University of Economics, Prague

Faculty of Economics

Study programme: Economics and Economic Administration

ANALYSIS OF THE DEPENDENCE OF ECONOMIC GROWTH ON THE LEVEL OF HIGHER EDUCATION OF THE POPULATION WITHIN THE EU COHESION REGIONS

Bachelor thesis

Author: Michal Votava

Thesis supervisor: Ing. Martin Lukavec, Ph.D.

Year: 2019

Prohlášení

Prohlašuji na svou čest, že jsem bakalářskou práci vypracoval samostatně a s použitím uvedené literatury.

Michal Votava

V Praze, dne 25. 8. 2019

Acknowledgements

I would like to express my gratitude to my supervisor Ing. Martin Lukavec, Ph.D. Without his ideas and assistance, this bachelor thesis would never have been written. His advice and willingness were invaluable and highly appreciated.

Last but not least, I would like to thank my beloved Bára for her incredible support during completing of this Bachelor Thesis. She supported me mentally and made much needed corrections. Thank you, my love!

Vysoká škola ekonomická v Praze
Katedra regionálních studií

Národohospodářská fakulta
Akademický rok: 2018/2019

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Zpracovatel: **Michal Votava**

Studijní program: Ekonomie a hospodářská správa

Obor: Veřejná správa a regionální rozvoj

Název tématu: **Analýza závislosti hospodářského růstu na míře vyššího vzdělání populace v rámci regionů soudržnosti na území EU**

Zásady pro vypracování:

1. 1. Cíl práce:

   Cílem bakalářské práce je nalézt a kriticky zhodnotit míru vlivu vyššího vzdělání populace na hospodářský růst, určit která odvětví vzdělání ovlivňuje hospodářský růst nejvíce a změřit další efekty, které s sebou vyšší poměr vzdělaného obyvatelstva přináší. Hypotézou, kterou budu ve své práci testovat je, že vyšší míra vysokoškolsky vzdělané populace pozitivně ovlivňuje hospodářský růst.
2. 2. Význam, aktuálnost nebo očekávaný přínos zvoleného tématu:

   Dnes se mění struktura zaměstnanosti velmi dynamicky a stále větší část populace pracuje v oborech vyžadujících vyšší vzdělání. Stoupající poměr vysokoškolsky vzdělaného obyvatelstva generuje efekty peněžní i nepeněžní povahy. Kromě toho, že více vzdělané obyvatelstvo se více podílí na veřejném životě nebo má vyšší sociální toleranci, tak zde existuje i významný efekt na hospodářství.
   Tyto faktory můžeme sledovat v časových řadách a zjistit tak vliv politik jednotlivých států Evropské unie na poměr vysokoškolsky vzdělaného obyvatelstva a jejího vlivu na růst hospodářství. Tímto způsobem tedy i předpovědět budoucí uspořádání Evropy a vyvodit z toho případná doporučení pro politiky jednotlivých států, či regionů.
3. 3. Charakteristika teoretické části práce:

   Teoretická část se zaměří na vymezení základních pojmů týkajících se problematiky zvyšování vzdělanosti a jejího efektu na růst hospodářství. Bude shrnovat názory ostatních autorů, kteří se touto problematikou zabývají na základě aktuální literatury. Dále se pokusím shrnout aktuální situaci, a poukázat na politiky, které jsou státy využívány pro zvýšení podílu vysokoškolsky vzdělaných obyvatel.
4. 4. Charakteristika praktické části práce:

   V praktické části budu zejména pracovat se sebranými statistickými daty na úrovni regionů soudržnosti (NUTS 2). S pomocí těchto dat budu popisovat vývoj jednotlivých ukazatelů a jejich závislosti, které s sebou přináší. Dále se pomocí multivariabilní regresní analýzy časových řad pokusím předpovědět budoucí vývoj vzdělanosti a růstu hospodářství jednotlivých regionů.
5. 5. Klíčová slova:

   Vzdělání, hospodářský růst, regionální rozvoj, regiony soudržnosti

Rozsah práce: min 45 normostran

Seznam odborné literatury:

1. FLORIDA, R L. *The rise of the creative class : and how it's transforming work, leisure, community and everyday life.* New York: Basic Books, 2004. ISBN 0-465-02477-7.
2. Glaeser, E. L., & Tobio, K. (2007). The Rise of the Sunbelt (Working Paper No. 13071). National Bureau of Economic Research. https://doi.org/10.3386/w13071
3. Greenstone, M., & Moretti, E. (2003). Bidding for Industrial Plants: Does Winning a 'Million Dollar Plant' Increase Welfare? (Working Paper No. 9844). National Bureau of Economic Research. https://doi.org/10.3386/w9844
4. Moretti, E. (2004a). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. Journal of Econometrics, 121(1–2), 175–212
5. Moretti, E. (2004b). Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions. The American Economic Review, 94(3), 656–690.
6. Moretti, E. (2010). Local Multipliers. American Economic Review, 100(2), 373–377. https://doi.org/10.1257/aer.100.2.373
7. Moretti, E. (2012). The New Geography of Jobs. Houghton Mifflin Harcourt.

Datum zadání bakalářské práce: leden 2019

Termín odevzdání bakalářské práce: květen 2019

Michal Votava
Řešitel

Ing. Martin Lukavec
Vedoucí práce

doc. Ing. Petr Toth, Ph.D.
Vedoucí katedry

prof. Ing. Zdeněk Chytil, CSc.
Děkan NF VŠE

Abstrakt

Tato bakalářská práce se zabývá vyšším vzděláním v kontextu hospodářského růstu. Cílem této práce je nalézt a kriticky zhodnotit vzájemnou spojitost vyššího vzdělání a regionálního rozvoje. V teoretické části se zabývám rozebráním pojmu vzdělání a jeho vlivů na regionální ekonomii. Tento vliv je zkoumán i v kontextu úbytku pracovních míst v průmyslu. V praktické části provádím regresní analýzu za pomocí panelových dat, konkrétně Pooled OLS, Fixed effects a Random effects modelů s cílem změřit korelaci mezi vzděláním, HDP, disponibilním příjmem, průměrnou teplotou, variačním koeficientem průměrné měsíční teploty a hustotou obyvatelstva. Mimo regresní analýzu zkoumám i návratnost vzdělání pro jednotlivce na základě porovnání ceny školného a průměrných budoucích výdělků.

Klíčová slova

Lidský kapitál, vyšší vzdělání, hospodářský růst, regionální ekonomie, regionální rozvoj, regiony soudržnosti, pracovní trh

JEL klasifikace

R11, H75, I21, J24

Abstract

This bachelor's thesis analyses higher education in the context of economic growth. The aim is to identify and evaluate the connection between higher education and regional development. In the theoretical part, I analyse what education is and what it affects. Specifically, I am studying its' effects on the regional economy and on the decline of manufacturing jobs. In the practical part, I am using panel data models, specifically the Pooled OLS, the Fixed effects and the Random effects regression models to quantify the correlation between education, GDP, disposable income, average temperature, variable coefficient of the average temperature and density of population. Besides this, I am examining the return of education comparing costs of education and future average earnings.

Keywords

Human capital, higher education, economic growth, regional economics, regional development, cohesion regions, labour market

JEL classification

R11, H75, I21, J24

Table of Contents

Introduction

The main purpose of this bachelor thesis is to examine and verify the theory describing that population with higher education in the region has a positive impact on its economy. The topicality and time relevance of this bachelor thesis is undeniable because manufacturing production is not the main engine of economic growth anymore. The population is more educated, and it is capable of inventing new technologies much faster than ever before. It is even possible today to substitute the workers in manufactures with new high-tech machines with higher profitability as a result. This fact, that education is an important factor for economic growth is going to be discussed in the following pages. The data relevant for this topic found in the literature shows that the decline in manufacturing jobs is enormous over the last decades and the divergence between the regions is growing. The main reason is globalization, i.e., big companies are outsourcing their production overseas, where the labour market is cheaper. Fortunately, there are more new jobs created instead of the ones that are disappearing. These new jobs are in the tertiary sector and specifically, I will be targeting new jobs in the innovation sector, especially well-paid jobs in high-tech firms. Nowadays, the innovation sector has become the true engine of economic growth. The key to its growth is the human capital which determines also the growth of all the regional economy itself. Because human capital is constructed mainly of education my goal is to identify which effects are brought to the economy by the higher education and also what are the side effects. The purpose is to examine the situation namely in the regions of the European Union. But due to the different quality of data, different data collection across the Europe and interests of most authors on this topic, my theoretical part is mostly using and focusing on the research and literature from the United States. The United States has many advantages when talking about the quality of data used for this kind of literature research. Most of the relevant literature on this topic was published there. All of the American states are connected together in one federation much longer and much more stable in comparison with the states of European Union, therefore we can better understand the regionals disparities caused by educational level of inhabitants in regions of US than in the European Union. Nevertheless, I believe that the lack of research in the European Union makes this bachelor thesis more interesting and relevant because it is at least partially filling this gap.

The theoretical part of this thesis is about literature research and focuses on current knowledge about the impact of education on economic growth. Firstly, the importance of human capital and how it influences productivity is explained. Then, the meaning of education as the most important part of human capital is described and what it means for population wealth. The emphasis is on education and thus the position and perception of the importance of education in history since Adam Smith is described there. Also, I will introduce two main models of the educational systems that could be found across the world as an important determinant of quality and costs of education. Then, it is discussed the decline of manufacturing jobs with a closer look at Detroit as an icon of this decline. Specifically, the reasons for this happening and the consequences of job loss is discussed. The theoretical part is closed with the section devoted to the topic of local employment multiplier.

The practical part consists of two main parts devoted to regression analysis and return of education for individuals. For this purpose, it was needed to gather a relatively large amount of data mainly collected from Eurostat and OECD databases. Firstly, I am examining the correlation of education with the variables such as the income of households and GDP per capita which should represent the regional development and also with other non-economic variables like the density of population, average temperature or standard deviation of average monthly temperature. These non-economic variables are examined because I suppose them to correlate with education and economic variables representing economic growth. All the data were collected on the level of NUTS II regions for the time period from 2007 to 2017 as there is the best possible quality of data. Even this section is divided into two parts: first examined the relationship between education and disposable income and the second examines the relationship between educational and GDP. To examine those relationships, I used regression analysis of time series run on three different models: Pooled OLS, Fixed effects and Random effects. The second part of the practical part is studying the return of education comparing the costs of education and average future earnings using mainly OECD statistics and simple linear regression.

1 Theoretical part

On the following pages of the theoretical part of my bachelor's thesis, I will sum up all the literature research I have done with the aim to verify that education has a positive effect on economic development. The purpose of this part is to get at least a basic understanding of given problematic and through that to establish some basis for the ongoing practical part. After reading the theoretical part, the reader should have a brief knowledge about what is human capital, what it is consist of and how and through which channels it affects both, directly and indirectly, economic growth.

1.1 Human capital

The definition of human capital according to OECD is: "*The knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being"* (OECD, 2001). According to this definition, human capital is a measure of skills of individuals that consist of factors like education, work experience, communication skills, intelligence, personal habits, creativity etc. We can acquire most of these factors from schooling, work, training courses or just by living and gaining experience. Together, the set of these factors gives us human capital that determines the productivity and earning potential of individuals.

As time goes and the engine of all advanced economies has changed from manufacturing to innovations, the human capital is becoming still more and more important for economic growth. Factors like education, schooling, science, technology, knowledge and training have become the most valuable to achieve the development of the country (Salgür, 2013). The higher level of human capital means that workers can learn faster, they are more responsible, able to adapt to different situations, handle more difficult and complicated tasks and is much easier for them to learn how to use new technologies. Despite the fact that human capital is intangible and can't be precisely measured, it is a priceless attribute of each human being, which contributes to economic growth. In other words, human capital is required to gain a higher level of labour productivity.

However, the increase in human capital bids costs. To increase the mentioned factors, we must take into consideration the costs of both time and financial nature. Usually, we get our starting skillset through schooling and after compulsory school attendance. Thus, it solely depends on our diligence, conscientiousness and determination, how we develop the skillset furthermore. The question that rises from the government point of view, is whether it pays off to have free public universities and how much to invest through them into the education of future workers. One of the indicators that the education is the important factor of economic growth is that developed countries with high economic growth have generally higher standards of schooling strategies leading to better human capital (Hanushek, 2013).

The firms similarly must deal with similar issues to have a functional structure with valuable and productive employees. For example, the training of employees is relevant for their contribution to the firm. However, the training can be expensive, time-demanding, and also the profit of the employees decreases during the ongoing training courses, i.e., they become non-productive during training. Therefore, the companies have to calculate the return of that investment in employees the same way just as if the company wants to buy new machines for its factory. Last but not least, a point of view of the people have to be considered. Depending on goals every person wants to accomplish, it is crucial for everyone to decide whether to invest into education, training or language courses, which university to choose and most importantly, which field of study they want to pursue. This issue was studied, and it is discussed thoroughly further in this thesis in the chapter Return of Education later on.

1.1.1 Influence of human capital on productivity

It is known, that human capital influences productivity and there have been many theories described in the literature. For example, as a suitable representative of this relation between output and human capital we could use neoclassical production formula:

$$Q = F(K; L)$$

Where:

- Q stands for products.
- L stands for labour (we could also perceive as human capital).
- K stands for capital.

If we want to describe this topic more complex, we could look at the Cobb-Douglas production function, which takes the neoclassical production function above and includes also the effect of technology. The formula itself looks like this:

$$Q = AL^{\propto}K^{(1-\alpha)}$$

Where:

- Q stands for products.
- L stands for labour (we could also perceive as human capital).
- K stands for capital.
- A is a positive constant that stands for total-factor productivity.
- α is constant between 0 and 1 and stands for output elasticities of capital and labour.

Constant A includes the positive effect of technology in the production function. This effect of technology can be also perceived as the output of human capital, e.g., when the firm wants to be innovative it needs to invest in highly educated or trained people, who are capable of developing new innovative improvements of the company interests such as new strategies of the firm, products, technologies, etc.

The neoclassical model of production function has one important issue that it ignores human capital, which Paul Romer was pointing out. Therefore, he adjusted the formula and Romer's model of production function includes also the level of knowledge in the firm's stock of capital (Romer, 1986).

In the article *Human Capital and Regional Development* (Gennaioli et al., 2013) where authors investigate the determinants of regional economic growth, the model describing how and through which channels does human capital influences productivity is presented. This model combines three features:

1. The human capital of workers X human capital of entrepreneurs

The first feature describes the difference between the human capital of workers, which enters as an input to neoclassical production function and human capital of managers or entrepreneurs. The human capital of entrepreneurs influences productivity independently on the firm-level. Also, the big difference is that investment into entrepreneurial human capital often returns as profit rather than wages.

2. Externalities of human capital

The second feature describes human capital externalities. It generally claims, that people in a given location are learning from each other, while they are spontaneously interacting mutually. This means the knowledge is transmitted, broadened and multiplied across the people without being paid for. That is one of the reasons, why companies today settle near to other companies with the same industrial interests, even though the rivalry between them can make more difficult becoming successful on the given market, especially when the new company settles down next to the global market leader. These externalities enable existing of regions like Silicon Valley, but it is incredibly expensive for both firms and workers. Firms have to pay for competent, well-educated people and estates, yet it still pays off to them. That is not only for the reason mentioned above but also because of the huge labour market, which arises in these locations. Hence, it is much easier for companies to find top-talented workers. It is not only beneficial for firms but also to workers to move to a location with a higher share of educated people as there are externalities on a social level that contribute to local economic growth. For example, there is less crime and racism in these regions (Hjalmarsson and Lochner, 2012) and other factors such as habits of people in the given region or patience, way of thinking, etc (Goleman, 1995).

3. Mobility of workers and firms

The last feature in our model, that also needs to be considered as it has become more prevalent than in the past, is the mobility of firms and workers across regions. In the previous point was mentioned, that the shift of workers and firms to the location, with a bigger density of the given industry, is advantageous not only for workers but also for the companies. However, that is not applicable for all of them and therefore many firms moved overseas driven by the cheaper labour market and estate market. Even though the transport costs are higher, it is still insignificant fact compared to savings from labour and estate costs.

1.2 Education

Education is the process of learning, especially through the school system or it could be also perceived as knowledge gained from other sources such as training or experience. Education is the most important factor that impacts the quality of human capital. It improves the way people live their lives, makes the decisions, how they interact with each other and also change the overall perception of the world. In other words, education helps with creating more individual people with better behaviour and habits, which leads to an improvement in the quality of family life and the neighbourhood environment. This is caused by educational externalities. It is known that education has many externalities. There is, for example, less crime in the regions with a higher share of educated people (Lochner and Moretti, 2003), more social equity, a stronger sense of nationhood, reduced risks from infectious diseases and so on (Mingat and Tan, 2012).

To be able to work with education as a measurable factor, we need in the first place to define which type of education will be examined, as there are many types of education. However, not all of them are trackable or measurable. For economic growth, all types of education like schooling, training and other gained knowledge are very important factors. Fortunately, only data of the measurable type of education, which are as well statistically tracked, were needed for further research in the practical part of this bachelor's thesis. Specifically, the practical part is using the Eurostat database, where shares of the population for each level of school education is tracked.

According to Selami Ahmet Salgür PhD, there are three main groups into which the perception of education or human capital can be separated (Salgür, 2013).

1. Education is considered as a separate component of the production process.

Specifically, education is perceived as an input of production function. Which is something that we mentioned above in the chapter *Influence of human capital on productivity*.

2. Educations acquired through practice.

Simply, education is also perceived as some kind of skill that is acquired by mastering some action. It is expected that a worker with 10 years of practice should be more skilled than a worker with just a year of practice. We can also call it as "learning by doing".

3. Cooperation of human capital, economic conditions and technology.

Especially, the third group is very important for us. It just says that educated workers have an advantage in implementing and using technologies. And more importantly, education is needed input for developing new technologies.

1.2.1 Position of education in history

The idea of education affecting economic growth is relatively new. People started to explore the importance of human capital and specifically education just before the turn of the millennium as the human capital has become a more important factor than physical capital. There are still plenty of effects and links of education that needs to be discovered as it is such a young field. On the other hand, the importance of education in economic growth has been known much earlier.

The first economist who somehow came up with the basic ideas of influencing economic growth through human capital was no one else than Adam Smith in his well-known work *Wealth of Nations* (Smith, 2000) from 1776. It means 200 years before the human capital has become such an essential element for all the advanced economies and before someone really formalized this theory. In his book, Adam Smith describes the differences between investing in human capital and physical capital. Specifically, he describes education and training as determinants of individual productivity and therefore also the wages of the workers. And implicitly, through aggregation are education and training also determinants of the wealth of nations (Demeulemeester and Diebolt, 2011).

Besides Adam Smith, just a few years after publishing *Wealth of Nations* at the beginning of the 19th century the Prussia had been dealing with reform of the educational system because of the defeat in 1807. Prussia blamed economic backwardness for their defeat in Napoleonic wars and therefore they aimed to modernize its economy through better educational system as they stressed the importance of human capital. They hoped to be able to assimilate new technologies (Gispen, 2002). It is interesting that 200 years ago, in such an industrial era the education had been related to economic growth on the state level and within its bureaucrats, Prussia started the reform of the educational system. In the lead of this reform was Wilhelm von Humboldt who changed dramatically whole Prussia's educational system and developed new teaching methods with new study plans and new fields of study. As a result, Prussia had become the leader in primary schooling in 1816 (Peter H. Lindert, 2006).

Another one who later dealt with human capital as a factor of production on some deeper level has been Robert Solow. In 1956 he introduced the today well-known neo-classical model of economic growth, a model with an accumulation of production factors. Only one year after in 1957 Robert Solow added an empirical methodology to his model that convinced other economists that conventionally measured factors are a reason for only a small part of growth progress (Solow, 1957). The rest of the progress is interpreted as exogenous technological progress, education etc. This breakthrough plays a key role in the evolution of economic thought concerning education and other exogenous factors as determinants of growth (Demeulemeester and Diebolt, 2011).

The effects of education itself on economics growth were not considered in the economic theories until the 1980s. In 1983 Findlay and Kierzkowski and Paul Romer in 1986 introduced the effect of human capital in terms of endogenous skills in economic development (Findlay and Kierzkowski, 1983; Romer, 1986). But they still did not describe how exactly education affects economic growth and till today there is not a simple answer to that question.

1.2.2 Educational systems

To understand the outcomes of education, it is important to know the cross-national differences in higher education systems. Return of education, which will be further discussed in the practical part, depends on factors like financial resources, management, performance and also on government funding and control.

The easiest way to separate systems of education is to divide them into two groups: private and state schools. This thesis is mainly focused on higher education, therefore the differences between educational systems will be discussed primarily for colleges. The biggest difference, that arises between private and public colleges, is the way of their funding. State schools are funded by the government as is commonly known, but they are also funded by tuition fees, but usually in reduced form. This funding system can be for example found in the United States, the situation in Europe is slightly different. The European system of higher education is mostly based on government funding and that is the reason why also top-tier universities are accessible for students without getting in debt. But of course, not in every European state is the higher education system the same. It also depends on whether you are a European citizen or not. Looking at average tuition fees in Europe the cost of bachelor's degree for European citizen is about 4.500 EUR per year (and 8.600 EUR for non-European citizen) which is still much less than tuition fees of US's universities. The average tuition fees for master's degrees is slightly higher than for the bachelor's studies. It's 5.000 EUR per year for European citizen and 10.170 for outside students (Studyportals B.V., 2019).

Of course, funding and costs of studying are not only factors that affect our decisions about which school choose. Another important difference between state and private colleges is usually the size of the college, meaning the buildings, number of students or size of classes. Private universities tend to be smaller and also have smaller classes than state universities. In the smaller class, it is much easier to interact with the professor, which is the premise of a better outcome from studying and is usually preferred by students. On the other hand, private schools cannot often offer you that many opportunities as public ones because of their size. Also, the fields of study are much more limited compared to which you can study at state university (Peterson's, 2017).

However, the top 10 universities in 2018 according to the *Academic Ranking of World Universities 2018* (shanghairanking.com, 2018) are mostly private.

To summarize this chapter, the two key differences between private and public schools were mentioned. The educational system in the US and Europe was briefly compared. Which bring us to the fact, that it really depends on whether you study private or state school and also whether you study in Europe or in the US.

1.3 The decline of manufacturing jobs

The fact that manufacturing is no longer the engine of economic growth and prosperity is shown by the gradual decline in manufacturing jobs. This is because the labour market has undergone and is still undergoing transformation and companies in advanced economies today are not looking primarily for "just" machine operators, but rather for highly skilled and highly educated workers. These educated workers are required by successful firms all over the world more than ever before, as they are trying to innovate as much as possible. Automatization and also outsourcing to other low-cost counties are destroying low-skill manufacturing jobs in advanced economies. On one hand, it is creating plenty of new jobs and on the other, these new jobs still often require high specialization and education, which is not easy nor cheap to get. For a better understanding of the development of manufacturing jobs see Figure 1 made by International Monetary Fund, which is showing the progress of share of jobs in aggregate manufacturing since 1970 to 2015 (IMFBlog, 2018).

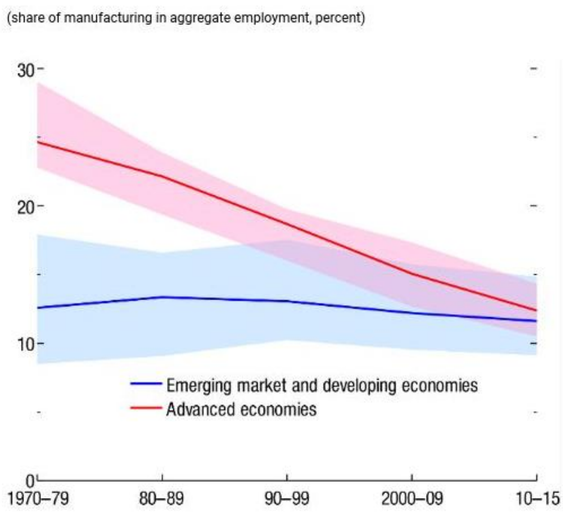

Figure 1: Overall decline of manufacturing jobs (source: IMFBlog, 2018)

As we can see, nowadays there are half of the jobs in manufacturing than it was in the 1970s in advanced economies. This decline still continues and did not stop until today. The time period from the 1970s to today is long enough to prove, that it is not just a short-term phenomenon such as a recession of manufacturing.

Quite interesting is the situation in the emerging markets where we could suggest a rise in manufacturing jobs. However, the situation there is different and the number of jobs in manufacturing is keeping its level. According to the IMF, the workers are shifting from agriculture to services, bypassing the manufacturing sector (IMFBlog, 2018). Therefore, we do not register the same waning of manufacturing jobs as in the case of advanced economies. These jobs have never been there.

To deepen understanding of the outcome seen in Figure 1, there is another figure showing the same decline but from a different point of view. Figure 2: Change of collar (source: The Economist, 2005) created by The Economist is based on the data from OECD and is showing us that the decline is happening across all advanced economies all over the world. The name of the figure comes from change of collar from blue-collar (manufacturing and other manual labour) to while collar (office employment). It derives from traditional attire workers in the United States wore, even though it doesn't apply so much anymore. In this figure, we can see the change in the number of employees in manufacturing as a share of total employment in a few selected advanced economies across the world. Countries like the United States or Great Britain, where the drop has been the sharpest, are now on less than half of the share of manufacturing employees they had in 1970. Another country with the loss of almost half of the employees in manufacturing is Germany. The Economist's outcome is more than less the same as the outcome of the IMF in Figure 1.

Figure 2: Change of collar (source: The Economist, 2005)

1.3.1 Consequences of the decline in manufacturing jobs

To understand deeper the problematic of job losses in manufacturing, let's take an example of the United States as one of the advanced economies where the decline was really strong, and some cities are dealing with the consequences till these days.

In 1978 there were 225,4 million inhabitants in the US, nowadays it is 328,6 million inhabitants (Woldometers, 2019). The population is now much larger than it was in 1978. But, even over the fact that there is that much larger labour market, there are just half as many jobs in manufacturing as in its peak in the late 70s. According to Enrico Moretti since 1985 the United States has lost an average of 372,000 manufacturing jobs every year (Moretti, 2012).

As mentioned above, the effect of a decline in manufacturing jobs is more than evident. As an example, we can look at American cities like Detroit, Cleveland, New Orleans and many others. These cities, the former big manufacturing centres of America are now struggling with a shrinking population. The most appropriate example is Detroit because it is a former icon of the US automotive industry and it is experiencing one of the worst population declines.

Detroit is a former automotive centre of the world and is also known as *"Motor City"* or *"Motown"*. In Figure 3: Population of Detroit from 1820 to 2015 (source: worldpopulationreview.com, 2019), we can see that the population of Detroit is now on the same level as one hundred years ago and it is declining for more than 60 years in a row. The shape of the curve in Figure 3 looks just astonishing, it shows us how an enormous population boom Detroit experienced and how quickly is its population disappearing. We can see that there was there is no calm and gradual progress after the 1900s at all. In its peak in 1950, Detroit had around 1,850,000 residents and just two years ago in 2017, it had only 673,000 people living there. It has been an enormous drop down in population and it still continues to decrease. According to most actual data, Detroit dropped in population between the 2000 and 2010 censuses by around 25 per cent. Other cities similar to Detroit are Cleveland that lost 17 per cent of its inhabitants or Cincinnati with loss of 10 per cent. Just behind them are Pittsburgh, Toledo and St. Louis which all lost 8 per cent of its population (Moretti, 2012).

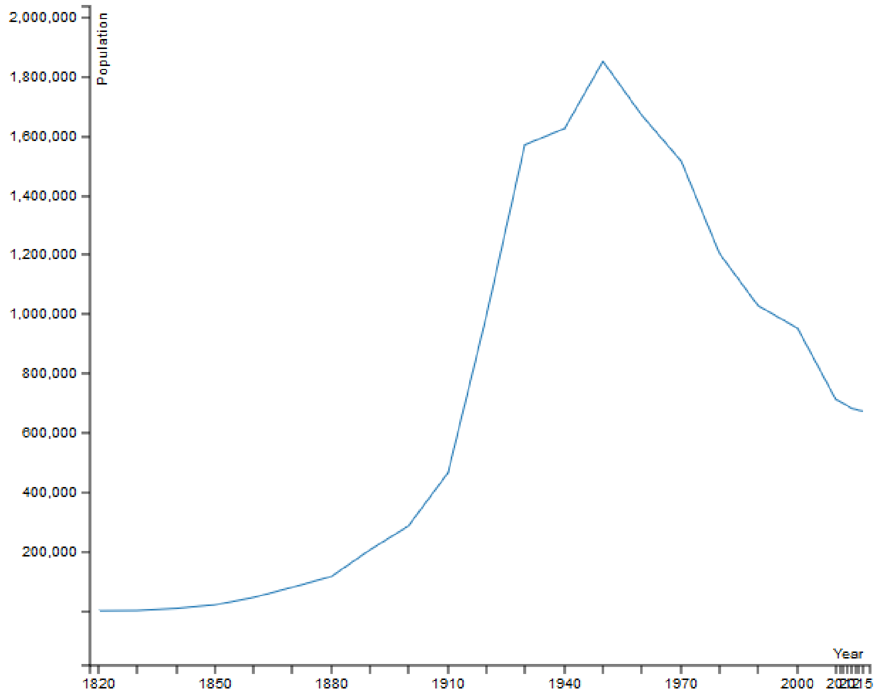

Figure 3: Population of Detroit from 1820 to 2015

(source: worldpopulationreview.com, 2019)

The reason for this huge drop is because most of the people have shifted from manufacturing to services. Nowadays, 80 % of all workforce in the United States works in services according to the journal The Economist. Although this number looks impressive, it won't look so when we realize that in the service sector have already been plenty of workers before. More than that, the United States and Great Britain already have had more workforce in services that in manufacturing since the early 1900s. On the other side stands manufacturing, where at its peak never worked more than one-third of America's workforce (The Economist, 2005).

Together with the decline of manufacturing jobs, also manufacturing output had fallen, measured as a proportion of GDP. In 1970, the manufacturing output was 26 % of GDP, in 2005 it has dropped down to just half, specifically to 13 %. The reason is growing productivity in manufacturing, because of what the prices of goods dropped and implicitly lowered the manufacturing output. Speaking of productivity in manufacturing, the interesting contrast is the United States with China in 2005 when there was a big difference in their productivity level. China had that time six-times more employees in manufacturing than the US, but the United States still had a bigger output. Which is remarkable in view of the fact that the US had much bigger productivity manufacturing even with declining manufacturing jobs, proving that bigger labour does not need to mean bigger GDP (The Economist, 2005).

Deindustrialization is happening also in the EU (Eurostat, 2017). But we know that some counties like Germany or Italy are protecting their manufacturing industry with the help of laws. But they do not realize, that this decline in manufacturing jobs does not have to necessarily mean a bad thing. The unemployment rates in most developed economies have not increased during the past decades which means that most of the laid-off factory workers have found new jobs as they are not needed anymore in the sector of manufacturing. This is not the first time in history that this is happening, pointing out the Industrial Revolution when people start leaving agriculture because of growing productivity gains from technologies like a tractor with plough, belt production and others.

As mentioned before, Detroit is well-known as a world-class icon of manufacturing in America. Therefore, I believe it is really important to understand what happened with Detroit and what it links to education. Despite its decline it is still 23rd most populous city in the US, which made Detroit one of the most important cities for the economy of the US (Bureau, 2019). Just to become aware of how much Detroit fell in ranks over the years there is a simple assumption: If Detroit would have today the same number of residents as in its top it wound placed as the 5th biggest city in the US.

If we look 70 years back, Detroit was, as the Edward Glaeser described in a podcast for EconTalk in 2013, among the most productive places on the planet. There were many entrepreneurs that were trying to construct solid automobiles in the most cost-effective ways. Detroit was the brain-hub where many small firms and entrepreneurs were collaborating on a local level to bring innovations to their sector. That was the factor that has run the local economy and well-educated labour were a crucial factor.

The gamechanger has become vertically integrated factories. Till that time all those entrepreneurs were constructing cars on the top-technological level with the aim of the quality of the final product. The difference is that factories are great for short-run productivity and do not need that much-educated workforce. But they are not that great for the urban area generally as they squeezed out all other sources and tried to be extremely productive for all costs. They moved the factories to the suburban area and then to another more low-costs area with aiming at cost-saving. The factories did not need the city and it's educated labour market anymore. Small companies, on the other hand, had been trying to be innovative as they didn't have that advantage in the form of range revenues. They also needed that interaction with other entrepreneurs to develop and that is generally how the regional economy is growing ("Glaeser on Cities", 2013), without that it could hardly work. We can compare Detroit's situation to that of today's Silicon Valley. Silicon Valley would turn out the same as Detroit if entrepreneurs in there stopped producing innovations. Nowadays, there are a lot of high-end technologies that others are still trying to adopt, but it takes some time. And as that happens another innovation will come along. This is kind of race that keeps these "brain hubs" on the top. Detroit was used to be the same, but when the big factories came everything changed. They aimed to low-costs and process effectivity, but the innovations stopped. And when new locations with lower costs appeared the end of Detroit started.

1.3.2 Reasons and consequences of job losses in manufacturing

There are many views on what is causing job losses in manufacturing. To generalize these opinions, the two most prevailing reasons why the loss is happening will be discussed and they are applicable for most advanced economies.

The first reason is the decline in demand for consumer physical goods compared to the demand for intangible goods. People are also becoming wealthier as their income is growing due to economies become richer and more advanced. So, they can afford to spend more money on all kinds of goods and services and implicitly, their consumption increases. But they mostly consume local services rather than goods. They maybe buy a new car or a fridge, but when they already have all physical goods, they turn their attention and money to the services like health, education, holidays etc. (Moretti, 2012). Of course, this does not explain such a mass job loss together with a decreasing of manufacturing gross output in the US, but the behavioural point of view also plays an important role.

The second prevailing and more important view is the growing productivity of factory workers. Increasement of productivity has been caused mainly by automatization and generally by better technology. Productivity grows faster and faster due to better technologies and on the top of that, it is much easier to automate manufacturing than services. So, when we go back to Figure 2: Change of collar (source: The Economist, 2005), we can see that such a restructuralization as Great Britain or United states went through, could be just result of growing economic strength. Yes, it caused that many people must shift from factories to services, but they moved in more productive areas and the unemployment rates have not significantly changed.

Although the manufacturing job loss is still happening and the output of the manufacturing sector measured as a share of national GDP of the US is declining, the real manufacturing output of the US is keeping up with the rest of the economy (Houseman, 2018). It means that growing productivity is a substitute for human capital and allows the sector to even grow while the workforce is shrinking. The fact that manufacturing in advanced economies is able to grow while losing jobs shows us that the manufacturing sector itself is still strong, just don't need that big labour market.

Today, people shift from manufacture to service sector, but the same has been happening to agriculture. Through automatization, new technologies and productivity growth, these sectors are becoming so productive, that do not need as big labour force.

But it also has its dark side. Concentrating on US employment market, among the main consequences of manufacturing job losses is that dislocated workers losses on average about 25-30 % of their immediate earnings and even 20% of long-term (15 – 20 year) earnings compared to workers who did not lose their job. The decline of long-term earnings is less but still significant. Also, these job losses may cause a shock for local economy depending on local labour market condition and how easily can be workers reemployed. In a weak labour market region workers suffer from much larger long-term earning losses. And not only them, but large shock can also devastate the local economy as in the example of Detroit which still not recovered from huge job losses (Houseman, 2019).

1.4 Local Employment Multiplier and Spillovers

Above, the negative effected of job loss was mentioned, but only to the workers themselves. However, it’s not only workers, but Moretti also estimates, that each manufacturing job creates approximately 1.6 nonmanufacturing jobs such as doctors, lawyers, waiters etc. This multiplier effect is higher for a more skilled and educated worker and could be up to 2.5 additional nonmanufacturing jobs (Moretti, 2010). When losing jobs in manufacturing, these workplaces are also vanishing. This described effect is called local employment multiplier and the effect is stronger when the education of worker is higher.

For better understanding, it is needed to describe further how the local multipliers work. Enrico Moretti is describing this effect in his paper *Local Multipliers*, where he claims that every new work-position that is created by the local economy is increasing the wages in the local economy and demand for local goods and services. And this positive effect is able to trigger creating new jobs in the non-tradable sector and only the non-tradable sector, the effect on the other part of the tradable sector has no significant effect. As I mention above the education is a very important factor for the strength of local multiplier’s effect. Also, the effect is different depending on the industry and Moretti is pointing out the high-tech industries to have the largest multiplier. However, we can generalize and point to most sectors that are human capital intensive. Once a new job appears, new money could be spent in the local economy, which means that a better-paid job could benefit its economy better.

The local employment multiplier is closely related to the spillover effect. This not only means that a new job is able to create more jobs, but it also attracts jobs of this particular type. It mainly works speaking about the tradable sector of the economy when the market is beyond the local region. For example, the programmer can work from almost everywhere no matter where is his client and he doesn’t care about the borders of the region. But, on the other hand, there are jobs like a hairdresser who has just the local market and can not provide services to someone in a different region. Larger market is able to generate more revenues and benefits for its economy (Moretti, 2004). That is the reason why do we have sectors-specific cities. In Detroit, it is the car industry and in Silicon Valley, it is the high-tech sector.

2 Practical part

My research is aiming to find the most relevant methods that are proving that education has a positive effect on economic growth. Because there is a lack of studies in the European environment, I chose for this purpose to collect and examine the data on the regional level of NUTS II regions in the European Union. The reason for choosing NUTS II regions is that I did not find any better territorial divisions in which the data are collected in the same way and in the same quality across all the regions. Therefore, it is the best available method, how to get the most consistent results from known available data.

2.1 The Data

My goal was to create a dataset that could be used to examine selected variables using regression analysis. The first method I used, was the regression analysis: a statistical method that seems to be the most applicable method for this study. It allows me to examine the relationship between many variables, such as education, disposable income, GDP per capita, the density of population, average temperature and standard deviation of temperature. It was needed to pick variables that presumably correlate with education and represent economic growth. I chose to examine the above-mentioned variables: GDP per capita and income per capita as they should represent a wealth of region and its inhabitants. The density of population was chosen because of assumed positive correlation with all mentioned variables: education, GDP and income. Last variables average temperature and temperature variation coefficient was chosen to be examined as controlling variables.

After the suitable variables were picked, I had to find an appropriate source of data. To cover most of the regions with useable data, Eurostat seemed to fit the best and became the main source of the dataset for my research. because Eurostat has the best quality of accessible data in the form matching my purpose and covering all the regions with yearly based data. I was able to collect most of the examined variables except the temperature-related data in Eurostat. There was a maximum effort to cover the longest time period as possible, but it was needed to cut many years from the created dataset, because of too many missing data in some regions, which created big gaps in the resulting dataset. Also, instead of the number of observations, the number of regions surveyed is preferred for my study to get more qualitative data for relevant results because I still could not get significantly more years of data when cutting regions instead of years of observation. In the end, all the data in the dataset are collected in the time period from 2007 to 2017 for 275 relevant regions of NUTS II level. This time period is providing the best possible compromise between the quality of data and number of observations as there are too many gaps to cover a longer time period and still have the relevant outcome.

Only after considering all those factors named above, I was able to construct the dataset promising relevant results. But, to understand the results of a practical part, there is a need to describe furthermore how the data are collected and why they were chosen. The description of each chosen variable is (gradually) discussed in the following chapters.

2.1.1 Educational attainment

There are quite difficult questions on how to measure education and in which form to expose it to testing. The easiest way to find an answer to them has seemed the usage of the share of people with higher education in the population gathered by Eurostat. Specifically, this indicator measures tertiary educational attainment from the age group from 25 to 64 years old within NUTS II regions.

Educational attainment is the crucial indicator for this research and because I am trying to quantify the relationship between education and all other indicators. This indicator was used in all analyses that made. The indicator is expressed as the percentage share of the population aged from 25 to 64 who have higher education, e.g. university or similar educational establishments with an appropriate level of education. This level of higher education also refers to ISCED (International Standard Classification of Education) where we examine level 5-8. These levels cover short-cycle tertiary education, bachelor's level of education, master's level of education and doctoral or equivalents to mentioned levels. Eurostat is using EU Labour Force Survey as a base for these data. This indicator is collected on an annual basis within the NUTS II regions (Eurostat, n.d.).

2.1.2 Gross domestic product per capita

GDP (Gross domestic product) is the final result of the production activity of the given region and should be telling us how fast the economy of the region is growing. Hence, I am using this regional account as it seems to fit the best to examine whether education affects the regional economy. It does not necessarily describe the wealth of the region, but it is a good indicator of economic year-to-year growth. Especially, when we can cover a longer time period.

For the purpose of my research, I am using GDP per capita at current market prices by NUTS II regions collected by Eurostat. I also tried to use GDP per capita converted to purchasing power standards (PPS), this method is similar and based on national purchasing power parities. Unfortunately, the results of its regression analysis were not that clear as the results, where current market prices (Eurostat, n.d.) were used.

To form of data, I used GDP per capita in EUR units per inhabitant as a percentage of the EU average. I use the data on an annual basis to keep the form of the dataset.

2.1.3 Disposable income per capita

Disposable income is the amount of money that households have available for spending or saving after deduction of taxes and the other expenses like social and health insurance. It is an indicator, where I expected a positive correlation with education because higher education should mean better-paid job as mentioned in the theoretical part.

I chose disposable income as another indicator sides to GDP that is showing us level people's well-being in the given region. Disposable income is used in real terms, expressed in Euros per inhabitant. The data are collected by Eurostat annually for NUTS II regions, as in the example of GDP. Also, I was testing disposable income converted to its PPS form too. But, the results of regression analysis were not that clear and have lower significance, as in the example of GDP.

2.1.4 Density of population

The density of population is the ratio of the annual average population of the NUTS II region expressed as the absolute value of the average people living in a square kilometre area (population/km^2). This variable is gathered as many other by Eurostat (Eurostat, n.d.) per NUTS II regions. Population density is used in this research because of the heterogeneity of NUTS II regions. This heterogeneity is caused due to the size of NUTS II regions, they are quite big to comfortably aggregate all density levels that could be found within the regions because we can often find both rural and metropolitan areas in the NUTS II regions and both have the different output of an economy. We cannot estimate rural area to be as productive and attractive for that many types of workers and firms as a metropolitan area. So, it makes no sense to compare these regions without concerning mentioned differences. I used this density of population as an independent variable because of clearing the results of regression analysis and I suggest a positive correlation with all other variables mentioned above. The purpose of using the density of population as an independent variable is to quantify the level of urbanization and examine it with other mentioned variables, especially education.

The idea is to examine whether there is a positive correlation between education and density of population and how strong this relationship is. Eventually, how strong is also the relationship of density of population with GDP and Disposable Income compared to educational attainment.

2.1.5 Average temperature and coefficient of variation

Last, I decided to add the average temperature and coefficient of variation of temperature. These two indicators were chosen as non-economic variables that should help with clearing the effects of other indicators on the results of regression analysis by adding non-economical characteristic of the region that is not changeable.

Temperature data were collected with big difficulties because there is no accessible dataset of average monthly temperatures for NUTS II regions. Even Eurostat does not gather this kind of data. Therefore, I had to create the dataset of temperature's data by myself with the help of two websites. Most of the data were collected on the Weatherbase.com and if the data there were missing, I used another website

Infoclimat.fr to find the rest of the desired data. But these websites contain only data for cities, and they are not aggregated to regional's data, hence I had to pick the biggest city in each of NUTS II region and use it to represent the weather data for the region. I used the biggest cities as the urban areas are producing a bigger share of regional GDP than rural areas.

The data are in the form of degrees Celsius (°C) and they were collected as average monthly temperature for last years. From these data, I was able to calculate the average annual temperature and also the standard deviation of average monthly temperatures using an electronic spreadsheet program MS Excel. The average temperature was calculated using the "AVERAGE" formula on the temperature data of all months per each region. The standard deviation of average monthly temperatures was calculated similarly, except that the "STDEV.S" formula was used. Both weather indicators are needed as there are big differences in the average temperates and even the standard deviations across Europe and both indicators need to be considered in the context of its influence on regional development as it has a confirmed impact on society and the nature of the region, but it really depends on how much developed the region is. Many wide-ranging effects of higher temperature have been proved with reduction of economic growth in poor counties and reduction of growth rates (not just the level of output) in the foreground (Dell et al., 2012).

2.2 Methodology

All collected data gave me a proper background to build up my analysis. But, before the analysis itself could be carried out, it was needed to straighten the data out to some usable format. All data must have been put in time series beginning in the year 2007 and ending in 2017. This way I made a dataset with sequential timing of variable values for each NUTS II region and each year separately. This dataset was created using MS Excel and it was necessary to sort the data specifically to see all the indicators for each year and region in one row. In such a way, I could get the regression analysis of time series to work.

Thus prepared dataset was ready to be analysed in Stata, which is software for statistics and data science. It is a very useful program due to its ability to perform regression analysis and other related tasks, such as Hausman Endogeneity Test. This test was also needed in my research to choose whether to interpret the random-effects model or fixed-effects model. Both of these models will be described later on. Stata works with input of the dataset in the form of a spreadsheet. The great advantage of using Stata is the simplicity of its coding commands. I needed only a few codes to perform the regression analysis, and the logic behind the code is quite easy to understand.

With the help of Stata, I could create a panel data and use regression analysis to examine the dataset. Panel data are multidimensional data that involve an element of time (time measurement) for some unit (regions, individuals, firms, etc.). It combines cross-sectional and time-series data models. Thus, the panel data contain one or more indicators of some phenomena for certain time periods for some unit (Mishra, 2018). Regression analysis is a statistical method providing an examination of the relationship between some range of indicators and identifying which variables correlate with the subject of interest (Foley, 2018). In this way, we can quantify the relationship between the indicators, then determine if there is any correlation and understand how strong the relationship is. Asides from that, we also get and the probability of existing relationship and therefore we know how much we can trust these calculations.

The regression analysis works by picking up the dependent variable, such as educational attainment for example. The dependent variable is some indicator that we are about to predict and which we are trying to understand. Together with this one dependent variable, we must pick one or more independent variables such as disposable income, the density of population, average temperature and others. The relationship of this variable is investigated by examining how these independent variables influence the dependent variable. In this hypothetical example, I gave above we would examine the impact of all independent variables (disposable income, the density of population, the average temperature, etc.) on educational attainment.

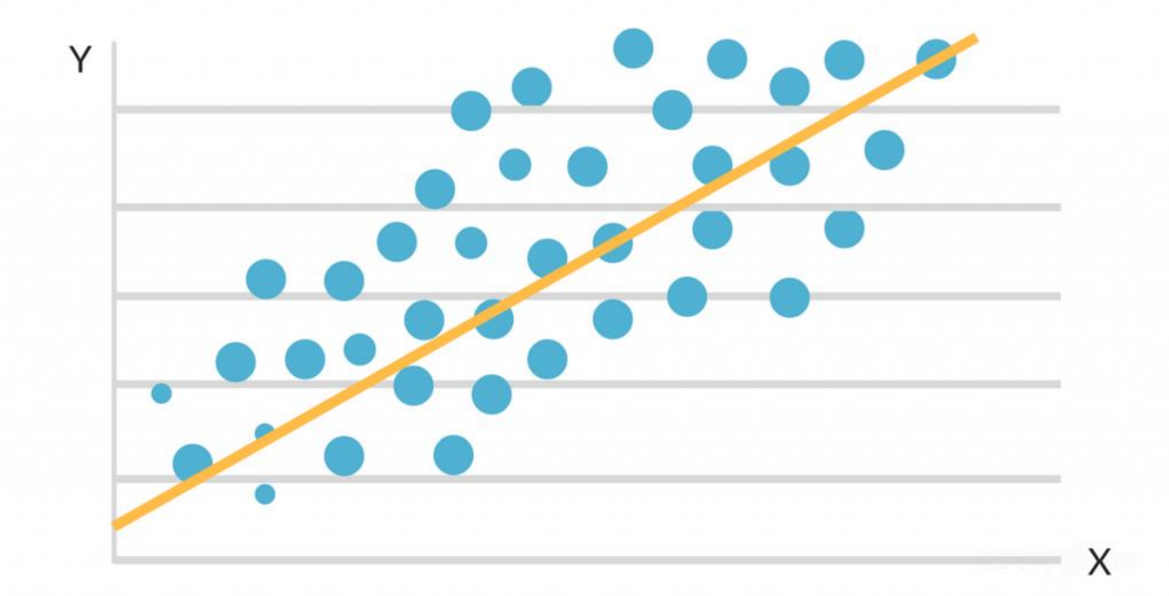

Figure 4: Chart of regression analysis (source: surveygizmo.com, 2018)

Figure 4 is a graphical visualization showing an example of data plotting. For better understanding, we can imagine one dependent variable that should be plotted on axis Y and one independent variable on axis X. For example, let educational attainment become a dependent variable and disposable income an independent variable. In other words, when we plot the data such as in Figure 4, we get the number of observations represented as blue points, where each point shows data for one specific year. These blue points are showing levels of educational attainment and disposable income for the given year. The yellow line (regression line) is showing us the trend and it is also the first explanation of the relationship between the dependent and the independent variables (Foley, 2018; Gallo, 2015). This yellow line is a graphical visualization of the regression analysis output, that can be also expressed by the following formula:

$$Y = a + bX + u$$

This is just an example of a simple linear regression showing that if X (disposable income) increase, then Y (educational attainment) goes up by coefficient b. The letter u stands for the error term, which is considered because in reality there is nothing like a perfect prediction (we can never now all variables that influence the dependent variable). It is just a very modest explanation of simple regression, but it was needed for ongoing research and to understand three methods for regression analysis that I used: Pooled regression by OLS (ordinary least squares), Fixed Effects Estimation and Random Effects Estimation.

2.2.1 Pooled-OLS

This type of method is the most basic one from the used methods in this thesis and is suitable for analysing the panel data and the estimation of unknown parameters (β) in a linear regression model. Specific effects of individuals are neglected here. Even over the fact that we are analysing the panel data, each row of the dataset is treated as a separate observation so it neglects the fact that the data is a panel data. Specifically, it leaves out the multidimensional structure of panel data (time and unit element) and analyses just the influence of independent variables on the dependent variable. If the explanatory variables are exogenous, then OLS will be unbiased and consistent (Hsiao, 1999; Mishra, 2018).

The example of Polled OLS model's formula suited for our model can be formulated like (Hsiao and Pesaran, 2004):

$$Y_{it} = \beta_1 + \beta_2 X_{it2} + \dots + \beta_k X_{itk} + u_{it}$$

Where Y stands for the dependent variable and X stands for the independent variables. The subscript i indicates the region in our example and the second subscript t indicates the year of observation. The Greek letter β stands for the unknown parameters as mentioned above and u is an error term.

2.2.2 Fixed Effects Estimation

The Fixed Effect Model works more sophisticated than Pooled OLS Model because it estimates the relationship between the dependent and independent variables within the groups (which consist of individual variables) for all reference years per individuals, such as the region in the example of this research. Each of these observed groups has its own individual characteristics that should have an impact on the outcome (Torres-Reyna, 2007). But, the fixed-effects model is a statistical method where variables that are constant across individuals like age, sex etc. are treated like fixed or non-random (Allison, 2009). Therefore, it is not suitable for estimating the effects of variables that do not change over time because the result of such variables is omitted (this behaviour of fixed effect model will be seen in the following regression tables). Also, it's not useable for estimating the effects of the average annual temperature of the regions and also the standard deviation of the average monthly temperature of the regions (these indicators were calculated from the aggregated data of many years and therefore their values are not changing over time - they are time invariate and this model considers their influence as equal to zero).

Inability to estimate the effects of temperature data is quite a disadvantage compared to Pooled OLS, where we can estimate these effects comfortably. On the other hand, the big advantage of this model in comparison with Pooled OLS is that it takes the regions and estimates effects in the context of these regions, so it should include more the characteristics of individual regions. And it is needed to mention that in comparison to the random-effects model (which is perceived as a contrast to the fixed-effects model) there is another disadvantage: in many cases, estimates may have larger standard errors causing bigger p values (Allison, 2009).

The equation for the fixed effects model looks like (Econometrics Academy, n.d.):

$$Y_{it} = \alpha_i + \beta_1 X_{it} + u_{it}$$

Where Y is the dependent variable for the entity i and time t, α is the unknown entity-specific intercept for each observed sample. Greek letter β represents the coefficient for each independent variable X. And u stands for the error term.

2.2.3 Random Effects Estimation:

The Random Effects Estimation model works like a contrast to the fixed-effects model as it threads all the variables as random variables (Ramsey and Schafer, 2012). This method of regression analysis is able to cover also the differences across observed groups (Torres-Reyna, 2007) and therefore I believe that this method is more suitable for this research because I assume the nature of regions (their characteristics) have an effect on our dependent variable (either it is education, disposable income or GDP). So, this model covers the effects of within-entity and also between-entity in comparison with Fixed-effects model that is able to estimates effects only within the entity. I am trying to partially cover this nature of region using temperature data (because the temperature itself, as non-economic factor have an effect on the economy outcome (Dell et al., 2012)) which I could not examine using Fixed-effects model due to its form as time-invariant variables. The big advantage of Random-effects model is the fact that I can examine these time-invariant variables.

The standard equation of the Random-Effects model looks like (Econometrics Academy, n.d.):

$$Y_{it} = \beta X_{it} + \alpha_r + u_{it} + \varepsilon_{it}$$

Where Y is an output of the equation is the dependent variable for the entity i and time t. Greek letter α is the unknown entity-specific intercept for each observed sample and β represents the coefficient (unknown vector) for each independent variable X for each entity i and time t. The error terms are more complicated here than in the fixed-effects model because it has two types of them. It has an error term u_{it} that stands for the between-entity error term and error term ε_{it} that specifies the within-entity error.

2.3 The relationship between education and disposable income per capita in European regions NUTS II

In this part, I am about to examine the relationship between educational attainment and disposable income. But, also the effects of other independent variables such as the density of population, average temperature and standard deviation of temperature on both of these mentioned variables will be examined. The first panel data regression was run in two variations of methods for panel data regression analysis: fixed-effects model and the random-effects model, which are most commonly used for this kind of research. Pooled-OLS model had not been included in analysis yet, because I am about to examine the changes in time and cover the regional differences. Mentioned models were used in the following forms of linear regression:

Fixed effects:

$$LnEdu_{rt} = \alpha_r + LnIncome_{rt}\beta + LnDensity_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + u_{it}$$

Random effects:

$$LnEdu_{rt} = LnIncome_{rt}\beta + LnDensity_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + (\alpha_r + e_{it})$$

Where:

Edu - Share of people with higher education in NUTS II region r in year t

Density - The number of persons per km^2 in a NUTS II region r in year t

Income - Average disposable income per capita in a NUTS II region in year t

AvTemp - Average temperature in the NUTS II region r

SDTemp - Standard deviation of average monthly temperature in the NUTS II region r

α_r - Region-specific intercept

u_{ti}, e_{it} - Region-specific error terms

β - slope of the independent variable

These models were run in Stata and the results could be seen below in Table 1: Education Panel data 2007 - 2017.

Table 1: Education Panel data 2007 - 2017

	Ln of Education				*Ln of Education*				*Ln of Education*			
	Fixed Effects		*Random effects*		*Fixed Effects*		*Random effects*		*Fixed Effects*		*Random effects*	
Ln of Income (per capita)	0.4300	***	0.3488	***					0.4223	***	0.3371	***
	(0.2842)		(0.2280)						(0.2844)		(0.0228)	
Ln of density					1.1493	***	0.7576	***	0.1098	***	0.0544	
					(0.0344)		(0.1396)		(0.0328)		(0.0135)	
Average temperature	omitted		-0.0559	***	omitted		-0.4779	***	omitted		-0.0331	***
			(0.0058)				(0.0058)				(0.1469)	
SD of Temperature	omitted		-0.5588	***	omitted		-0.1047	***	omitted		-0.3987	**
			(0.0144)				(0.1456)				(0.1469)	
Intercept	-0.8729	***	0.6513	*	2.4673		4.0483		-1.3574	***	0.3801	
	(0.2709)		(0.2863)		(0.1751)		(0.1621)		(0.3066)		(0.2927)	
Nr of observations	2 302		2 302		2 302		2 302		2 302		2 302	
Nr of groups	275		275		275		275		275		275	
F-stat (p-value)	70.39				77.08				68.20			
	(0.00)				(0.00)				(0.00)			
R2 within	0.1016		0.1016		0.0092		0.0092		0.1065		0.1058	
R2 between	0.2371		0.3417		0.1072		0.3294		0.2733		0.3658	
R2 overall	0.2377		0.3328		0.0871		0.3078		0.2580		0.3478	
Rho	0.8928		0.8692		0.8991		0.8631		0.8972		0.8652	
Hausman Test: Chi-square (p-value)			22.95				5.46				30.17	
			(0.00)				(0.02)				(0.00)	
Wald Test: Chi-square (p-value)			357.52				148.28				379.38	
			(0.00)				(0.00)				(0.00)	

robust errors in parentheses

* p<0.1 **p<0.01 ***p<0.001

(note: Hausman tests run on standard, non-robust fe and re models)

In Table 1, there are six columns with calculated regression in each of those. The results can be seen for the natural logarithm of Educational attainment as the dependent variable and natural logarithm of disposable income, the density of population, average temperature and standard deviation of the average monthly temperature as the independent variables.

The most important (as it covers all selected variables) results could be seen in the last two columns for both fixed-effects and random-effects models. These columns will be described first. Another four columns are tested with the omission of some of the independent variable and their purpose is to get some further information about the behaviour of these variables. They will be described when needed (if some interesting result appears).

All regression analysis in Table 1 were made from 2302 observations divided into 275 groups. To choose which model should be interpreted as most appropriate, I used the Hausman Test for Endogeneity, which showed p-value = 0,00. This value, which is less than 0,05 means the rejection of the null hypothesis and recommends the fixed-effects model to be described.

The results of the Fixed-effects model showed highly significant effects when F-stat's p-value is equal to 0,00. The regression showed that 26 % of educational attainment is explained by all examined independent variables ($R^2 = 0{,}2580$). Other 74 % stays unexplained. The main outcome of this model is that when the disposable income will double, we estimate an increase in education attainment by 42 %, ceteris paribus. For example, if we take the hypothetic region with the average disposable income of 10,000 EUR and educational attainment 30 % and if the disposable income would double the educational attainment will increase by 42% to 42,6 % ($30 \cdot 1{,}42 = 42{,}6$). This result is shown by values of slope β for Ln of the disposable income. This relationship is highly significant as the p-value is less than 0,001 (this relationship exists with the possibility of 99,9 %).

The estimation for the density of population is also highly significant (99,9% possibility of existing relationship) and has a strong positive correlation shown by the value of slope $\beta = 0{,}1098$. It shows that an increase in the density of population by 100 % should cause an increase in the education attainment by 11 %. Because the Hausmann test rejects the null hypothesis and I should interpret the fixed-effects model there are no results for the average temperature and standard deviation of average monthly temperature. Results of both of these variables were omitted, as there is no change in the time and fixed-effects can't calculate with them due to it. But when we use the results of the Random-effects model we can see the really high significance for the average temperature and the strong negative correlation with the educational attainment. The similar results show the standard deviation of the average monthly temperature but with lower significance but much stronger correlation. These results show that higher temperature and higher standard deviations have a negative impact on educational attainment.

In the first two columns, I run the regression when the effect of density of population is omitted. The formulas for these two columns look like:

Fixed effects:

$$LnEdu_{rt} = \alpha_r + LnIncome_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + u_{it}$$

Random effects:

$$LnEdu_{rt} = LnIncome_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + (\alpha_r + e_{it})$$

The Hausmann test for this model shows that the fixed-effects model is more suitable (p-value $< 0{,}05$) than the random-effects model. The regression shows that the R^2 has declined to 23,77 value, which means that almost 24 % of education attainment is explained by examined independent variables. But the effects of the disposable income on educational attainment increase to exactly 43 % in educational attainment when the disposable income increase to its double.

I decided to examine also the effect of sole density of population on educational attainment. This can be seen in the third and fourth column. The formulas for this relationship are:

Fixed effects:

$$LnEdu_{rt} = \alpha_r + LnDensity_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + u_{it}$$

Random effects:

$$LnEdu_{rt} = LnDensity_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + (\alpha_r + e_{it})$$

The results of this model show that again the fixed-effects model should be described as a more appropriate one. This model is also highly significant and shows a highly positive correlation between the density of population and educational attainment. The interesting fact is, that the intensity of this variable increase incredibly. In the first described model the slope β was equal to just 0.1098 and now it goes up to

1.1493 and again with 99,9 % possibility of the existing relationship. This means that the effect increased 10 times when omitting the effect of the other independent variables. This is likely to be due to intercorrelation between disposable income and density of population. This intercorrelation is, however, not strong enough to cause a major multicollinearity issue. The result should be read as that the third model is showing the effect of density controlled for disposable income or vice versa. Also, as we omitted other independent variables the R^2 has dropped down to just 9 % and letting other 91 % of the relationship unexplained, which make sense when omitting the effects of disposable income.

In Table 2: Disposable income Panel data 2007-2017 I made similar regression analysis with the addition of Pooled OLS model and I also moved the natural logarithm of education to a place of the independent variable and moved natural logarithm of disposable income on the place of the dependent variable. Furthermore, this model was made to examine the relationship from the other side and to see the differences. In this table below we can see regression run on three models: Pooled OLS, Fixed-effects, Random-effects. Let's again start with the right side of Table 2, specifically the last three columns which covers all examined variables. The formula for each model looks like:

Pooled-OLS:

$$LnIncome_{rt} = \alpha + LnEdu_{rt}\beta + LnDensity_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + u_{it}$$

Fixed-effects:

$$LnIncome_{rt} = \alpha_r + LnEdu_{rt}\beta + LnDensity_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + u_{it}$$

Random-effects:

$$LnIncome_{rt} = LnEdu_{rt}\beta + LnDensity_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + (\alpha_r + e_{it})$$

Hausmann test shows the Fixed-effects is more appropriate to be described. The overall significance of the model is high as well as the relationship between disposable income and educational attainment. The results show that when educational attainment double the disposable income grow by 23 % which more than less fits the previous results, but the equation cannot be reversed due to strong regional-specific intercepts. R^2 results also did not change much, just dropped approximately by 1%.

Table 2: Disposable income Panel data 2007-2017

	LnIncome			*LnIncome*		
	Pooled OLS	*Fixed Effects*	*Random effects*	*Pooled OLS*	*Fixed Effects*	*Random effects*
LnEducation	0.4476***	0.2362***	0.2436***	0.4167***	0.2325***	0.2392***
	(0.0288)	(0.1561)	(0.0154)	(0.2923)	(0.1566)	(0.0154)
LnDensity				0.0425***	0.0589*	0.0500**
				(0.0081)	(0.0244)	(0.0167)
Average temperature	-0.3111***	omitted	-0.0324***	-0.0322***	omitted	-0.0317***
	(0.0033)		(0.0088)	(0.0032)		(0.0088)
SD of Temperature	-0.1431***	omitted	-0.1825***	-0.1323***	omitted	-0.1661***
	(0.0079)		(0.0207)	(0.0082)		(0.0214)
Intercept	9.3312***	8.7703***	10.2273***	9.1575***	8.4827***	9.8761***
	(0.1465)	(0.0504)	(0.1996)	(0.1495)	(0.1293)	(0.2312)
Nr of observations	2 302	2 302	2 302	2 302	2 302	2 302
Nr of groups		275	275		275	275
F-stat (p-value)	381.58	282.94		296.24	274.22	
	(0.00)	(0.00)		(0.00)	(0.00)	
R2 within		0.1016	0.1016		0.1041	0.1041
R2 between		0.2371	0.3249		0.2343	0.3320
R2 overall		0.2377	0.3125	0.3403	0.2454	0.3238
Rho		0.9743	10.2273		0.9730	0.9665
Hausman Test: Chi-square (p-value)			8.24			6.34
			(0.00)			(0.04)
Breusch-Pagan LM test			736.90			787.79
			(0.00)			(0.00)
Wald Test: Chi-square (p-value)			361.00			371.23
			(0.00)			(0.00)

robust errors in parentheses

* p<0.1 **p<0.01 ***p<0.001

(note: Hausman and BPLM tests run on standard, non-robust fe and re models)

As well as within the previous model I am not able to describe the effects of the temperature as they are omitted in the Fixed-effects model. But unlike in Table 1, here I can use Pooled-OLS model which shows almost the same results as Random-effects for the temperature data, which supports the results. It shows us that both average temperature and standard deviation of the average monthly temperature have a strong negative effect on disposable income, but not that strong as on educational attainment.

Since the Pooled-OLS model has now been used, we can also describe what does that mean in the context of other models. We can see, that when we do not include the effects within and between regions, the R^2 increase to 0,3403 while having still the very significant model. Also, the effects of education and the density of population are still significant, but the effect of educational attainment is much stronger now with the slope $\beta = 0{,}4167$ which is almost double than the values in Fixed-effects or Random-effects models. Considering the differences between the models, I can state that educational attainment has a very significant and strong effect on disposable income. But, the effect of education has not been that strong in the examined period as it should be (according to the difference between estimations of Pooled-OLS and Fixed-effects models described above in the methodology part). It probably could be caused by the post-recession period after the big financial crisis in 2008 when the recovery of regions was not only led by technological progress or high-education requiring positions. Possibly, the increment of the disposable income was caused by other unknown variables that were not examined here. However, the regions with higher educational attainment still grew better, just not that much as estimated.

Aggregating the results of both tables, I can state that there is a strong positive correlation between disposable income and educational attainment. But we can't define the direction of causality, whether educational attainment affects disposable income or vice versa. The relationships probably exist in both directions because the higher educated person could get a better job and likewise, a person with higher disposable income than the average can spend more money on education. On the other hand, there is quite a lot of unknown (because the R^2 is just about 24-34 % for described models) for the third variable such as race or average age. I missed those variables, but it is at least a good opportunity for further studies. Also, we should keep in mind that investing in higher education is not an immediate solution and it takes some time to effects the regional economy, which also influences the results.

2.4 The relationship between education and GDP per capita in European regions NUTS II

In this part, I will use Stata again and same models of regression analysis as in the previous part to examine the relationship between GDP per capita, the density of population, average temperature and standard deviation of the temperature as independent variables on educational attainment as the dependent variable. In Table 3: Education Panel data 2007-2017 is the first tested model. The most relevant results can be seen in the last two columns on the right side of the table. Formulas of this regression look like:

Fixed effects:

$$Edu_{rt} = \alpha_r + Density_{rt}\beta + GDP_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + u_{it}$$

Random effects:

$$Edu_{rt} = Density_{rt}\beta + GDP_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + (\alpha_r + e_{it})$$

Where:

Edu	- Share of people with higher education in NUTS II region r in year t
Density	- The number of persons per km^2 in a NUTS II region r in year t
GDP	- Gross domestic product per capita in a NUTS II region in year t
$AvTemp$	- Average temperature in the NUTS II region r
SDTemp	- Standard deviation of average monthly temperature in the NUTS II region r
α_r	- Region-specific intercept
u_{ti}, e_{it}	- Region-specific error terms

The results of this regression analysis are unfortunately not that clear as in the case of disposable income because the Hausmann model proposes the fixed-effects model to be more appropriate. But this model shows the low significance for the effect of GDP on education ($p<0,1$).

Table 3: Education Panel data 2007-2017

	Education		*Education*		*Education*	
	Fixed Effects	*Random effects*	*Fixed Effects*	*Random effects*	*Fixed Effects*	*Random effects*
GDP (per capita)	0.0035	0.0386 ***			0.0034	0.0383 ***
	(0.0084)	(0.0059)			(0.0084)	(0.0058)
Density of population			0.0001 **	0.0002 ***	0.0001 **	0.0002 ***
			(0.0000)	(0.0000)	(0.0000)	(0.0000)
Average temperature	omitted	-1.0410 ***	omitted	-1.2423 ***	omitted	-1.0389 ***
		(0.1475)		(0.1530)		(0.1447)
SD of Temperature	omitted	-2.8912 ***	omitted	-3.4053 ***	omitted	-2.8590 ***
		(0.3500)		(0.3608)		(0.3435)
Intercept	26.6942 ***	53.5510 ***	26.9774 ***	62.6126 ***	26.6407 ***	53.2665 ***
	(0.8271)	(3.3970)	(0.0625)	(3.2689)	(0.8259)	(3.3374)
Nr of observations	2 302	2 302	2 302	2 302	2 302	2 302
Nr of groups	275	275	275	275	275	275
F-stat (p-value)	67.00		91.27		65.69	
	(0.00)		(0.00)		(0.00)	
R2 within	0.0001	0.0001	0.0040	0.0040	0.0041	0.0021
R2 between	0.3201	0.4070	0.1443	0.2896	0.3217	0.4167
R2 overall	0.2976	0.3901	0.0721	0.2883	0.2380	0.3985
Rho	0.9198	0.8671	0.9208	0.8810	0.9191	0.8625
Hausman Test: Chi-square (p-value)		34.44		124.16		202.33
		(0.00)		(0.00)		(0.00)
Wald Test: Chi-square (p-value)		174.73		130.14		194.49
		(0.00)		(0.00)		(0.00)

robust errors in parentheses

* p<0.1 **p<0.01 ***p<0.001

(note: Hausman tests run on standard, non-robust fe and re models)

Because of these results, I won't describe the effects as detailed as in the case of disposable income. From the results shown in Table 3, we know that there is a positive causality between GDP per capita and educational attainment (both models suggest this positive correlation, they only differ in the intensity of this correlation). The R^2 of the fixed-effects model is 0,2380, which shows that 24% of educational attainment is explained by examined independent variables. The random-effect model provides much higher R^2 value explaining almost 40% of educational attainment. This shows us that the average temperature and standard deviation of the average monthly temperature play an important role in the explanation of educational attainment. In the rest of the table, I will not describe any results because there is no interesting result, besides already described facts.

To get some better results, I decided to switch the dependent variable from education to GDP and put education into the model as an independent variable. Also, I decided to use Pooled-OLS model to get some further knowledge of the examined relationship. The outcome is shown in Table 4 and the last three columns on the right side will be described first. Formulas for these models look like:

Pooled-OLS:

$$GDP_{rt} = \alpha_r + Edu_{rt}\beta + Density_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + u_{it}$$

Fixed-effects:

$$GDP_{rt} = \alpha_r + Edu_{rt}\beta + Density_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + u_{it}$$

Random-effects:

$$GDP_{rt} = Edu_{rt}\beta + Density_{rt}\beta + AvTemp_{rt}\beta + SDTemp_{rt}\beta + (\alpha_r + e_{it})$$

Hausmann again suggests using the fixed-effects model as it is more appropriate. But, the significance of the education's effect on GDP did not get better and in addition, the significance of the whole model is low as the p-value of F-stat is 0,9028. However, the model is at least proving the positive correlation between these variables. The interesting fact is that the R^2 value increased. Which means that education explains more of GDP than vice versa.

The results of Pooled-OLS are much more useful as all of the effects are very significant. The probability that the relationship exists is more than 99,9% (P-value for all independent variables is lower than 0,001). The R^2 value is also much higher as it also includes the effects of temperature and therefore, I will describe this model as it is the most informative and stable. The effect of education on GDP is shown as positive and very strong. If educational attainment increases by 1%, then GDP per capita grows by 2,5% (because both of variables are expressed as a percentage).

Table 4: GDP per capita Panel data 2007-2017

	GDP per capita			*GDP per capita*		
	Pooled OLS	*Fixed Effects*	*Random effects*	*Pooled OLS*	*Fixed Effects*	*Random effects*
Education	3.1089 ***	0.0247	0.1500 *	2.4586 ***	0.0241	0.1119 *
	(0.0996)	(0.0589)	(0.0593)	(0.1175)	(0.0590)	(0.0594)
Density of population				0.0045 ***	0.0000	0.0001
				(0.0005)	(0.0001)	(0.0001)
Average temperature				-2.7852 ***	omitted	-5.1515 ***
				(0.3326)		(0.9337)
SD of Temperature				-3.2772 ***	omitted	-13.9437 ***
				(0.8130)		(2.2038)
Intercept	14.3392 ***	97.7375 ***	93.2175 ***	80.5251 ***	97.7444 ***	236.9565 ***
	(2.8697)	(1.5993)	(3.3897)	(9.6174)	(1.6002)	(20.2220)
Nr of observations	2 302	2 302	2 302	2 302	2 302	2 302
Nr of groups		275	275		275	275
F-stat (p-value)	974.67	0.18		291.40	0.10	
	(0.0000)	(0.6743)		(0.0000)	(0.9028)	
R2 within		0.0001	0.0001		0.0001	0.0001
R2 between		0.3201	0.3201		0.3423	0.1498
R2 overall	0.2976	0.2976	0.2976	0.3366	0.3132	0.1605
Rho					0.9836	0.9746
Hausman Test: Chi-square (p-value)			332.14			204.73
			(0.00)			(0.00)
Breusch-Pagan LM test			2029.82			8139.77
			(0.00)			(0.00)
Wald Test: Chi-square (p-value)			6.41			59.59
			(0.01)			(0.00)

robust errors in parentheses

* p<0.1 **p<0.01 ***p<0.001

(note: Hausman and BPLM tests run on standard, non-robust fe and re models)

The effects of temperature are also interesting here because we got very significant results for both variables. And both variables influence the GDP very strongly. As for income, even here the correlation is negative, and the effects are even stronger than the effect of education.

Let's now focus on the left side of the table, there are models when omitting the effects of all other independent variables except education. Formulas for this regressions looks very simple here:

Pooled-OLS:

$$GDP_{rt} = \alpha_r + Edu_{rt}\beta + u_{it}$$

Fixed-effects:

$$GDP_{rt} = \alpha_r + Edu_{rt}\beta + u_{it}$$

Random-effects:

$$GDP_{rt} = Edu_{rt}\beta + (\alpha_r + e_{it})$$

I will continue here with describing Pooled-PLS model as it was described also above, and it is providing the highest significance. Because of omitting other independent variables, the R^2 values decreased but the effect of the education on the GDP increased. Here, we can also nicely see that the effects are much stronger when ignoring time factor (Pooled OLS is showing much stronger correlation than other two models), which is pointing at non-nature behaviour of GDP increment in the examined time period (the increment were not probably caused as much by educational attainment as it should). The interpretation of this model is that if we increase educational attainment by 1%, then GDP per capita increases, ceteris paribus by 3,1% (for Pooled-OLS model).

From all of the outcomes that were provided in tables 3 and 4, I can state that there is a positive correlation between education and GDP. But again, I can't say the direction of the causality. We can only estimate that the causality works on both sides, but it is much stronger in the direction from education to GDP. For some reasons, the fixed-effects and random-effects model very not much very useable here, but I was not able to discover the reason. I estimate that it was caused probably by the examined time period, for the same reasons as in the previous case when I was testing disposable income. The effects of rising economies in Europe after the recession were stronger than in the case of income and therefore, I could only use Pooled-OLS model to get some significant results. It's much more significant here, that education is a long-term investment that pays much later for GDP than for disposable income.

2.5 Return of Education

The purpose of this part is to examine if it pays off for students to spend their time and money for university education. This issue is affected by many variables. For example, it is needed to discuss what is the motivation for the people to get higher education and what drives and motivates them to continue with schooling and getting higher degrees. There is something like common knowledge about which everyone has heard from someone during their childhood, that studying, and school is essential for us if we want to succeed. We are convinced, that if we will be hard-working students and get good marks then we will become successful, respected and get a high-paid job, just like that. Simply, the idea is that good marks are key to become whoever we want, at least here in the Czech Republic it is a common attitude. There are many questions on this topic that needs to be answered and discussed for better understanding the importance of education, i.e., "What is the link between education and future earnings?", "Does studying really pays off ?" or "Do you really need a degree?" etc. To challenge this idea and to find the answers I tried to discover how strong is the link between education and future earnings and whether the studying really pays off.

What is the link between education and future earning?

According to the article by Patrick Gleeson, PhD the link between education and earnings is more than clear. He claims that the more education you have, the more money you will make in the future (Gleeson, 2018). Gleeson bases his opinion mainly on the article *Education and Lifetime Earnings in the United States* (Tamborini et al., 2015). In this article, Tamborini, Kim and Sakamoto found out that the lifetime earning a gap in the United States between high school graduates and college graduates is around $1,130,000 for men and $792,000 for women on average. So, as can been seen there is some correlation between the level of education and earnings, but for better understanding, it has to be investigated much deeper.

Does studying really pay off?

When we take a closer look at cost of a college education at very top-tier US universities, which is calculated including fees and living expenses, students costs go up to around $50,000 per year (The College Board, 2019). In comparison with the average lifetime earning a gap, it seems like a really good investment, because the cost of university studies returns to students throughout their lives in multiple higher income. The return of investment into higher education looks just amazing even when the average lifetime gap calculated for average graduates as compared to the cost calculated for the top-tier universities. According to The College Board statistics, the costs for the average university drop down to half of that cost. This statement is supported by another report published by HSBC Holdings PLC. According to this report, the average amount of money that is spent by students during their studies is $99,417 (HSBC Bank USA, 2018). But that amount of money is not calculated the same way as the one from The College Board. The HSBC Bank's list of items covers all the student's expenses. The biggest items on the list are of course tuition fees and accommodation, which together make about half of the total amount. These two biggest items bring us near to the number given by The College Board. Now, we can be pretty confident, how high is the cost of education for US students. The difference between the reports is that the one published by HSBC Bank includes also other living expenses like food and groceries, entertainment credit cards, transport and others non-study relating items, which drive up the costs much higher. However, even if you would study the top-tier university and then you would get a salary of the average college graduates, it still pays off. But these numbers are just for universities in the United States. To found out what is the situation in the rest of the world and mainly in the European Union, I made a simple analysis of the return of investment in education.

Do you really need a degree?

My analysis is based on the article "*Why you REALLY need a degree"* written by Ing. Martin Lukavec, PhD. in 2017. I updated the charts from his article with the fresh data that I got from Organisation for Economic Cooperation and Development (OECD) to show the size of the pay gap between holders of tertiary education over secondary or lower education. Figure 5 shows us, how much more holders of bachelor's degree earn in comparison with the workers with secondary or lower education.

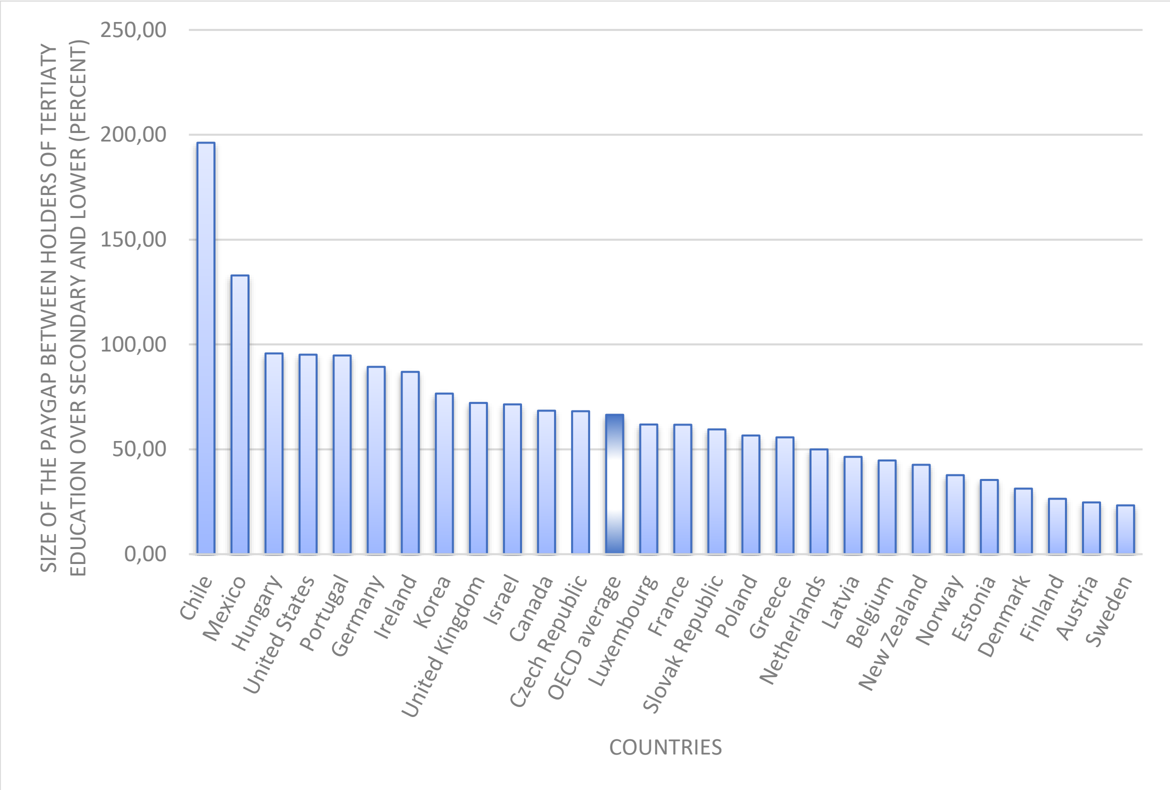

Figure 5: How much more you earn if you have a bachelor's degree (source: OECD, 2018)

I constructed this chart using the data from OECD's *Education at Glance 2018* (OECD, 2018), which is using data from 2016. The OECD average gap, highlighted in the chart, is 66,47 %, which is a really stunning number. It means that getting a higher education will give you 66,47 % higher earnings on average across the OECD countries. But what is even more interesting, and staggering are the differences between OECD counties. The biggest earning gap can be found in Chile, where the size of the pay gap is astonishing 196 %, which is three times bigger gap than the OECD average. Mexico is not far behind with a size of the pay gap of 133 %. But they are not alone on this side of the list, there are other countries where the earning gap does not differ much from them: Hungary (96 %), the United States (95 %), Portuguese (95 %), etc. In these all mentioned countries, it really pays off to get a bachelor's degree as you have a high probability to get a much higher salary than workers without a college degree. On the other side of the chart, there are countries like Sweden (23 %), Austria (25 %), Finland (26 %) or Denmark (31 %). For these countries, you really have to think thoroughly before investing in education as the difference between salaries of college degree holders and workers without higher education is not that high. It really needs to be mentioned that the dispersion of income gaps is truly staggering, and it says to us that it really depends, where you live.

As we compared the salaries of bachelor's degree holders to workers with secondary education or lower, it is logical to compare also master's degree holders to workers with no college education and see the difference between impacts of each level of the higher education on the salary. For this purpose, I made another chart. Figure 6 shows the same earning gaps as shown in Figure 5, but for master's degree holders.

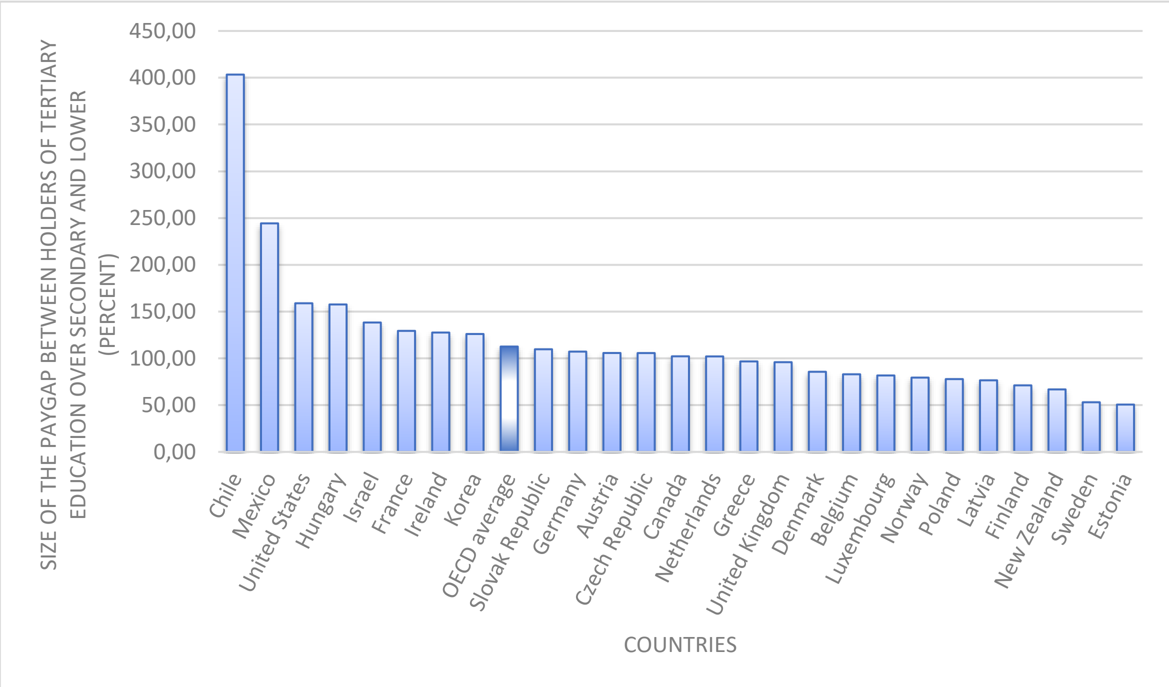

Figure 6: How much more you earn if you have a master's degree (source: OECD, 2018)

We can see that differences are logically higher than in Figure 5 with the bachelor's degree workers as the workers with the master's degree are more educated and can be assumed that they are capable of handling more complicated tasks. However, although the earning gaps of countries look similar as in Figure 5, mainly because the countries keep the trend, it is even more staggering when we look at for example Chile. We can see, that the earning gap increased from 196 % to unbelievable 403 %, which is more than twice that high as it is for bachelor's degree holders. Mexico changed similarly from 132 % to almost double 244 %. Generally, the difference in most of the countries is almost doubled, even in the countries with the lowest gap like Estonia (51 %) or Sweden (53 %). OECD average pay gap is 112,6 %, which gives the worker with a master's degree additional potential earning of 46,1 % on the worker with just bachelor's degree.

As the numbers for Chile and Mexico are almost unbelievable, there is a need to find an explanation. The unreal earning gaps in Chile and Mexico could mostly be influenced by the status of their educational system, which is not able to provide its counties with educated labour supply. There are just not as many university graduates to cover the demand of the labour market. Figure 7 shows a regression model with R2 (determination coefficient) equal to 0,4238, meaning that 42.38% of the variability of the dependent variable was included. This means that the model shows a strong regression and can be described.

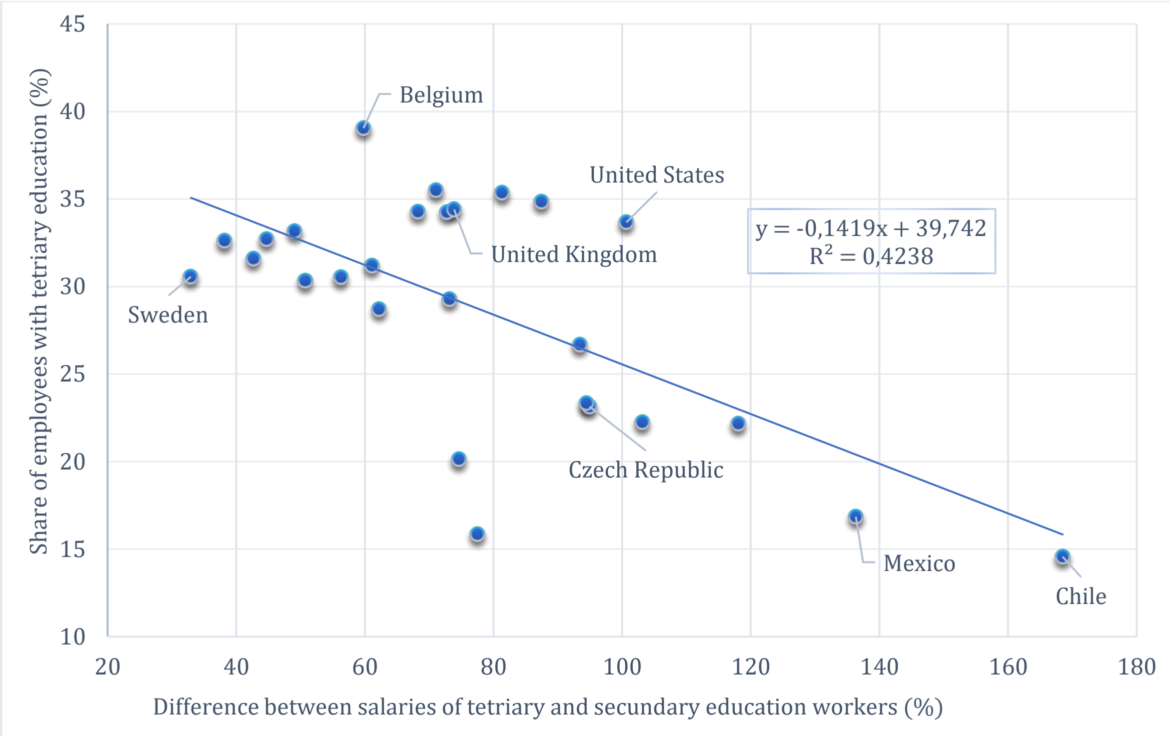

Figure 7: Pay gap between education levels (source: OECD, 2018)

The figure explains the relationship between the share of employees with tertiary education on the vertical axis and the difference between salaries of tertiary and secondary education workers on the horizontal axis. The regression model tells that a higher share of employees with higher education the smaller the pay gap between the levels of education. The regression models also show that the economy of countries above the regression line such as Belgium, the United States or the United Kingdom does need much more educated workforce than the economy of countries that are under the regression line like Chile and Mexico, and even the Czech Republic does not need that much-educated workforce.

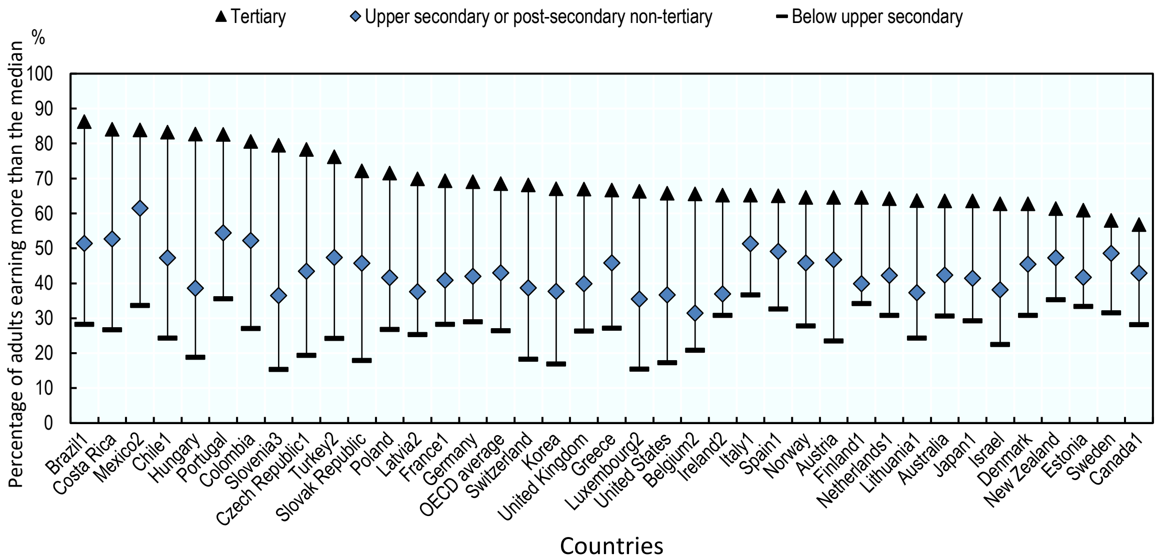

Figure 8: Percentage of adults earning more than the median (source: OECD, 2018)

In Figure 8: Percentage of adults earning more than the median (source: OECD, 2018), we can see the percentage of adults earning more than the median in each of the selected countries measured by OECD divided by educational attainment. It supports, even more, the idea that education influences future income. It shows, that the higher level of education means a higher share of people earning more than the median in the given country. Also, as we can see all of the selected countries follow the trend and not even one differs. So, the answer to the question whether you need a degree or not is more about where you live, because the earning gap significantly differs across the countries.

Does it matter which university do you study?

In addition to the location that we discussed above, there is another view that we must discuss in greater depth. It is crucial to focus on the fields of study that people choose. The most suitable tool that can be used to look deeper into this issue is the return of investment (ROI) of each university, which should show if there is some return from higher education and also whether the field of study is important in any way.

According to PayScale's 2018 report, which is comparing average earnings of graduates to the cost of degree (net of financial aid), some of the American universities have annual ROI 19 % which is just a dream invest compared to average annual returns for the S&P 500. Looking into S&P 500's annual returns the average ROI from 1957 to 2018 is 7,96 % (Maverick, 2019). This finding makes the investment into education more than twice as profitable as investing in the biggest American companies listed on stock exchanges.

From the above-mentioned report, we can proclaim that college is usually worth it, and by studying it you will get a pretty nice annual ROI. But there are also universities, where you can get negative annual ROI of even -16%, which is a huge difference showing that it really matters, where you study.

There are not many universities that get you to negative annual ROI. But most of those who do, have mostly supper high costs for studying (tuition fees). In 2015 The Economist took this report of PayScale and in Figure 9: Average annual return on a degree (source: The Economist, 2015) is shown that it really depends on the field of study more than on the specific university.

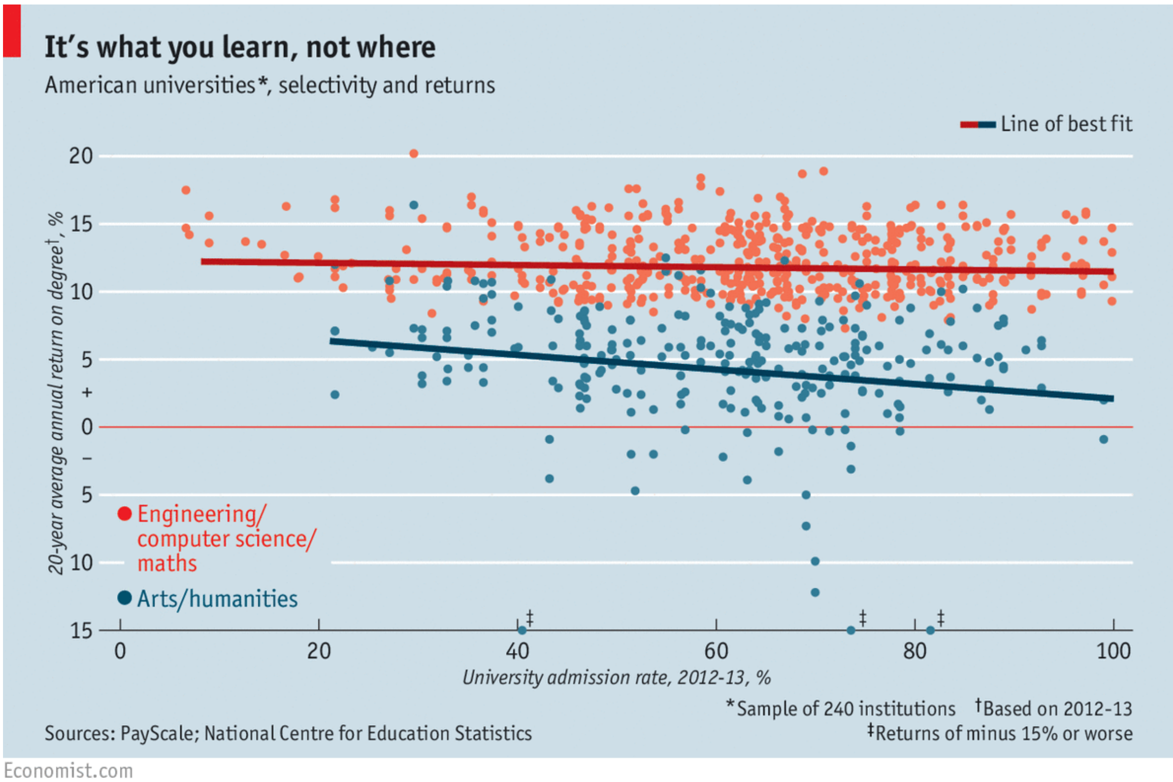

Figure 9: Average annual return on a degree (source: The Economist, 2015)

In this figure, we can see that all universities were divided into two groups based on the field of study: orange - Engineering/computer science/math field and blue - Art/humanities field. It is pretty clear, that it does not much matter what university admission rate is as what is the main field of study of the university. The orange group of universities have on average 5 % higher 20-year average annual return on degree than the blue group.

In a more detailed look, the best fields of study are engineering and computer science. They both have a 20-year annual return of 12% on average. Also, both business and economics degrees have a solid annual return of 8,7%. On the other hand, there are universities specialized mainly on art education that has negative 20-year net ROI. For example Maine College of Art with annual ROI -10,4 % and graduation rate of 57 % (PayScale, 2018). It means that slightly more than half of the students graduate from this university and then don't even profit from it.

Conclusion

The question of whether higher education pays off has been around for some years, but now when technologies go forward faster than ever before it's becoming very crucial to research this issue. It should also be said that this issue is not only addressed by policy-makers but even by individuals who are dealing with it when choosing whether to study or not (or what to study). That is why I focused my research in both directions and focused on the impact of education on the regional economy and on individuals. These directions are also very closely linked insomuch as if people get a higher education, they tend to get a better-paid position, as I confirmed this statement in my theoretical and even practical part. And if people do good in their professional life, also the regional economic benefits from it (through regional multiplication). As manufacturing is still losing jobs, this issue is very actual but not medialized and researched much. Therefore, there is still a big gap in research of educational effects on regional development. Also, it's sad to say that after my literature research I found out that most authors that are publishing in this field are from the United States and research of comparable quality is missing in the European context. Therefore, my literature research is based mainly on US authors.

In the theoretical part, I tried to summarize my research. But very soon I found out that it is almost impossible to cover the area of study within range and sources provided by the Bachelor Thesis. So, I did my best to cover the most important parts and give a brief overview of the researched topic. Despite that, I feel that some parts would need further explanation and there are many issues that I haven't mentioned there. However, I do believe that the theoretical part is fulfilling its purpose and it's giving the reader a proper and brief insight into the research topic. I also believe that the theoretical part prepares the reader enough to fully understand the outcomes of the practical part and ensures him of the importance of this research. The theoretical part starts with the basics, the reader should get enough knowledge to be able to describe what is human capital and education, what are differences between these two and why education plays so important role in the regional development and personal life and through which channels it influences the productivity. In this part, I would like to give an insight into externalities of human capital and education as I tried to fit these effects only to one paragraph and I do believe that education itself has many externalities that play a very important role in the development of the region and its inhabitants as individuals. I dedicated thorough description to the perception of education and its effects in our history, so the reader can get an idea of how the perception of today has been created. This perception made a long way since it had started with Adam Smith who came up with the idea of education influencing economic growth in his book Wealth of Nations (Smith, 2000). And since then, there are many interesting ideas that can be seen in many modifications to the production function. Another part describes the differences between the two main educational systems that differ mainly by funding. I discussed the differences there and gave an example of why it matters. But, even here I can't state which of the system brings more benefits into the economy as much deeper insight would be necessary. I think it would be very interesting to investigate which types of schooling bring more benefits to the economy. Probably the most intensive part comes after this because I do think that understanding of manufacturing jobs is essential for ongoing research. It shows the importance of the research in this field because of the essentiality of education for the region. I gave an example of the decline in Detroit because I believe (and my knowledge comes mainly from Edward Glaeser's work) that the decline in Detroit began because innovation has disappeared. And it is impossible to innovate without knowledge and education. Therefore, I summarized what is happening around the decline in the manufacturing sector, where it is still happening and provided the reader with the reasons and implications of job losses. I ended my theoretical part with a section covering local employment multiplier, this part was inspired by Enrico Moretti's work. In this part, I described what is happening with the region when the new job is created through multiplication effects. When a new worker appears, more money is spread around the region and it allows creations of another job, mostly in the non-tradable sector.

The purpose of my practical part was to quantify the effects of educational attainment on GDP and disposable income and also to cover the effects of the density of population, average temperature and standard deviation of the average monthly temperatures. I have to say that using three different models of regression analysis

(Pooled OLS, Fixed effects and Random effects models) were very useful as it brings more than just one point of view on the researched topic.

Firstly, I examined the relationship between education and disposable income. I found out the very significant and strong positive correlation between these two variables. However, an interesting finding is that the effects of education on disposable income were just about half as strong as it should be in the time period from 2007 to 2017. This result comes from differences between Fixed or Random effects models which includes the element of time into the regression and Pooled effects model which do not work with time in the regression analysis. Therefore, there is probably another important variable that plays an important role in this topic and which caused these lower effects during the examined time period. It could also be caused by the state of European economies after the financial crisis in 2008. How countries have been recovering from it than for example the output of manufacturing could help them to recover faster which lowered the effect of education. However, countries with higher educational attainment have still grown faster than others. But, as I have already mentioned with the description of this model, we can't state in which direction the causality works. Even though I proved the correlation between educational attainment and disposable income, I can only estimate that the causality works in both directions. I did also found out that there is a strong and significant correlation between the temperature of the region and educational attainment. The regions with higher temperature and a higher standard deviation of the temperature tend to have smaller educational attainment which shows that even geographic characteristics of the regions are important for the educational attainment and even for the disposable income. I am satisfied with the results of the regression analysis of this model as I got highly significant results and correlation between education and income has been proven.

Secondly, I examined the relationship between educational attainment and GDP per capita. From Pooled OLS model I got very nice and significant results again. Which confirmed the correlation between education and GDP. Omitting the time element, the correlation is very strong: when educational attainment increases by 1%, then GDP increases by 2,5% (GDP per capita is measured as the percentage of EU average). But, the results of fixed effects and random effects models were not that promising at all. Significance of these models is much lower and even strength of effects dropped. Henceforth I estimate that the effects of the post-recession era appeared here much stronger than in the example of the relationship between income and education. However, a positive correlation between education and GDP was still proved. But because of unknown effects, the relationship was much weaker than estimated in the examined period.

Last part of my practical part discusses the return of education for individuals. I have already proved a positive correlation between education and income on the regional level, but my intention here was to investigate the effects of education for individuals. I was successful with my goal to find the link between education and future earnings. I proved in that part, that higher education the individual person has, the higher earnings he can expect. I also discussed and compared the pay-gaps in selected OECD countries and found out that there is a big difference between the biggest and smallest pay-gaps across OECD countries. For example, leading Chile has the pay-gap between the average salary of the master's degree holder and the average salary of a worker with secondary or lower education 403 %. On the other side of the selected is Estonia with the pay-gap only 51 %. But we can find a bigger pay-gaps also in other countries in Europe. For example, in Hungary where is pay-gap equal to 158 %. This fact affirms, that there is low labour mobility across Europe. To get some further results I made a simple linear regression and found the link between the share of employees with tertiary education and the difference between salaries of tertiary and secondary education workers. The higher share of employees with tertiary education the smaller is the pay-gap. Additionally, I was able to state in which countries there is enough of the educated labour force and which countries would need more of it. The link between education and future earning proved even the share of adults earning more than the median across the OECD countries. Not only education level matters, but it also matters what field of study are people interested in. The biggest return of education we can find in fields like engineering, computer science and math.

I was successful in fulfilling the purpose of my bachelor's thesis. I confirmed the hypothesis that the higher share of people with tertiary education positively affects the regional economy. More than that, I found out which fields of study have a stronger effect on future earnings and even on the local economy through local multiplier effects. However, I was nicely surprised by the outcome of the regression analysis and the whole practical part itself. The results are very interesting and come with other questions that would be interesting to solve if I would have more resources. It definitely overcame my expectations. I would be pleased to continue with this kind of research in the future.

Even though, as regards to this topic, there are still more than enough questions to deal with beyond the bachelor thesis assignment. For example, I would like to discuss more side effects of education on the local economy and people living there. Or, research some longer time period effects. Which I was not able to examine because of the accessible data for NUTS II regions in Europe. I really enjoyed studying this topic and would be more than pleased to study this issue any further in future.

Table of Figures

Table of Tables

Bibliography

Allison, P.D., 2009. Fixed Effects Regression Models. SAGE Publications.

Bureau, U.S.C., 2019. American FactFinder - Results [WWW Document]. URL https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk (accessed 5.15.19).

Dell, M., Jones, B.F., Olken, B.A., 2012. Temperature Shocks and Economic Growth: Evidence from the Last Half Century. American Economic Journal: Macroeconomics 4, 66–95. https://doi.org/10.1257/mac.4.3.66

Demeulemeester, J.L., Diebolt, C., 2011. Education and Growth: What Links for Which Policy. <p>Historical Social Research Vol. 36 No. 4, Volumes per year: 1</p>-. https://doi.org/10.12759/hsr.36.2011.4.323-346

Econometrics Academy, n.d. Panel Data Models - Econometrics Academy [WWW Document]. URL https://sites.google.com/site/econometricsacademy/econometrics-models/panel-data-models (accessed 8.20.19).

Eurostat, 2017. Which sector is the main employer in the EU Member States? [WWW Document]. URL https://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20171024-1 (accessed 8.24.19).

Eurostat, n.d. Eurostat - Data Explorer [WWW Document]. Population aged 25-64 by educational attainment level, sex and NUTS 2 regions (%). URL http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=edat_lfse_04&lang=en (accessed 7.28.19a).

Eurostat, n.d. Eurostat - Data Explorer [WWW Document]. Gross domestic product (GDP) at current market prices by NUTS 2 regions. URL http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10r_2gdp&lang=en (accessed 7.28.19b).

Eurostat, n.d. Eurostat - Data Explorer [WWW Document]. Population density by NUTS 3 region. URL http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_r_d3dens&lang=en (accessed 7.29.19c).

Findlay, R., Kierzkowski, H., 1983. International Trade and Human Capital: A Simple General Equilibrium Model. Journal of Political Economy 91, 957–978. https://doi.org/10.1086/261195

Foley, B., 2018. What is Regression Analysis and Why Should I Use It? | SurveyGizmo Blog [WWW Document]. SurveyGizmo. URL https://www.surveygizmo.com/resources/blog/regression-analysis/ (accessed 7.31.19).

Gallo, A., 2015. A Refresher on Regression Analysis. Harvard Business Review.

Gennaioli, N., La Porta, R., Lopez-de-Silanes, F., Shleifer, A., 2013. HUMAN CAPITAL AND REGIONAL DEVELOPMENT. The Quarterly Journal of Economics 128, 105–164.

Gispen, K., 2002. New Profession, Old Order: Engineers and German Society, 1815-1914. Cambridge University Press.

Glaeser on Cities, 2013.

Gleeson, P., 2018. Comparison of Level of Education to Salary [WWW Document]. URL https://work.chron.com/comparison-level-education-salary-6533.html (accessed 4.9.19).

Goleman, D., 1995. Emotional intelligence. Bantam Books, New York.

Hanushek, E.A., 2013. Economic growth in developing countries: The role of human capital. Economics of Education Review 37, 204–212. https://doi.org/10.1016/j.econedurev.2013.04.005
Hjalmarsson, R., Lochner, L., 2012. THE IMPACT OF EDUCATION ON CRIME: INTERNATIONAL EVIDENCE. Research Report 7.
Houseman, S., 2019. Understanding the Decline in Manufacturing Employment | Research Highlights | Upjohn Institute [WWW Document]. URL https://www.upjohn.org/research-highlights/understanding-decline-manufacturing-employment (accessed 5.9.19).
Houseman, S.N., 2018. Understanding the Decline of U.S. Manufacturing Employment. W.E. Upjohn Institute. https://doi.org/10.17848/wp18-287
HSBC Bank USA, 2018. The Value of Education- The price of success 10.
Hsiao, C. (Ed.), 1999. Analysis of panels and limited dependent variable models: in honour of G.S. Maddala. Cambridge University Press, Cambridge, [U.K.].
Hsiao, C., Pesaran, M.H., 2004. Random coefficient panel data models.
IMFBlog, 2018. The Decline in Manufacturing Jobs: Not Necessarily a Cause for Concern. IMF Blog. URL https://blogs.imf.org/2018/04/09/the-decline-in-manufacturing-jobs-not-necessarily-a-cause-for-concern/ (accessed 4.7.19).
Infoclimat - la météo en temps réel : observations météo en direct, prévisions, archives climatologiques, photos et vidéos... [WWW Document], n.d. URL https://www.infoclimat.fr/ (accessed 7.29.19).
Lochner, L., Moretti, E., 2003. The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports 51.
Lukavec, M., 2017. Why you really need a degree. URL (accessed 4.20.19).
Maverick, J.B., 2019. What is the average annual return for the S&P 500? [WWW Document]. Investopedia. URL https://www.investopedia.com/ask/answers/042415/what-average-annual-return-sp-500.asp (accessed 7.16.19).
Mingat, A., Tan, J.-P., 2012. The full social returns to education : estimates based on countries' economic growth performance. The World Bank.
Mishra, M., 2018. Understanding Panel Data Regression [WWW Document]. Medium. URL https://towardsdatascience.com/understanding-panel-data-regression-c24cd6c5151e (accessed 8.3.19).
Moretti, E., 2012. The New Geography of Jobs. Houghton Mifflin Harcourt.
Moretti, E., 2010. Local Multipliers. American Economic Review 100, 373–377. https://doi.org/10.1257/aer.100.2.373
Moretti, E., 2004. Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions. The American Economic Review 94, 656–690.
OECD, 2018. Education at a Glance 2018 - Data and Methodology - OECD [WWW Document]. OECD. URL http://www.oecd.org/education/education-at-a-glance-2018-data-and-methodology.htm (accessed 4.18.19).
OECD, 2001. The Well-being of Nations: The Role of Human and Social Capital. OECD. https://doi.org/10.1787/9789264189515-en
PayScale, 2018. Best Value Colleges [WWW Document]. PayScale. URL https://www.payscale.com/college-roi (accessed 4.20.19).
Peter H. Lindert, 2006. Growing Public: Social Spending and Economic Growth since the Eighteenth Century. 2 Volumes. Cambridge University Press, Cambridge, 2004. J Econ Inequal 4, 251–252. https://doi.org/10.1007/s10888-006-9027-7

Peterson’s, 2017. Private School vs. State School: Which Would Work Best for You? Peterson’s. URL https://www.petersons.com/blog/private-school-vs-state-school-which-would-work-best-for-you/ (accessed 7.17.19).
Ramsey, F., Schafer, D., 2012. The Statistical Sleuth: A Course in Methods of Data Analysis. Cengage Learning.
Romer, P.M., 1986. Increasing Returns and Long-Run Growth. Journal of Political Economy 94, 1002–1037.
Salgür, S.A., 2013. The Importance of Education in Economic Growth. Euromentor Journal; Bucharest 4, 50–57.
shanghairanking.com, 2018. ARWU World University Rankings 2018 | Academic Ranking of World Universities 2018 | Top 500 universities | Shanghai Ranking - 2018 [WWW Document]. URL http://www.shanghairanking.com/ARWU2018.html (accessed 8.24.19).
Smith, A., 1723-1790, 2000. The wealth of nations / Adam Smith ; introduction by Robert Reich ; edited, with notes, marginal summary, and enlarged index by Edwin Cannan. New York : Modern Library, 2000.
Solow, R.M., 1957. Technical Change and the Aggregate Production Function. The Review of Economics and Statistics 39, 312–320. https://doi.org/10.2307/1926047
Studyportals B.V., 2019. Tuition Fees at Universities in Europe in 2019 - Overview and Comparison - MastersPortal.com [WWW Document]. URL https://www.mastersportal.com/articles/405/tuition-fees-at-universities-in-europe-in-2019-overview-and-comparison.html (accessed 7.17.19).
Tamborini, C.R., Kim, C., Sakamoto, A., 2015. Education and Lifetime Earnings in the United States. Demography 52, 1383–1407. https://doi.org/10.1007/s13524-015-0407-0
The College Board, 2019. Average Published Undergraduate Charges by Sector and by Carnegie Classification, 2018-19 - Trends in Higher Education [WWW Document]. URL https://trends.collegeboard.org/college-pricing/figures-tables/average-published-undergraduate-charges-sector-2018-19 (accessed 4.9.19).
The Economist, 2015. It depends what you study, not where. The Economist.
The Economist, 2005. Industrial metamorphosis - Manufacturing employment. The Economist.
Torres-Reyna, O., 2007. Panel data analysis fixed and random effects using Stata (v. 4.2). Data & Statistical Services, Priceton University.
Travel Weather Averages (Weatherbase) [WWW Document], n.d. . Weatherbase. URL http://www.weatherbase.com/ (accessed 7.29.19).
Woldometers, 2019. U.S. Population (2019) - Worldometers [WWW Document]. URL http://www.worldometers.info/world-population/us-population/ (accessed 4.9.19).
worldpopulationreview.com, 2019. Detroit, Michigan Population 2019 (Demographics, Maps, Graphs) [WWW Document]. URL http://worldpopulationreview.com/us-cities/detroit-population/ (accessed 4.14.19).