisites that have to be fulfilled or controlled for before applying the OLS.
Let us look at our final model and its results, afterwards we will go through those so called assumption and see that all of them are fulfilled in our model, namely: linearity in parametrs, random sampling, no perfect collinearity, zero conditional mean, homoskedasticity, normality.

3.6 Results

To estimate an ideal model was not such an easy task as in the case of Shaw et. al. (2005) . They in their paper came to a conclusion that they were able first of all to pool the data across gender, even though as we saw earlier many studies had studied the phenomenon of differences between genders, and even pool the data within the gender across time.
In our case after considering various possibilities and estimating different types of models we estimated four different models, where three of them are quite similar.
Models and results are:

For male at the age 45

$$lnLEM45 = \beta_0 + \beta_1 NPpDoc + \beta_2 PoiOx + \beta_3 U + \beta_4 M65 + u_i \tag{3.1}$$

For male at the age 65

$$lnLEM65 = \beta_0 + \beta_1 NPpDoc + \beta_2 PoiOx + \beta_3 U + \beta_4 M65 + u_i \tag{3.2}$$

For female at the age 45

$$lnLEF45 = \beta_0 + \beta_1 NPpDoc + \beta_2 PoiOx + \beta_3 U + \beta_4 F65 + u_i \tag{3.3}$$

For female at the age 65

$$lnLEF65 = \beta_0 + \beta_1 RPha + \beta_2 PoiOx + \beta_3 InvEnv + \beta_4 F65 + u_i \qquad (3.4)$$

Table 3.2: Dep = lLE45M

Variable	**Coefficient** (Std. Err.)
lPoiOx	-0.006** (0.002)
lM65r	0.107** (0.021)
lNPpDoc	-0.020** (0.007)
lU	-0.036** (0.005)
Intercept	3.423** (0.086)
N	77
R^2	0.753
$F_{(4,72)}$	54.81

Table 3.3: Dep = lLE65M

Variable	**Coefficient** (Std. Err.)
lPoiOx	-0.008* (0.003)
lM65r	0.124** (0.036)
lNPpDoc	-0.046** (0.012)
lU	-0.033** (0.009)
Intercept	2.828** (0.147)
N	77
R^2	0.585
$F_{(4,72)}$	25.391

Table 3.4: Dep = lLE45F

Variable	**Coefficient** (Std. Err.)
lPoiOx	-0.006** (0.002)
lF65r	0.093** (0.022)
lNPpDoc	-0.013† (0.007)
lU	-0.013** (0.005)
Intercept	3.481** (0.091)
N	77
R^2	0.521
$F_{(4,72)}$	19.568

Table 3.5: Dep = lLE65F

Variable	**Coefficient** (Std. Err.)
lPoiOx	-0.012** (0.004)
lF65r	0.127** (0.037)
lRPhapP	0.028* (0.013)
lInvEnv	0.008* (0.003)
Intercept	2.341** (0.114)
N	77
R^2	0.398
$F_{(4,72)}$	11.893

Table 3.6: Test's results

test	M45	M65	F45	F65
Variance inflation factor (Mean VIF)	1,41	1,41	1,27	1,29
Cameron & Trivedi's decomposition of IM-test (total p-value)	0.6525	0.2802	0.6252	0.6474
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity ($Prob > \chi^2$)	0.2047	0.4286	0.6650	0.2235
Shapiro-Wilk W test for normal data ($prob > z$)	0.43036	0.39701	0.93985	0.81736

3.7 Testing the model

We will focus on interpreting the results in the next chapter. For now let us look at the econometrical formality that a model has to fulfil before we can say whether the model is valid.

It may be a little harder to test everything for four models so if needed to show something we are going to indicate each model by its dependent variable (LE45M, LE65M, LE45F, LE65F).

Before starting with testing our model from various perspectives it is useful to look at the data on its own. In the previous chapter we have been talking about outliers, points that have exceptionally big value of independent variable considered the value of independent variables, those points can very easily change the whole model thanks to the nature of OLS. That is the reason why looking for those outliers is going to be our first goal. First of all we can use the easiest way and that is visual observation. If we wanted to we could list all the values of those residuals and see which ones are oddly big or small.

Other type of such points is so called leverage point; this point has high values of independent variable, so it lies once again far away from our points set but in the other direction than the outliers. Luckily Stata (Data analysis and statistical software) can measure leverage for each of our observations, which means we can easily look at which distance is abnormally big.

The most dangerous points are those that are outliers and leverage points at the same time. Once again we can observe this visually from the picture in Appendix. (Appendix: Figure A.1 - A.4)

The worst point in all of the four cases seems to be "Prague-West", however because this district is not exceptional from any point of view and the statistics for outliers and leverage are not as odd as in other extreme cases we have decided not to drop this observation.

Now we shall continue with testing for fulfilling the assumption that are ensuring that our estimates are not going to be only BLUE (best linear unbiased estimators) but we are looking for fulfilling one extra assumption. If this one assumptions are satisfied we can also use the law of large numbers. So lets continue.

3.7.1 B.L.U.E.

Let us start with making sure that our estimates are BLUE. We are going to test for those assumptions: linearity in parametrs, random sampling, no perfect collinearity, homoskedasticity, normality.
We are not going to test for zero conditional mean because this assumption is expected to be fulfiled from the nature of OLS.
First Gaus-Markow assumption is in connection with the second issue I have mentioned - linearity. Because OLS is linear regression we want the relationship between dependent and independent variable to be linear. One thing is to assume it; the other is to test it. Unfortunately in the case of linearity it is hard to find appropriate test, so the only way how we can judge if the model is linear or not is graphical way. Because in our case the data are not exactly linear we have decided to use logarithms. Logarithms are not just good to help make our data more linear, but difference of logarithms can be very easily interpreted as a percentage change, which is convenient in interpretation phase.
Second assumption is one that is in our case impossible to fulfil. Our study is focusing on a phenomenon of "life expectancy" within one country, we believe that this is not going to be an issue for the reason that we collected data from all districts within the Czech Republic, we do not actually need random sampling because we have "all possible" observations.
It is also important look at the independent variables on their own. An issue can easily arise if some variables are highly correlated; in our case of multiple regressions this potential problem is called multicollinearity. Fortunately we can use the variance inflation factor VIF that is index measuring how the variance of an estimate change because of some collinearity. In our case we have not found any proof for multicollinearity so we can continue without any hesitation.
The last assumption out of the basic five assuring that the estimators are BLUE is homoscedasticity which is saying that the variance of residuals does

not change with the observations. Because homoscedasticity is one of the most important assumptions we have to pay special attention to it. Once again the first, quick help is to look at the graphical expression (see Appendix: Figure A.5 - A.8). This can give us the first clue what we are dealing with. Other way how to make sure that we do not have to be afraid of heteroscedasticity is to use one of the following tests White's test or Breusch-Pagan test. Luckily both of them ended up in our favor in all of our cases.

3.7.2 CLT and LLN

At this point we can confidently state that our model is fulfilling five basic assumptions of OLS, which means that our estimates are BLUE. As we anticipated at the beginning of this chapter we need to test one more assumption to make sure that we can use the Central limit theorem (CLT) as well as the Law of large numbers (LLN).

The assumption we are talking about is assumption of normality of residuals. Luckily for us if we test for the normality of residuals in graphical way (see Appendix: Figure A.9 - A.12) or numerical form we always get to the same result that this assumption is fulfilled and we can easily test our hypothesis as we need.

At this moment we have our final models and as a goals of the next chapters are to interpret the results and comment on the differences among our own models as well as look at them from a bigger perspective and compare them to the model of Shaw et. al. (2005) and explain the possible reasons for the differences we came across.

Chapter 4

Interpretation of the results

In this part we would like to look at the results, make sure it is clear what they present and how each of the explanatory variables influence the dependent variable in form of the life expectancy. It is important to highlight the fact that we are going to interpret the influnce of only significant variables. For the complete list of explanatory variables, which we have tried you can see Appendix A.0.4.
Because the first two models for male at the age of 45 and 65 are exactly the same and the model for women at the age of 45 has only one different variable - instead of percentage male over 65 years old we are considering female with the same specification - we have decided to interpret those models together and comment on the results as follows - first result from the model of M45, second M65 and as last F45. The last model, that is more different, we are going to interpret separately at the end in a separate part.
For the simplicity let us sum up the coefficients from all of the four models in one synoptic table.

We will start gradually from the smallest to the biggest coefficient. The smallest influence has "amount of poisonous oxides into the environment". The coefficients of -0.006; -0.008 and -0.006 respectively are stating that if we increase the pollution of the environment by poisonous oxides by 10% our life expectancy will decrease by 0.06%; 0.08% or 0.06% which means on average for male at the age of 45 the life expectancy would decrease by approximately one week (average LE is 31 years), for male at the age 65 it would be 4 days (average LE is 15.1) and for female at the age of 45 it would be by 8 days (average LE is 36.3). Even though those numbers seem small enough to be ignore we have to relate them to the amount of poisonous oxides issued in the environment. The

Table 4.1: Coefiicient's overview

Results				
variable	M45	M65	F45	F65
logU	-0.036	-0.033 -0.013	x	
logPoiOx	-0.006	-0.008	-0.006	-0.012
logInvEnv	x	x	x	0.008
logNPpDoc	-0.020	-0.046	-0.013	x
logRPhapP	x	x	x	0.028
logM65	0.107	0.124	x	x
logF65	x	x	0.093	0.127
intercept	3.423	2.828	3.481	2.341

average amount is 8580.8 t/km^2 but the maximum value in "Ostrava-city" is 107727.8, which means it is more than 12 times the average. This means that if every districts was issuing the same amount as "Ostrava-city" our LE would decrease by 84 days, 48 days or 96 days respectively. This is a big change.
It is impossible to distinguish which is the second and the third most influencing factor in all those three models because the role is not the same. We will start with the unemployment rate. "Unemployment rate" is a replacement for GDP in the national model and it is supposed to represent the economic situation in the particular region. In our case the effect of logU are -0.036; -0.033; -0.013 respectively. The interpretation in this case is as follows. Average unemployment rate was 10.88% so if we take this as a unit and we increase this unit by 1%we will get the unemployment rate of 10.9888%. In this situation the life expectancy decreases by 0.036%; 0.033% or 0.013% respectively. If we want to see the connections with the amount of days it is 4; 1.8 and 1.7 days. And this was an increase in unemployment rate by only 0.1088%. We can easily conclude that unemployment rate can have disastrous consequences on LE.
The third variable which coefficient we are going to interpret is "number of people per one doctor". As we can see above the coefficients are -0.020; -0.046 and -0.013. Once again we are using log transformation but because the basis is in numbers, the interpretation is easier than in case of unemployment rate. If we increased the number of people per one doctor by 10 percent we would get decrease on the LE by 0.2%; 0.46% and 0.13% respectively. Once again if we recall the average LEs we would get that LE would decrease in form of days by 23; 25 and 17 days. To put a picture behind those 10% increase of people per one doctor we should say that the average number of people per one doctor is 323 (where the minimum value is 123 and the maximum is 651).

The last and the most influential variable it seems to be "percentage of male/female over 65" in the population. The coefficients speak for themselves: 0.107; 0.124 and 0.93. Those coefficients have the biggest influence and the most disturbing influence at the same time. Let us look in this chapter only at the statistical interpretation. Because this variable is once again in the percentage form we will use the same interpretation as with the unemployment rate. For male the average percentage is 10.9% and for female it is 16.4%. Let me talk about the males and I will put the results for females into brackets so. If we consider 10.9 (16.4) as a unit and we increase this unit by one percent calculated from this unit, we will get 11.009% (16.564%). But what will happen with LE? In the case of male at the age of 45 it will increase by 0.12% which means by 12 days in case of male at the age of 65 it is nearly half - 7 days - and in the case of female at the age of 45 it is the same as for male at the same age - 12 days.
Now when we have interpreted the three models that have basically the same explanatory variables, except the percentage of people over 65% according to the gender, we will continue with the last remaining model for female at the age of 65.
We will start in the same manner with "amount of poisonous oxides" issued into the environment. As we can see above the coefficient is -0.012. So if we increase the amount of poisonous oxides issued into the environment by ten 10% we will get 0.12% decrease in LE, which corresponds to 8 days.
Unemployment rate as an economical variable does not seem to have any influence on female at the age of 65, instead of that we can observe bigger influence of environment. One of the variables that were not significant in previous models is "investments into the environment". Coefficient in front of this variable is 0.008. It is saying if we increase the amount of investments into the environment by 10% (average "investment into the environment" were 193 746 and excluding Prague - as an exceptionally high value - 170 314) our LE increases by 0.08% or 5 days.
Instead of "number of people per one doctor" we have in this mode as a health indicator "revenue from pharmaceuticals per person". The effect is on average the same as the effect of "number of people per one doctor" in the previous three models. Concretely the coefficient is 0.028. Once again simple interpretation saying that 10% increase of revenue from pharmaceuticals per one person (average 3667, with values higher in the biggest cities) which is approximately 367 CZK means LE longer by 19 days.
The last and once again the most influential variable is "percentage of female

over 65" in the population. The value of 0.127 means that if we take the average percentage of female in the population as a unit (16.4) and we increase it by one percent (0.164) we get 16.564% and the consequences on LE will be in form of 0.127% or 17 days.
Now there are some issues connected not only straight with our models but differences in the results of our and previous works that should be remit to analysis in the next chapter.

Chapter 5

Discussion

5.1 Advantages of our model

It is important to realize that there have been many different studies which aim was to determine factors influencing life expectancy or mortality. *"Auster, Leveson and Sarchek (1969) were the first economists to study a population production function for health"* (Shaw et. al., p. 768, 2005) . This citation shows how long period it is since the first study on this topic. It is almost a half a century and as I indicated earlier it is half a century that meant the biggest change in the life expectancy of the human history (Walker, 2013) .
Many things have changed in the field of various factors during this half a century-life style, economic, political, informational or health factors. So it is not as unexpected that all of those models conducted during this period had different outcomes. They all used the same categories of factors and they were all international models. This is the reason why we believe our model is special in the area of one country, concretely the Czech Republic. Country that went through changes of political ordering, which as we talked at the beginning may be an important deterministic factor (Lin et al., 2012 and Mackenbach, Hu, & Looman, 2013) . The second thing that our model stands above the previous models is the inclusion of environmental pollution factors and not only environmental factors in general as political, economic etc. environments but we have included the state of the air and water pollution as well as money invested in fighting with those issues. And not so surprisingly we have found out that those factors are significant in our models.
The first point that has to be highlighted about our model is omission of life style variables. This category has been often present in the previous model

focusing on life style expectancy and the model of Shaw et. al. (2005) was not an exception. However this fact can not be considered as a drawback of our model, this omission has a reasoning that lies in the way how those variables are calculated. Those factor can be calculated in two different ways, unfortunately both of those ways were not applicable in our case unless we wanted to present statistically indefensible data. As an explanation please find an answer of the supervisor of social surveys at the Czech statistical office Mr. Jaromir Kalmus in Appendix A.0.6.

5.2 Health care and its role

Asiskovitch (2010); Papavlassopulos & Keppler (2011) and Retzlaff-Roberts, Chang, & Rubin (2004) are three of the many studies highlighting the importance of health care on the life expectancy, the theory is that when a country has adopted basic hygiene manners the main way how to increase life expectancy even more is to streamline the health care. The original idea behind this study was to find which factors do influence the life expectancy the most with special focus on health care factors.

In the phase of collecting the data we have gathered quite a big amount of possible health care factors - number of hospitals, number of individual outpatient facilities, number of prahmacists, revenue from pharmaceuticals, number of people per one doctor etc. there are many other factors that could have been influencing life expectancy. And we have tested all of them to make sure to find the ones that do influence the life expectancy. At the end we came to a quite suprising conclusion. The only explanatory variable that showed to be significant was “number of people per one doctor” in three out of four cases and “revenue from pharmaceuticals” in the case of women at 65.

Those findings deserve an explanation. If we look at the significance of the “number of people per one doctor” we can come to a conclusion that doctors are the ground stones of the health care system. Hospitals or outpatient facilities are useless without doctors same as nobody can prescribe a prescriptions except the doctor. This means that the number of doctors is the essential number, any money invested in the health care system has to go through hands of a doctor or some doctor has a use of them. This proves the logical fact that doctors are the alfa and omega of the health care system. The lack of importance of other explanatory variables from category of health factors can be explained from the statistical point of view due to the high correalations of any other

factor with the "number of people per one doctor".
More interesting case is the life expectancy of women at the age of 65. The possible explanation of replacement of the variable "number of people per one doctor" by "revenue from pharmaceuticals" may lie in the text of Asiskovitch (p. 888, 2010) he is stating *"...women live longer than men and have lower risk of fatal health diseas."* This together with the fact that *"Women pursue more healthy life - on average, they smoke and drink less than men and are less involved in physically demanding and risky jobs."* Asiskovitch (p. 888, 2010) may mean that for women at the age 65 is care of a docto less important than for a man because they need just the right pharmaceuticals that they have been using for quite some time.
This preceeding argument explains the fact that we found the "revenue from pharmaceuticals" an insignificant factor compare to the study of Shaw et. al. (2005) who has been highlighting its importance. We may just wonder if he tried other health factors that may be even better than revenue of pharmaceuticals.

5.3 Enviromental pollution factors

The second surprise of our model is the importance of environmental polutionl factors. The preceding studies understood under the term "environmental factors" factors that are creating the whole political and economic environmentn not in the sense of ecological and natural environment. In our study we have decided to focus more on the ecological side of our environment. On the pollution and the fight with it and it payed back. We have found out that those factors are significantly important for explanation of the life expectancy. We can find backing for our theory in the Mariani et al., (2010) or Pautrel, (2009) who both state the environment is an important factor for life expectancy especially in the sense of investments in the environment.

5.4 Age distribution - positive or negative effect

The last point that should be made before moving to the next chapter is about the age distribution of the population. In all of four models we have included

an explanatory variable indicating the percentage of the population in and over the age of 65. We would not find anything strange about our results if we did not have the outcome of Shaw et. al. (2005) .
They found that the coefficients of age distribution are significant but they found the coefficients to be negative. It means that if the relative population of the people over the age of 65 to the whole population would increase the life expectancy of the whole population would decrease. Unfortunately Shaw et, al. (2005) do not explain why this is an expected outcome.
In our model we found coefficient in front of the same variable to be positive and it is our duty to give an explanation. We can find our explanation in the text of Crimmins (2004) says: *"The general picture is that older people of today are healthier than older people of two decades ago. There have been improvements in most dimensions of health. People live longer and have fewer disabilities, have less functioning loss, and report themselves to be in better health. Over time there has been some reduction in risk from smoking and a lowering of cholesterol and average triglyceride levels. However, weight increase has been notable during this period. Because people live longer, a greater percentage of people have some specific common diseases, and, on average, older people live with more diseases."* This short citation is pointing out the main differences between the older people of today and thirty years ago. Older people of today may have more diseases but they are generally in better shape and because those diseases are common there are usually efficient ways to get them under control.
We can even assume that the higher the frequency of appearance of fatal disease the more focus is paid to it. It means that those common diseases mentioned above are in the center of attention of the researchers.
Based on the facts above we can only wonder why Shaw et. al. (2005) found the negative coefficient of age distribution comforting and did not decide for deeper analysis.

5.5 Main findings

To summarize our model it is important to highligh four main findings.

- First of all, health care factors are very important but thanks to the interconnections among them it is enough to include only one, unfortunately

in most cases we were not able to find a significant influence of the consumption of pharmaceutical on the life expectancy, which is denying in three out of four cases our hypothesis H.1.

- Second of all different kinds of pollution and investments in fighting with them have not gained only the political strength during the past 20 year but it showed to be a significant factor determining the life expectancy of the whole population, which is a confirmation of our hypothesis H.2.

- Thirdly we found a difference in the way how the population of people 65 and older influence the society and thanks to previous research we are able to explain our findings in a logical manner. This means approval of our hypothesis H.3 that the amount of people over 65 in the population is important factor for life expectancy.

- The last but not the least we are finally able to come up with an idea for increasing the life expectancy. First of all we can pay more attention the the environment and try to keep it clean as much as possible. Other possibility indicated by the importance of the amount of people over 65 we should pay more attention to this group of people and their problems, solving their problems should increase implicitly the life expectancy. For some categories of people it is a good way to increase life expectancy through decreasing the unemployment rate but it does not work for everybody.

Chapter 6

Robustness test

There is one more thing to try before coming to a conclusion. Our models are solid, they have fulfilled all of the assumptions necessary for the OLS to work. Results seem reasonable and we were able to explain all the differences compare to the model of Shaw et. al. (2005). The only thing we are missing now is the robust test.

This is not an official test, there are hardly any rules to fulfill when we are running it. The aim of this test is to see if our model only holds in vacuum or if it works if we change the initial conditions and add some extra explanatory variables. So we are looking at the same model but we will add other variables and see what happens with the coefficients.

Our situation is a little more difficult due to the fact that we are running four models. To make it as clear as possible we are going to create four tables - for each of our original model one. In the first column of the table there are explanatory variables and the dependent variable is the variable indicated in the heading of each table. In each column of the table we added one variable to the original model and we will be able to see how the results have changed. In the last column we will estimate one model with all of the explanatory together. Because we had a lot of different variables and it would be very complicated to test all of them, we are going to pick one or two additional factors from each category of environmental (socio-economical), health care and environmental pollution factors and see how it will change.

Table 6.1: Robustness test LE45M

	LE45M					
lU	-0.0356271	-0.0360216	-0.0350928	-0.0352762	-0.0356364	-0.0359475
SE	(-0.0054298)	(-0.006605)	(-0.0055355)	(-0.0056165)	(-0.0054709)	(-0.0073087)
lPoiOx	-0.0056185	-0.0054865	-0.0060384	-0.0061515	-0.0056352	-0.0065471
SE	(-0.0019768)	(-0.002345)	(-0.0021186)	(-0.0027987)	(-0.002018)	(-0.0035317)
lNPpDoc	-0.0200022	-0.0203528	-0.0186644	-0.0203841	-0.0193127	-0.0167959
SE	(-0.0071423)	(-0.0079102)	(-0.0075506)	(-0.0073257)	(-0.0153864)	(-0.0160255)
lM65	0.1068915	0.1063434	0.1092842	0.1074188	0.1064547	0.1071193
SE	(-0.0209966)	(-0.0217599)	(-0.0215099)	(-0.0212225)	(-0.0228323)	(-0.0245746)
lAW		-0.0029893				-0.0119964
SE		(-0.0280776)				-0.0339684
lInvEnv			0.0012066			0.0018446
SE			(-0.0021183)			(-0.0024813)
lSolid				0.0014588		0.0020275
SE				(-0.0053862)		(-0.0058395)
lRpPhapP					0.0008474	0.0038094
SE					(0.0167185)	(-0.0176849)
cons	3.422541	3.454356	3.397386	3.417886	3.412865	3.461798
SE	(-0.0856641)	(-0.3110308)	(-0.0967367)	(-0.0879169)	(-0.2094809)	(-3717298)
R^2	0.7528	0.7528	0.7539	0.753	0.7528	0.7551
$adjustedR^2$	-0.739	-0.7354	-0.7366	-0.7356	-0.7354	-0.7262

Table 6.2: Robustness test LE65M

	LE65M					
lU	-0.0333348	-0.0328319	-0.0308204	-0.0333018	-0.0331082	-0.0333722
SE	(0.0093012)	(0.0113147)	(0.0093394)	(0.0096261)	(0.009337)	(0.0122906)
lPoiOx	-0.0077253	-0.0078936	-0.0097017	-0.0077755	-0.0073146	-0.0093962
SE	(0.0033862)	(0.0040171)	(0.0035744)	(0.0047967)	(0.003444)	(0.0059391)
lNPpDoc	-0.0463952	-0.0459482	-0.0401	-0.0464311	-0.0632717	-0.0537698
SE	(0.0122348)	(0.0135507)	(0.0127391)	(0.0125554)	(0.0262597)	(0.026949)
lM65	0.1240479	0.1247465	0.135308	0.1240975	0.1347403	0.1400674
SE	(0.035967)	(0.037276)	(0.036291)	(0.0363728)	(0.0389679)	(0.0413255)
lAW	0.0038102					-0.0282501
SE		(0.0480985)				(0.0571224)
lInvEnv			0.0056784			0.006771
SE			(0.003574)			(0.0041727)
lSolid				0.0001373		0.0022677
SE				(0.0092313)		(0.0098199)
lRpPhapP					-0.0207434	-0.0134877
SE					(0.0285332)	(0.0297395)
cons	2.827503	2.786951	2.709126	2.827065	3.064356	3.133785
SE	(0.1467422)	(0.5328132)	(0.1632116)	(0.1506787)	(0.357518)	(0.6251131)
R^2	0.5852	0.5852	0.5994	0.5852	0.5882	0.06039
$adjustedR^2$	0.5621	0.556	0.5712	0.556	0.5592	0.5573

Table 6.3: Robustness test LE45F

	LE45F					
lU	-0.0126644	-0.0145025	-0.0120322	-0.0124802	-0.0126024	-0.0162573
SE	(0.0046448)	(0.0056583)	(0.0046502)	(0.0048127)	(0.0046945)	(0.0061388)
lPoiOx	-0.0055056	-0.0048581	-0.0063726	-0.0058052	-0.0055587	-0.0058161
SE	(0.0018949)	(0.0022125)	(0.0020036)	(0.0026562)	(0.0019397)	(0.0032441)
lNPpDoc	-0.0126508	-0.0143842	-0.0101822	-0.0128688	-0.0107166	-0.0050967
SE	(0.0067156)	(0.0073912)	(0.0069565)	(0.006894)	(0.014431)	0.0147484
lF65	0.0926524	0.089884	0.0954668	0.0929838	0.0916471	0.0855989
SE	(0.0215433)	(0.0221743)	(0.0215586)	(0.0217866)	(0.0226806)	(0.0236074)
lAW		-0.0151783				-0.0430381
SE		(0.0264239)				(0.0314914)
lInvEnv			0.0025177			0.0042753
SE			(0.0019607)			(0.0022997)
lSolid				0.0008276		0.0012117
SE				(0.0051065)		(0.0053995)
lRpPhaP					0.002326	0.010341
SE					(0.0153372)	(0.0158815)
cons	3.480754	3.642941	3.435175	3.477918	3.453759	3.739104
SE	(0.0909071)	(0.2967535)	(0.0972128)	(0.093186)	(0.2000944)	(0.351415)
R^2	0.05209	0.5231	0.5317	0.521	0.521	0.5466
$adjustedR^2$	0.4943	0.4895	0.4988	0.4873	0.4873	0.4931

Table 6.4: Robustness test LE65F

	LE65F					
lU		-0.0116115				-0.0188174
SE		(0.008199)				(0.0108833)
lPoiOx	-0.0121421	-0.0110075	-0.0119805	-0.0119319	-0.0159165	-0.0105556
SE	(0.003542)	(0.0036076)	(0.0035752)	(0.0036427)	(0.0049047)	(0.0057514)
lNPpDoc			-0.0129887			-0.0165066
SE			(0.0259767)			(0.0261473)
lF65	0.1266814	0.1052391	0.1295073	0.1268028	0.1256457	0.0957435
SE	(0.0370663)	(0.0398024)	(0.0376871)	(0.0373083)	(0.0370182)	(0.0418533)
lAW				-0.0124875		-0.0676845
SE				(0.0444909)		(0.0558307)
lInvEnv	0.0076871	0.0072836	0.0073258	0.0081835	0.0085297	0.009621
SE	(0.0034278)	(0.003416)	(0.0035207)	(0.0038768)	(0.0035053)	(0.004077)
lSolid					(0.0099989)	(0.0042301)
SE					(0.0090034)	(0.0095726)
lRpPhapP	0.028338	0.0298911	0.0163437	0.0294306	0.0311612	0.0227285
SE	(0.0129233)	(0.0128807)	(0.0272798)	(0.0135768)	(0.0131504)	(0.0281561)
cons	2.341179	2.40983	2.508682	2.444295	2.279315	3.198035
SE	(0.1136816)	(0.122863)	(0.3539506)	(0.3847882)	(0.1264308)	(0.6230195)
R^2	0.03479	0.4144	0.4	0.3985	0.4081	0.4353
$adjustedR^2$	0.3644	0.3732	0.3577	0.356	0.3665	0.3689

After inspecting all of the four tables we can conclude remarks for each of the model individually.

- lLE45M: lPoiOx reacts on change in lInvEnv and lSolid
- lLE65M: here we can see that the inclusion of lInvEnv changes the the influence of lPoiOx, we cannot say the same thing about lSolid. The other remark that should be made here is that the addition of lRPhapP changes the influence of other health factor lNPpDoc
- lLE45F: same as in the previous model lPoiOx reacts on inclusion of lInvEnv. The other interesting change happens when we include economic factor of lAW, in that case we can see increase in coefficients of lU
- lLE65F: the situation is similar to the case of model of male at the age 65. Environmental pollution factors lPoiOx and lInvEnv are responding to addition of lSolid and the health factor lRPhapP changes due to the inclusion of lNPpDoc

When we look at the fact that the changes in the coeficients are small we can observe that our models work even as a part of a bigger picture. The last good news we can conclude from the pictures above is what we can observe from R^2s. We can easily see that by adding different variables they do not change so much, so our models are actually the best we could came up with.

Chapter 7

Conclusion

The aim of our study was simple. To find determining factors of life expectancy within the Czech Republic. It is a first study of this kind for our country and we believe it might be extremely helpful not only for future research in this field but at the same time it can be practically applied in the decision process about pension and health care system in the future.

Life expectancy as a measure of the quality of life has been in use for some time but studies still do not have a clear opinion what the factors determining this widely used method are.

There are clear outcome of longer or shorter life expectancy on the economy in form of pension and health care system but even though today's economy is based on precise models there are still ways how to improve the formula for calculating the life expectancy. And the outcomes could be great.

There have been many studies trying to find out the essential factors determining the life expectancy but without any general luck that would apply in every situation. This is in fact understandable outcome due to the fact of big changes in the field of healthcare during the past century.

There have been other changes as well - democratization, introduction of basic hygiene and sanitization and increase in education. Except those changes that have been all mentioned above we have come up with a whole new category. Environment in the sense of pollution has been a factor which is playing an increasing role in all of our lives and it is not for so long when it came into general consciousness that we have to pay attention to it.

Except the fact that our study had pointed out the importance of environmental polluiton factors determining the life expectancy it come across of other interesting issues.

Beside our research question we had three hypothesis which were based on previous work of Shaw et. al. (2005) and our expectations. We have found two out of three hypotheses to be correct - importance of environmental pollution factors (as said above) and percentage of people over 65 in the population. The last hypothesis stating pharmaceuticals consumption as a crucial explanatory variable was rejected in three out of four models, which is different outcome in three out of four cases from the research done by Miller Jr & Frech (2000) , who found "consumption of pharmaceuticals" as a crucial variable.

We have found out that the factors influencing the life expectancy are in the Czech Republic the same for male at the age of 45 and 65 and female at the age of 45 with the small exception for female at the age of 65. For female at the age of 65 we have discovered differences in form of lacking importance of economic factors and bigger importance of environmental pollution factors had been revealed.

Even though we should point out that the model works the best for male at the age of 45, for the other group the R^2 is smaller, which means that we could improve those models. Unfortunately it seems the best option would be to try to include life-style factors that are unavailable for the homogeneous Czech Republic.

This is one of our recommendations for the future to deepen our research by using life-style factors in form of alcohol, tobacco and fat consumption on one hand and fruit, vegetable consumption and the average length of exercise on the other. (Tamakoshi et. al. 2010) This could reveal the differences that some category may live in a healthier way than other.

We have already mentioned the importance of health care factors but in the case of the Czech Republic our work has revealed an important fact. We have found the probable connection between all of the health care factors. This fact is represented by that in each of the four models we have found exactly one health care factor to be significant. In three of the models the only significant factor was"number of people per one doctor" this would imply that doctors are the means distributing the finance putted into the system in the most effective and efficient way. The fourth model - female at the age of 65 - considers the only significant health care variable "revenue from pharmaceuticals". This fact may indicate that for female at the age of 65 the most important thing is to keep paying attention to the already discovered problems so the assistance of the doctors is not the essential part.

To make sure that our explanation is the right one we would recommend con-

ducting a research that would aim for distribution of diseases among the population within the Czech Republic.

One more propositions in the field of health care system would be to find a way to cooperate with insurance companies to gain access to the evidence of money invested into the different health care departments to see which investments are the most effective and efficient. Unfortunately we were not able to cooperate with the insurance companies on such level so we have to acknowledge that this may once again increase R^2 of all of our models. On the other hand we should point out that we have tried all of the factors on the district level accessible from the Institute of Health Information and Statistics of the Czech Republic and we still found significant only the variables mentioned above.

Our model is the first one to focus solely on the Czech Republic. Most of the models before were international models that were actually polling data that have already been pooled for one country. Moreover the samples of the countries usually did not exceed 30 which is a small sample. Our sample is more than two times that and almost three times as big as the sample of the study of Shaw et. al. (2005) we have used as a pre-image for our model. The impact of our study can be great not only for the Czech Republic.

Thanks to the model for the life expectancy constructed especially for the Czech Republic we will be able to predict impacts of different changes in the economy, health care system or environmental pollution to the extent to adjust the whole pension system or payments within the health care system for balanced future economic situation.

Bibliography

Å berg, R. (2014): "Unemployment Employment Chances and the of the Less Educated." **19(2)**: pp. 199–216.

Ashraf, Q. H., A. Lester, & D. N. Weil (2013): "When Does Improving Health Raise GDP ?" **23(1)**: pp. 157–204.

Asiskovitch, S. (2010): "Gender and health outcomes: the impact of health-care systems and their financing on life expectancies of women and men." *Social science & medicine (1982)* **70(6)**: pp. 886–95.

Auster, Richard; Leveson, I. & D. Sarachek (????): *The Journal of Human Resources* .

Berndt, E. R. (2002): "Pharmaceuticals in U.S. Health Care: Determinants of Quantity and Price." *Journal of Economic Perspectives* **16(4)**: pp. 45–66.

Blackman, T. (2011): "Public spending and life expectancy." *Maturitas* **70(2)**: pp. 95–6.

Bulled, N. L. & R. Sosis (2010): "Examining the Relationship between Life Expectancy, Reproduction, and Educational Attainment." *Human Nature* **21(3)**: pp. 269–289.

Chan, M. F. & M. Kamala Devi (2012): "Factors Affecting Life Expectancy: Evidence From 1980-2009 Data in Singapore, Malaysia, and Thailand." *Asia-Pacific journal of public health / Asia-Pacific Academic Consortium for Public Health* .

Crimmins, E. M. (2004): "Trends in the health of the elderly." pp. 79–98.

CzechStatisticalOffice (2013): " Life expectancy in districts (LAU 1) in 2008-2012 (data extra regular outputs)." *Technical report*, Czech Statistical Office.

CzechStatisticalOffice (2014): "Statistics for the czech republic."

Echevarr, C. A. (2003): "Life expectancy , retirement and endogenous growth." **21**: pp. 147–174.

Eggleston, K. N. & V. R. Fuchs (2013): "American Economic Association The New Demographic Transition : Most Gains in Life Expectancy Now Realized Late in Life." **26(3)**: pp. 137–156.

Egidi, V. & D. Spizzichino (2008): "ITALIAN OF PUBLIC HEALTH Life expectancy : complex measures of the length and the health related quality of life OF PUBLIC." **5(2)**: pp. 91–95.

Griffin, B., V. Loh, & B. Hesketh (2013): "A mental model of factors associated with subjective life expectancy." *Social science & medicine (1982)* **82**: pp. 79–86.

Grootendorst, P., E. Piérard, & M. Shim (2009): "Life-expectancy gains from pharmaceutical drugs: a critical appraisal of the literature." *Expert review of pharmacoeconomics & outcomes research* **9(4)**: pp. 353–364.

InstituteOfHealthInformationAndStatisticsOfTheCzechRepublic (2014): "Statistics for the czech republic."

Leung, M. C. M. & Y. Wang (2010): "Endogenous Health Care, Life Expectancy and Economic Growth." *Pacific Economic Review* **15(1)**: pp. 11–31.

Lin, R.-T., Y.-M. Chen, L.-C. Chien, & C.-C. Chan (2012): "Political and social determinants of life expectancy in less developed countries: a longitudinal study." *BMC public health* **12**: p. 85.

Mackenbach, J. P. (2013): "Convergence and divergence of life expectancy in Europe: a centennial view." *European journal of epidemiology* **28(3)**: pp. 229–40.

Mackenbach, J. P., Y. Hu, & C. W. N. Looman (2013): "Democratization and life expectancy in Europe, 1960-2008." *Social science & medicine (1982)* **93**: pp. 166–75.

Mariani, F., A. Pérez-Barahona, & N. Raffin (2010): "Life expectancy and the environment." *Journal of Economic Dynamics and Control* **34(4)**: pp. 798–815.

Miller Jr, R. D. & H. E. Frech (2000): "Is there a link between pharmaceutical consumption and improved health in OECD countries?" *PharmacoEconomics* **18 Suppl 1**: pp. 33–45.

Mirowsky, J. (1999): "Subjective life expectancy in the US: correspondence to actuarial estimates by age, sex and race." *Social science & medicine (1982)* **49(7)**: pp. 967–79.

Mirowsky, J. & C. E. Ross (2013): "Life Expectancy * Socioeconomic Status and Subjective." **63(2)**: pp. 133–151.

Okamoto, K. (2006): "Life expectancy at the age of 65 years and environmental factors: an ecological study in Japan." *Archives of gerontology and geriatrics* **43(1)**: pp. 85–91.

OxfordUniversityPress (2014): "Definition of life expectancy in English."

Pautrel, X. (2009): "Pollution and life expectancy: How environmental policy can promote growth." *Ecological Economics* **68(4)**: pp. 1040–1051.

Pope, C. A., M. Ezzati, & D. W. Dockery (2009): "Fine-particulate air pollution and life expectancy in the United States." *The New England journal of medicine* **360(4)**: pp. 376–86.

Retzlaff-Roberts, D., C. F. Chang, & R. M. Rubin (2004): "Technical efficiency in the use of health care resources: a comparison of OECD countries." *Health Policy* **69(1)**: pp. 55–72.

Seifarth, J. E., C. L. McGowan, & K. J. Milne (2012): "Sex and life expectancy." *Gender medicine* **9(6)**: pp. 390–401.

Shaw, J. W., W. C. Horrace, & R. J. Vogel (2005): "The Determinants of Life Expectancy : An Analysis of the OECD Health Data." **71(December 2003)**: pp. 768–783.

Sögner, L. (2001): "Okun's Law Does the Austrian unemployment-GDP relationship exhibit structural breaks?" *Empirical Economics* **26(3)**: pp. 553–564.

Sung, H.-E. (2004): "Democracy and political corruption: A cross-national comparison." *Crime, Law and Social Change* **41(2)**: pp. 179–193.

Tamakoshi, A., M. Kawado, K. Ozasa, K. Tamakoshi, Y. Lin, K. Yagyu, S. Kikuchi, & S. Hashimoto (2010): "Impact of Smoking and Other Lifestyle Factors on Life Expectancy among Japanese: Findings from the Japan Collaborative Cohort (JACC) Study." *Journal of Epidemiology* **20(5)**: pp. 370–376.

Tsang, M. U. N. C. & H. M. Levin (1990): "The Economics of Overeducation." .

Walker, L. J. (2013): "Life Expectancy and Health Care Planning." **(February)**: pp. 2013–2016.

Yakita, A. (2005): "Life expectancy, money, and growth." *Journal of Population Economics* **19(3)**: pp. 579–592.

Appendix A

Appendix

A.0.1 Leverage points

Figure A.1: Leverage points M45

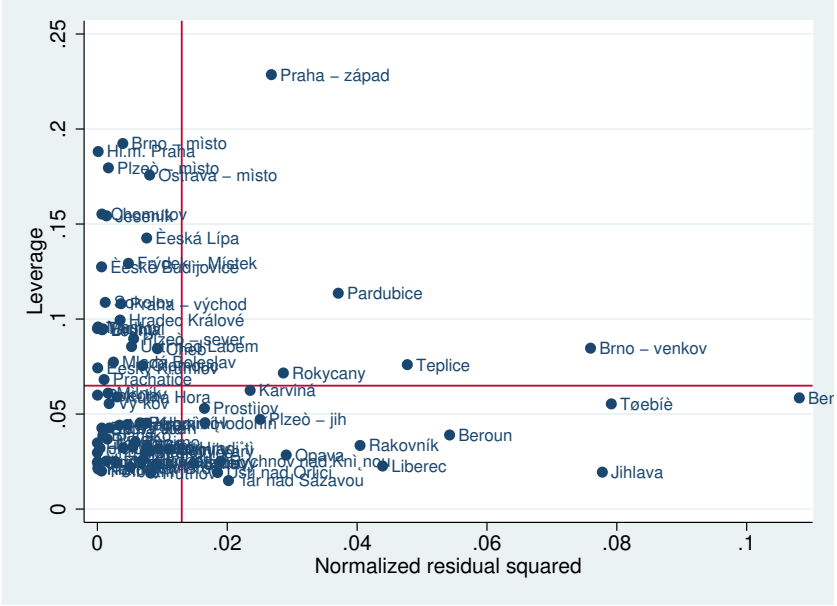


Figure A.2: Leverage points M65

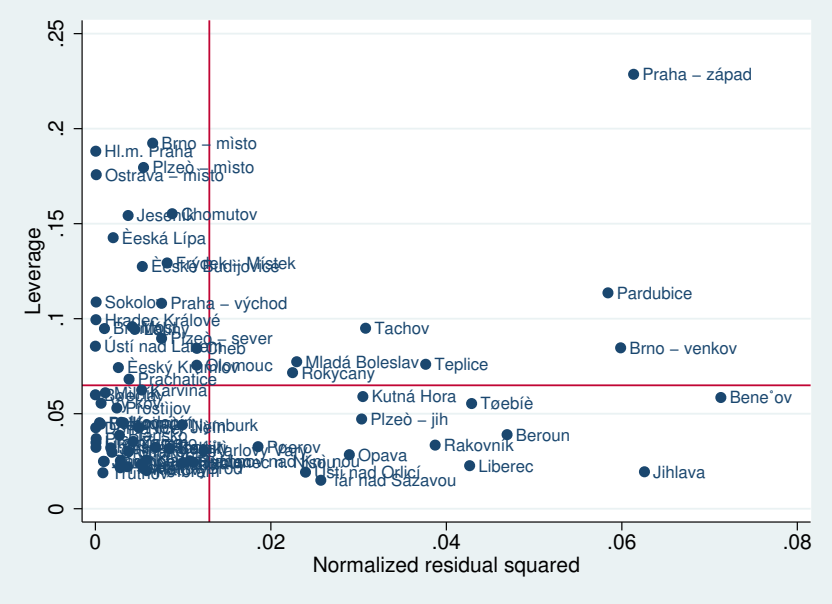


Figure A.3: Leverage points F45

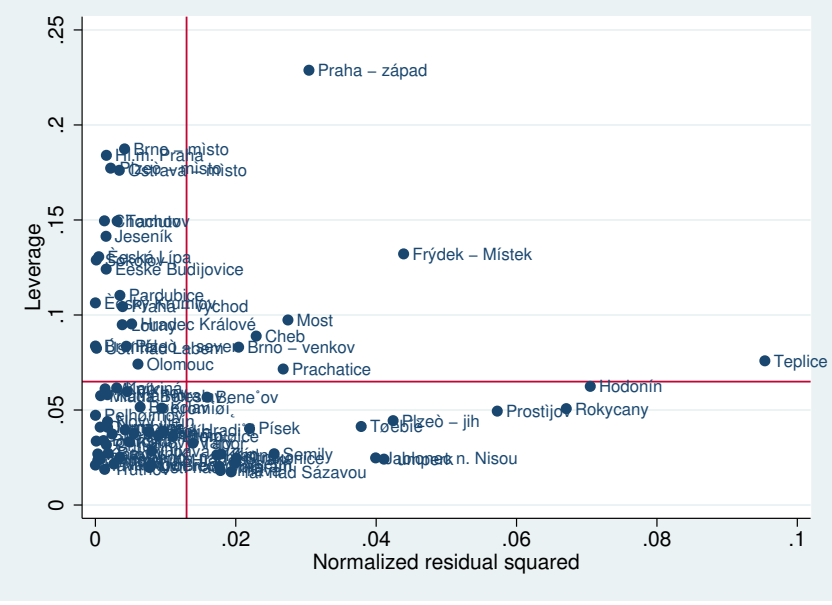
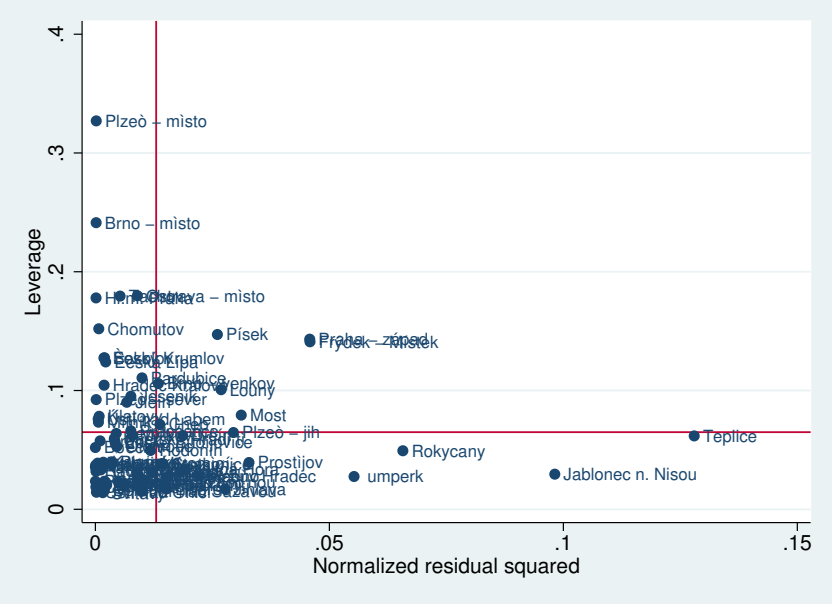


Figure A.4: Leverage points F65



A.0.2 Homoscedasticity

Figure A.5: Hommoscedasticity M45

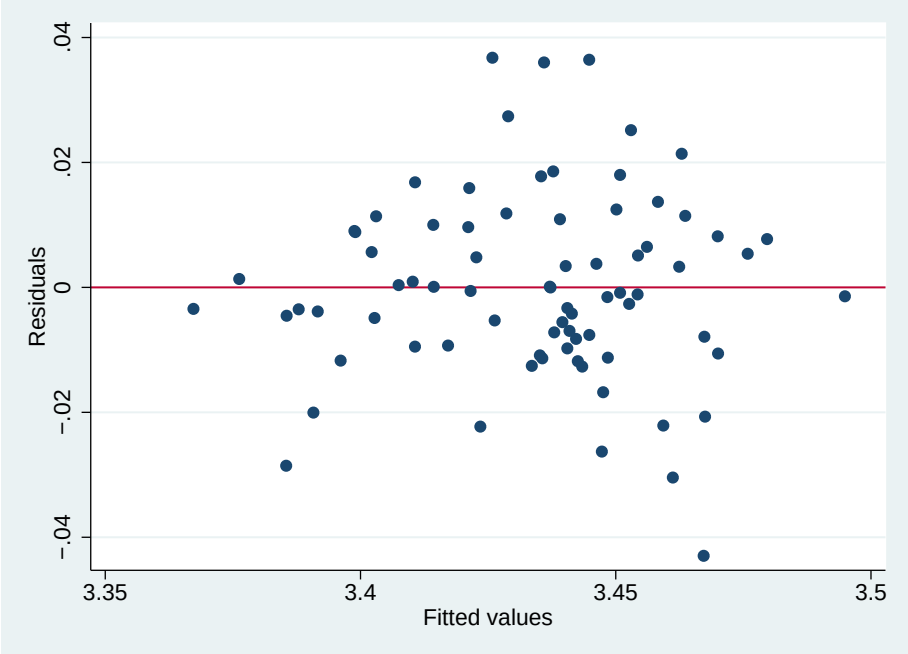


Figure A.6: Hommoscedasticity M65

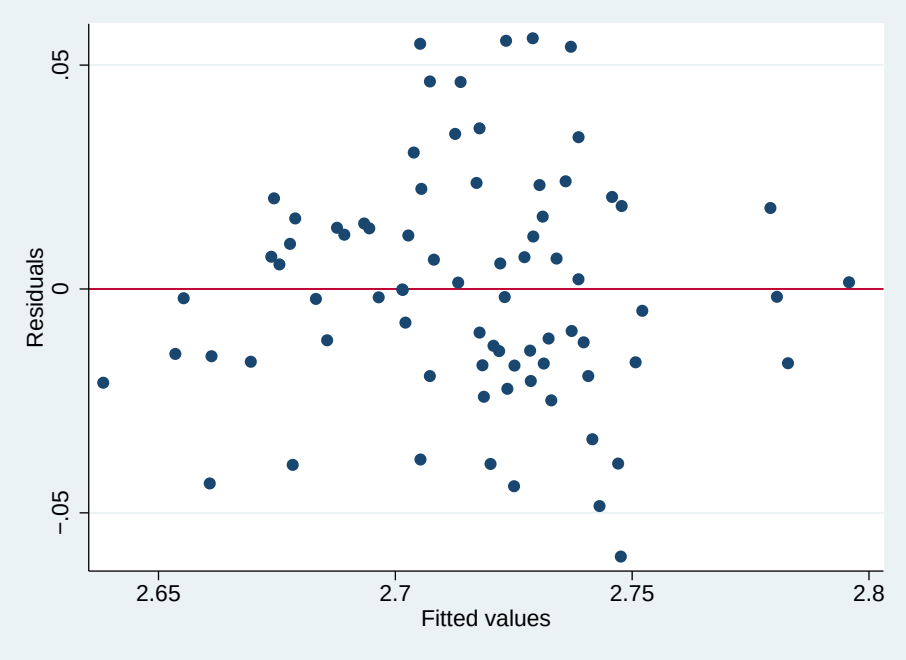


Figure A.7: Hommoscedasticity F45

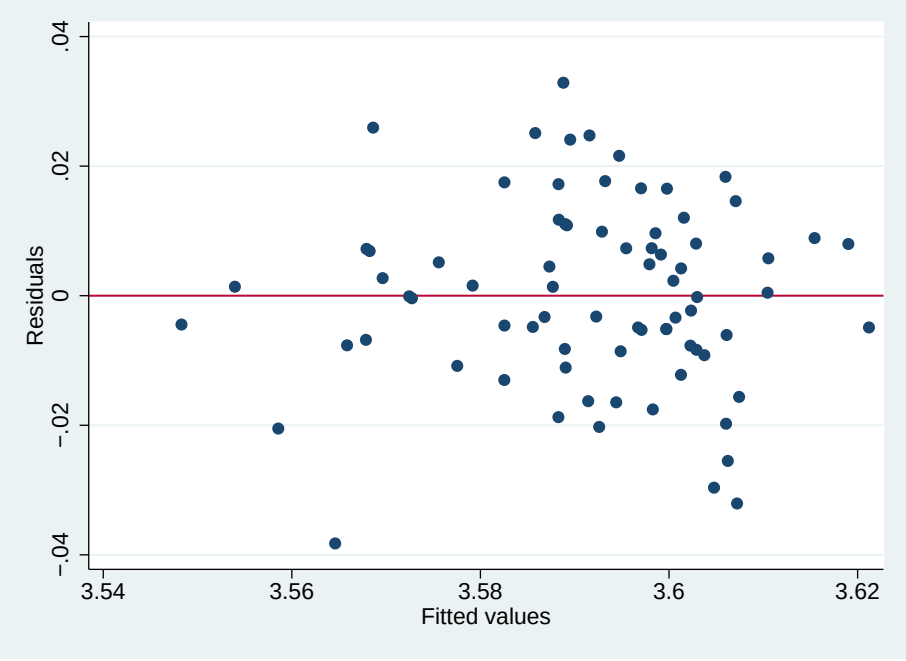
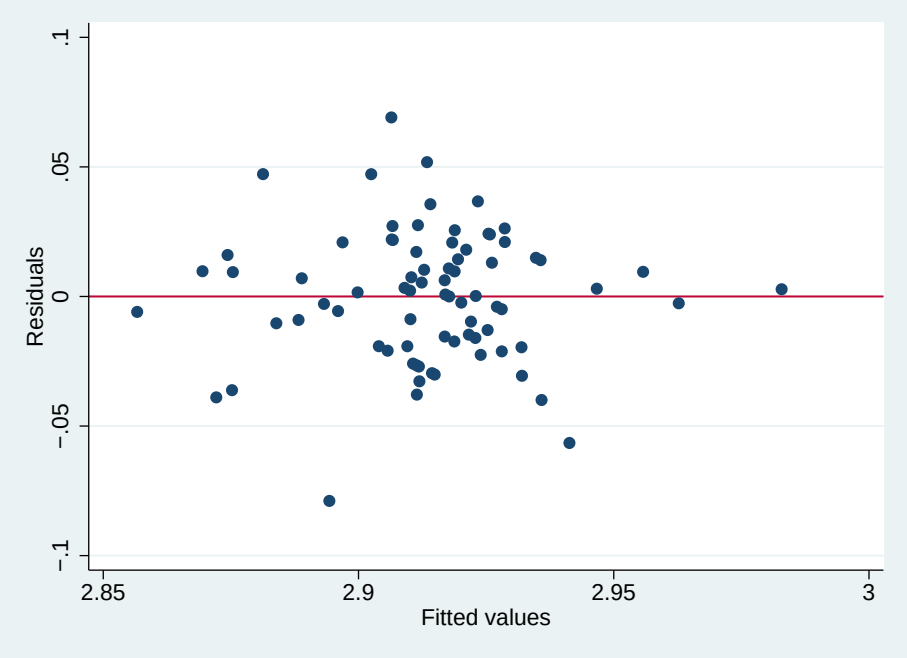


Figure A.8: Hommoscedasticity F65



A.0.3 Normality

Figure A.9: Normality M45

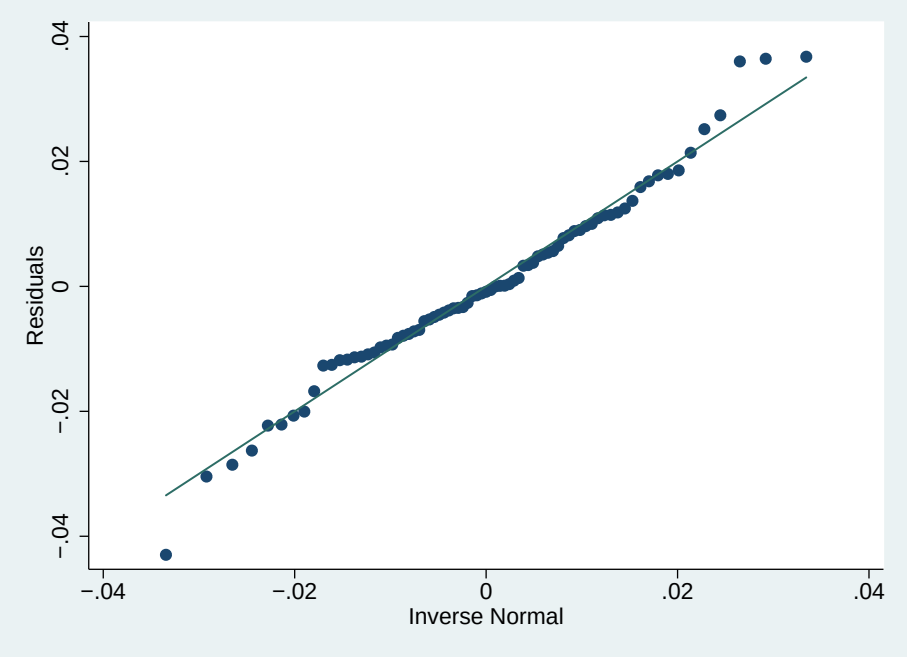


Figure A.10: Normality M65

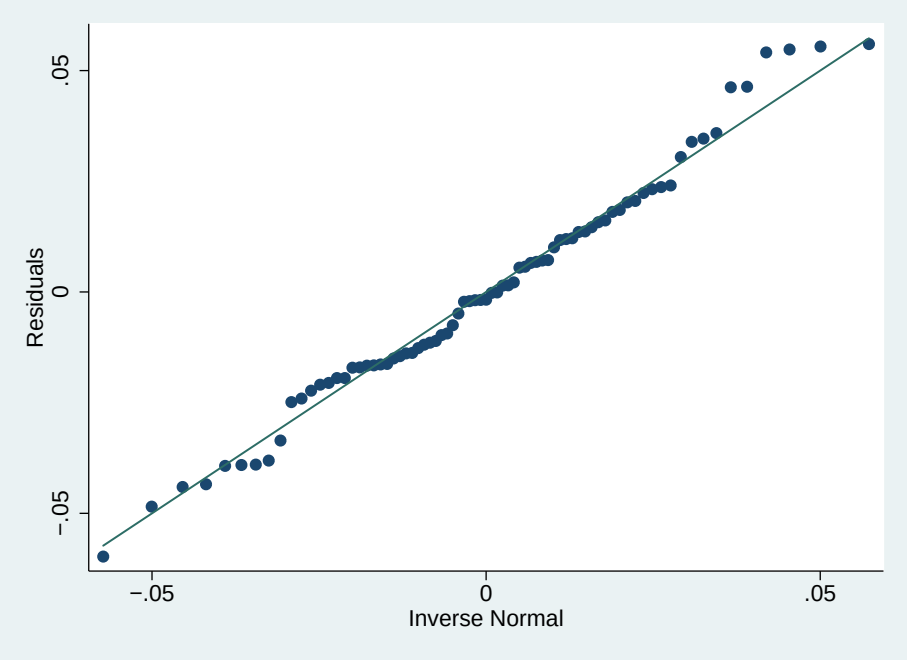


Figure A.11: Normality F45

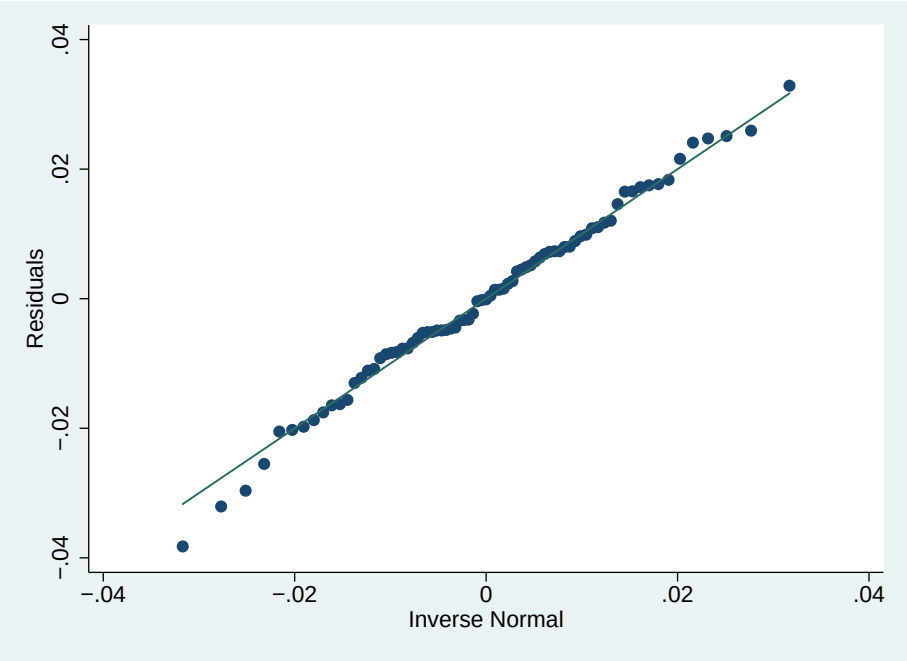
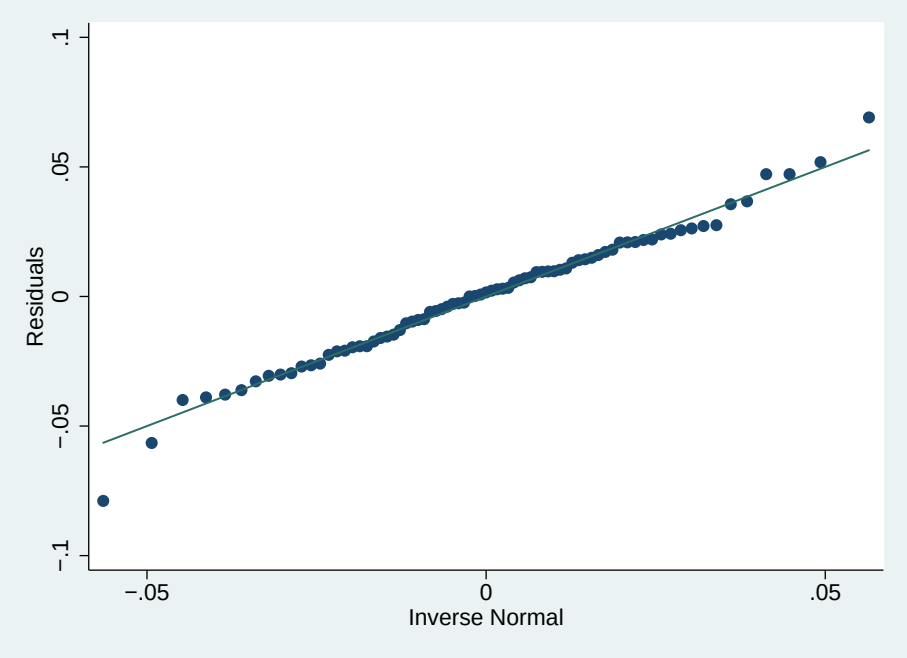


Figure A.12: Normality F65



A.0.4 Complete list of variables

- Number of people per hospitals (NPpH)
- Number of people per outpatient facilities (NPpOF)
- Number of people per one doctors (NPpDoc)
- Number of people per one pharmacy (NPpPh)
- Number of people per prescriptions (NPpPre)
- Revenue for pharmaceuticals per one person (RpPhapP)
- Number of people per one pharmacists (NPpPhar)
- Number of hospital's bed per 10000 people (NPp1HP)
- Average monthly salary before taxes (AW)
- Unemployment rate (U)
- Percentage of deaths in the population (D)
- Percentage of people over 65 in the population (M65+ or F65+ depends on the sex)

- The amount of solid particels issued into the environment (t/km^2) (solid)
- The amount of poisonous oxides (SO_2, NO_2, CO) issued into the environment (t/km^2) (PoiOx)

A.0.5 Detail list of used variables

- Expected length of life at 45/65 - "The life expectancies at the age 45 and 65 years express the average number of years to be lived by the table person at given age. They are the results of completed life tables prepared for each administrative districts of municipalities with extended powers separately for men and women using the indirect method of calculating the probability of dying, one-year age interval, and for the reason to eliminate random fluctuations they are calculated for five-year calendar period. Detailed methodology of computation of all indicators of life tables is the part of the regular annual publication Life tables for the Czech Republic, Areas and Regions (http://www.czso.cz/csu/2013edicniplan.nsf/engp/4002-13; as "alpha" fixed value (0.86) was used, corresponding to the long-term average of this indicator for the whole Czech Republic). The numbers of deaths, live births and inhabitants (as of midyear) by sex and age in each calendar year of given period were input data. Population figures used since 2011 are based on the final results of the Population and Housing Census 2011."
- Average salary before taxes - This variable is expressing average salary of employees during the year 2003 before taxes, excluding enterprises with 20 employees and fewer. (Enterprise method)
- Unemployment rate - Unemployment rate is expressing the percentage of people in a work force that on average did not work during the year 2003.
- Index of aging - This variable is describing the average percentage of people over 65 in the population during the year 2003.
- Number of people per one doctor - By this variable we can measure the number of people that one doctor has to take care of on average during the year 2003.
- Poisonous oxides - By this variable we are expressing the amount (in tons

per square kilometer) of poisonous oxides (SO2, NO2, CO) that had been drained into the atmosphere during the year 2003.

- Solid particles - This is as well environmental variable revealing the amount (in tons per square kilometer) of solid particles that had been drained into the atmosphere during the year 2003.
- Investment into the environment - Amount of money invested into the environment in CZK during the year 2003.
- Pharmaceutical revenue - This variable shows the amount of money paid on average by one person for pharmaceuticals during the year 2003.

A.0.6 Email from the CSO

"Hello, I believe that the original answer was quite sufficient. To complement I can only give these brief arguments: Data on the consumption of alcohol and tobacco, comes either from the selection survey of statistics of family accounts (SFA) or from the statistic on consumption of supplies. In the case of SFA required items of consumption collected from households are undervalued and a sample size (about 3000 households) is too small, so that the detailed breakdown would provide sufficiently reliable data for modeling. In the second case, the published data processed on the basis of data on production, stocks, imports and exports, which are aggregated statistics for the whole Czech Republic, which is not possible to breakdown into regional statistics."