Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Overeducation in the Czech Labour Market

Author: **Ľubica Laslopová**

Supervisor: **Mgr. Barbara Pertold-Gebicka M.A., Ph.D.**

Academic Year: **2014/2015**

Declaration of Authorship

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

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

Prague, May 14, 2015

Signature

Acknowledgments

I am especially grateful to my supervisor, Mgr. Barbara Pertold-Gebicka, M.A., Ph.D. for her useful advice and for her willingness to read all my long e-mails and an even longer thesis.

Also, special thanks are addressed to Zuzana Laslopová M.A. for correcting my English and to Matej Marčišin for his patience.

Bibliographic record

Laslopová, Ľ., 2014. *Overeducation in the Czech Labour Market.* Bachelor thesis. Charles University in Prague.

Character count: 88 665

Abstract

In one way or another, labour market theories link additional years of education to better jobs and higher earnings. The empirical research, however, identified systematic distortions of this rule. Some workers are considered to be overeducated, as their education exceeds the level required for their job. This thesis deals with overeducation in the Czech labour market, as there are still few studies considering this topic. First we estimate incidence of overeducation using objective measures of overeducation. Then we perform an analysis of its determinants and possible consequences, namely pay penalty and lower job satisfaction. The results are to a large extent dependent on the measurement method. Despite these differences, we found that certain factors, such as general health, place of residence, ethnicity or managerial positions may influence the probability of mismatch. We provide some evidence that the returns to "surplus" education might be systematically lower than returns to required education. Also, overeducation can be linked to a lower job satisfaction.

JEL Classification	J82, J31, J28, J24
Keywords	incidence of overeducation, determinants of overeducation, consequences of overeducation, Czech labour market, education in the Czech Republic

Author's email	llaslopova@gmail.com
Supervisor's email	barbara.pertold-gebicka@fsv.cuni.cz

Abstrakt

Teórie trhu práce spájajú vyššie vzdelanie s lepším uplatnením na trhu práce a vyšším finančným ohodnotením. Empirické výskumy však poukazujú na to, že tento princíp často neplatí. Prevzdelanosť predstavuje situáciu, v ktorej vzdelanie pracovníka presahuje požadovanú úroveň na danej pracovnej pozícii. Vzhľadom na stále nízke množstvo štúdií venujúcich sa tejto problematike, táto práca rieši prevzdelanosť na českom pracovnom trhu. V prvom rade odhadujeme jej výskyt použitím takzvaných objektívnych metód merania prevzdelanosti. Následne sa venujeme jej determinantom a taktiež možným následkom. Výsledky do veľkej miery závisia na zvolenej metóde merania prevzdelanosti. Napriek tomu sa dá z našej analýzy usúdiť, že pravdepodobnosť prevzdelanosti môže byť ovplyvnená viacerými faktormi, okrem iného celkovým zdravím, bydliskom, etnicitou či faktom, že pracovník je na manažérskej pozícii. Môžeme pozorovať, že návratnosť investície do dodatočného vzdelania je výrazne nižšia po presiahnutí úrovne požadovanej pre výkon zamestnania. Prevzdelanosť taktiež môže prispievať k nižšej spokojnosti s prácou.

Klasifikácia JEL	J82, J31, J28, J24
Kľúčové slová	výskyt prevzdelanosti, determinanty prevzdelanosti, následky prevzdelanosti, Český pracovný trh, vzdelanie v Českej Republike

E-mail autora	`llaslopova@gmail.com`
E-mail vedúceho práce	`barbara.pertold-gebicka@fsv.cuni.cz`

Contents

List of Tables **ix**

List of Figures **xi**

Acronyms **xii**

Thesis Proposal **xiii**

1 Introduction **1**

2 Overeducation Literature **3**

2.1 Labour Market Theories and Overeducation 3

2.2 Results of the Past Studies 5

2.2.1 Extent of Overeducation 5

2.2.2 Determinants of Overeducation 7

2.2.3 Consequences of Overeducation 9

3 Focus on the Czech Republic **11**

3.1 Czech Educational System 11

3.2 Czech Labour Market 15

3.3 Previous Evidence of Overeducation in the Czech Republic . . . 20

3.4 Hypothesis Statements 21

4 Data and Methodology **22**

4.1 Description of the Data 22

4.2 Selected Measures of Overeducation 23

4.3 Determinants of Overeducation: Multinomial Logit 26

4.4 Wage effects of Overeducation: Mincer Model Extensions 28

4.5 Overeducation and Job Satisfaction: Linear Regression 32

5 Results **34**
5.1 Incidence of Overeducation 34
5.2 Determinants of Overeducation 35
5.3 Wage Effects of Overeducation 40
5.4 Overeducation and Job Satisfaction 43

6 Conclusion **46**

Bibliography **53**

A Imputation of the Variable Years of Completed Education **I**

B Small and Hsiao Tests for the Independence of Irrelevant Alternatives **III**

C Cross-correlation Tables **V**

List of Tables

3.1 Assignment of Workers to the Occupations, by Education Level 17
3.2 Relative Wages by Educational Attainment *Secondary Education = 100%* 18

4.1 Descriptive Statistics of the Variable Years of Education by ISCO Major Groups 24
4.2 Normative Measure: Skill Requirements by 1-digit occupational categories 25
4.3 Normative Measure: Skill Levels corresponding to Education Levels 26

5.1 Incidence of Overeducation and Undereducation 35
5.2 Multinomial logit: Estimated Coefficients of the Outcome "Overeducated" and their Standard Errors 37
5.3 Average Marginal Effects after Multinomial Logit, for the Outcome "Overeducated" 39
5.4 Wage Effects of Overeducation: Verdugo and Verdugo Specification 41
5.5 Wage Effects of Overeducation: ORU Specification 42
5.6 Overeducation and Job satisfaction: OLS Coefficient Estimates and Robust Standard Errors 44

C.1 Cross-correlation Table for Verdugo-Verdugo Model, Statistical Measure V
C.2 Cross-correlation Table for Verdugo-Verdugo Model, Normative Measure VI
C.3 Cross-correlation Table for ORU Model VII
C.4 Cross-correlation Table for Job Satisfaction Model, Statistical Measure VIII

C.5 Cross-correlation table for job satisfaction model, normative measure IX

List of Figures

3.1 Educational Attainment of the Population older than 15, *in thousands* 13
3.2 Change in Number of Graduates by Field of Study 14

B.1 Small- Hsiao Test, Specification 1 III
B.2 Small-Hsiao Test, Specification 2 III
B.3 Small-Hsiao Test, Specification 3 IV
B.4 Small-Hsiao Test, Specification 4 IV

Acronyms

CZSO Czech Statistical Office

MEYS Ministry of Education, Youth and Sports

MoLSA Ministry of Labour and Social Affairs

OECD Organisation for Economic Cooperaction and Devvelopment

UNESCO United Nations Educational, Scientific and Cultural Organization

EU European Union

EU-SILC European Union Statistics on Income and Living Conditions

ISCED International Standard Classification of Education

ISCO International Standard Classification of Occupations

CZK Czech Koruna

IIA Independence of irrelevant alternatives

CLM Classical Linear Model

OLS Ordinary Least Squares

Bachelor Thesis Proposal

Author	Ľubica Laslopová
Supervisor	Mgr. Barbara Pertold-Gebicka M.A., Ph.D.
Proposed topic	Overeducation in the Czech Labour Market

Topic characteristics Overeducation represents a mismatch between the qualification necessary for a particular job and qualification actually acquired by an individual working on the position. This phenomenon demonstrates that returns to additional years of education can be questionable. Since the introduction of this concept by Richard Freeman in 1976, many researchers attempted to set diverse empirical measurements to the theoretical frameworks. As there is no unified theory, these researches lead to different outcomes.

In this thesis, we will try to evaluate the extent to which overeducation represents a real problem in Czech labor market. After a brief summary of knowledge about this phenomenon, empirical analysis will follow. Mainly we will focus on measurement of the extent of overeducation, then we will estimate its consequences in form of costs, using the data from the European Union Statistics on Income and Living Conditions (EU-SILC).

Core bibliography

1. McGuiness, S. (2006): "Overeducation in the Labour Market." *Journal of Economic Surveys* **20(3)**: pp. 387–417.
2. Tsang, M.C. & A. Levin (1995): "The Economics of Overeducation." *Economics of Education Review* **4(2)**: pp. 93–104.
3. Chevalier, A.& H.M. Levin (2003): "Measuring over-education". *Economica* **70(279)**: pp. 509–531.
4. Dolton, P. & A. Vignoles(2000): " The incidence and effects of overeducation in the U.K. graduate labour market." *Economics of Education* **19(2)**: pp. 179–198.

5. Battu, H. & C. Belfield& P. Sloane (2000): "How well can we measure graduate overeducation and its effects?." *National Institute Economic Review* **171(1)**: pp. 82–93.

6. Barone, C. L. Ortiz (2011): "Overeducation among European University Graduates: a comparative analysis of its incidence and the importance of higher education differentiation." *Economics of Education Review* **61(3)**: pp. 325–337.

Author

Supervisor

Chapter 1

Introduction

"You can never be overdressed
or overeducated"

Oscar Wilde

Successful entry to the job market has always been one of the motivations to study. It is commonly believed that more educated individuals earn higher wages and are less likely to be unemployed (Card 1999). Therefore it is not surprising that in many economically developed countries the educational level of the population has increased considerably over the past few decades. Czech Republic is no exception, the same trend has been detected (OECD 2014a). With higher share of educated workers in the labour market, there is a concern that returns to additional education become questionable. Evidence has been found that more years of schooling do not ensure access to more challenging and better paid jobs. Overeducation represents a mismatch between qualification necessary for a particular job and qualification actually acquired by an individual working on this position (McGuiness 2006). Alternatively, it can be defined as unfulfilled expectations of the educated concerning their career attainments. Finally, the term overeducation may denote a decline in economic position of qualified individuals relative to its historical level (Tsang & Levin 1985). This phenomenon can be compared to the inflation of educational attainment, as the increasing number of credentials can lead to a decrease in their relative economic value. For the purposes of our analysis, we mainly refer to the first of these three definitions.

As there is no unified theory, many different explanations for the overeducation exist. Although the theoretical debate is unlikely to lead to a consensus

any time soon, the econometric analysis allows us to interpret the evidence and consequently, contributes to a better understanding of the issue.

This thesis aims to explore current situation in the Czech labour market concerning overeducation. Many studies have already described this educational mismatch from different perspectives. With respect to the specifics of the Czech labour market and educational system, we build on these studies. More specifically, we attempt to quantify the extent of overeducation using different methods of measurement, and to identify the determinants and consequences of the educational mismatch. The objective of this analysis is to discuss the importance of the overeducation problem in the Czech Republic.

The thesis is organized the following way. After an overview of existing literature on overeducation in Chapter 2, where we briefly summarize previous knowledge about the phenomenon in general, a discussion of the Czech environment in Chapter 3 follows. In Chapter 4, we describe the data and methodology used for our calculations. In Chapter 5 we present our results. Firstly, it is the estimated incidence of both overeducation and undereducation. Secondly, we report the estimated influence of diverse factors on the probability of mismatch between education and employment requirements. Finally we focus on the consequences of overeducation: the effect on wages and less discussed link to the worker's satisfaction with the job. In Chapter 6, we conclude on our findings and discuss main ideas.

Chapter 2

Overeducation Literature

The term overeducation was introduced by Freeman (1976) in his paper " The overeducated American". It was written in reaction to a post-war expansion of higher education in American population. Freeman warned about the danger of decrease in rate of return to investment in academic degrees due to the possible oversupply of graduates (Dolton & Vignoles 2000). Since then, educational mismatches have been discussed and studied to a large extent. There are several important questions to deal with while analyzing this phenomenon. Is the existence of overeducation consistent with theories of the labour market? What is the extent of overeducation and how is it changing in time? What are the consequences and determinants of overeducation? We provide a brief overview of the debate concerning overeducation.

2.1 Labour Market Theories and Overeducation

To set the economic context of our analysis, we hereby discuss different theoretical models of the job market. Each perspective provides another explanation for the overeducation, its origins and dynamics.

Firstly, we can look at the job market from neoclassical point of view, by means of the Human Capital Theory. The idea has been presented in the influential work "Human Capital", written by Gary S. Becker in 1964. Becker considers the education as an investment to the human capital, which consequently raises marginal product of the worker. The key assumption is the flexibility of firms and workers, in terms of their ability to adapt to the changes in labour market. There are more possible correction mechanisms. Excess supply of graduates can lead to a relative decline in their wages. Individuals then

tend to reconsider their investment in education and they lower their expectations (Becker 1964) (Tsang & Levin 1985). Alternatively, overeducation can be corrected through a job change or an internal promotion (Barone & Ortiz 2011). In the long run, workers are paid their marginal product influenced by their human capital, and their skills are fully utilized in their occupation.

Therefore any mismatch between required and acquired human capital is considered temporary, or as Barone & Ortiz (2011) point out, "negligible, or even deceitful". This naturally brought several researchers to a conclusion that overeducation can be explained by Human Capital Theory only in the short-run. However, as McGuiness (2006) suggests, the imperfections of empirical measurements such as neglecting less formal forms of human capital can imply that even persistent overeducation could be consistent with this framework. In this case, the overeducation would compensate for the lack of other skills, which are unobserved.

Different perspective is provided by the Job Competition Model, based on Lester Thurrow's work "Generating Inequality"(Thurrow 1975). According to the model, a worker's productivity cannot be directly assessed from his education level. However, the education serves as an indicator of the amount of further training needed (Barone & Ortiz 2011). Individuals applying for a job form a "queue", and their relative position in this queue is determined by several factors including education (Tsang & Levin 1985). This could easily provide an explanation for the excessive investment in qualification, as the individuals are trying to improve their position in this job competition. Also, the firms are not so flexible in adjusting production and wages to current situation on the job market. From this perspective, job characteristics may be regarded as the only determinant of wages, education only influences access to the job. Similar perspective is provided by Job market signaling theory formulated by Spence (1973). Because of the fact that the job market has to deal with imperfect information, education can serve as a signal from applicants to employers about their ability level and motivation.

The Assignment models represent a compromise between the Human capital theory and the Thurrow's model. These models are suppose numerous factors that have effect on the way how the workers are assigned to jobs. Moreover, wage is not a directly observable function, but it is regarded as one of the outcomes of equilibrium solution. It is due to the fact that the marginal product of labour (and consequently wage) is dependent on the combination of worker's characteristics and the nature of the job. This framework appears

to be consistent with the overeducation phenomenon, as the education is only one of the many factors that determine which individuasl get which jobs and how much are they paid (McGuiness 2006).

Last but not least, according to the Theory of career mobility presented by Sicherman & Galor (1990), returns to education can be in form of "higher probabilities of occupational upgrading" so the individuals may accept the jobs with lower direct returns to education, but higher chance to upgrade. This theory believes that the overeducation is a rational outcome for both sides of the labour market: employers and employees. In this framework, the educational mismatch is mainly a short-run phenomenon. This theory is very attractive for the researchers, as it seems to be in accordance with empirical results (Buchel & Mertens 2010).

These are the theories that provide powerful tools for an empirical research of educational mismatches. All of these models offer different explanations to the allocation of labour force on the labour market. But what they do not consider is a discussion of the role of educational institutions. In contrast with the economic point of view, sociological research of overeducation discusses the importance of these institutions. As suggested by Barone and Ortiz (2010), their specifics can influence the economic value of education considerably. More precisely, the supply of highly educated workers is linked to the selectivity and stratification of the institutions providing higher education. High selectivity may to some extent prevent overeducation by keeping the number of academic degrees low, which helps to preserve their value. Moreover, highly stratified educational systems may contribute to a better match between the supply and demand of the human capital (Barone & Ortiz 2011).

2.2 Results of the Past Studies

2.2.1 Extent of Overeducation

Numerous studies explore the extent of overeducation. When comparing the results of these studies, one must keep in mind that measurement methods may be different, which naturally affects the results. In general, four different measures of overeducation are used. There are two "subjective" measures based on self-assessment of workers. Workers can be either asked about minimum required education in their position which is then compared to their acquired education, or they can be directly asked if they consider themselves to be

overeducated or undereducated (McGuiness 2006). The advantage of these methods of measurement is that they do not have to deal with any aggregates, which increases their accuracy. However, the reliability of self-assessment can be questionable. For instance, as Hartog (2000) points out, requirements of a job can be overstated by the workers.

On the other hand, "objective" measures do not rely on worker's judgment. Instead, the information about necessary education can be derived from realized matches. If worker's acquired education differs from the mean or median educational level of his or her occupation by more than one standard deviation, he or she is qualified as overeducated or undereducated. This approach is called "statistical". Alternatively, level of education required for different occupations can be determined by job analysts and then compared to individual's education. This method is also called "normative" or "dictionary-based" (Hartog 2000) (Dolton & Vignoles 2000). We provide a more detailed discussion of objective methods in Chapter 4, as we will utilize them in our analysis.

The question thus stands: Do these different measures lead to significantly different outcomes? McGuiness (2006) has come to a conclusion:

> "(...) on balance, and despite concerns relating to poor correlations between the various approaches, there is no consistent evidence to suggest that any of the subjective or occupational dictionary-based measurement frameworks result in a systematic and significant underestimate of either the incidence, or wage effects, associated with overeducation."

But what about statistical method? Meta-analysis elaborated by Groot & van den Brink (2000) demonstrates that only this measure "based on within occupation variation in years of education appears to yield to lower estimates of overeducation than studies that use another definition". However, this does not necessarily mean that statistical measure underestimates the real extent of overeducation. In fact, all of these widely used measures implicitly assume identically educated workers to be equally skilled. This assumption is at least questionable. Consecutive failure to control for this kind of heterogeneity may in fact overestimate real values of incidence and consequences of overeducation. For this reason, a different approach was applied by Chevalier (2003). His model distinguished two types of workers by their skill level which lead to a separation of individuals into two different groups: good graduates and underachievers. When this heterogeneity has been taken into account, the individuals traditionally defined as overeducated were classified as either "genuinely" overeducated or "apparently" overeducated.

Being aware of the existence of different measurement methods, we can now summarize the evidence from previous studies. Apart from the United States, where the problem of overschooling has been recognized for the first time, researches confirmed its existence in several European countries, in Canada, Australia, Latin America, and some Asian countries, too. According to Groot's and Maasen van den Brink's (2000) meta-analysis, the mean estimated extent of overeducation among studies in United States is higher than the mean value for European countries. This conclusion is in accordance with the findings of Leuven and Oosterbeek (2011), who made synthesis of the studies from all around the world dating from 1970s to 2000s. The result of comparing the average measured extent of overeducation by continents was that the greatest estimated fraction of overeducated workers was in United States, followed by Europe. As for the differences between European countries, we can mention the comparative study of Barone & Ortiz (2011), who explored the educational mismatch in Spain, Italy, Austria, Finland, Norway, Netherlands, Germany and the Czech Republic. Among all of these countries, only Spain had considerably higher number of overeducated workers. All of the other countries had less than 10%.

Given the increasing number of graduates, another question arises: Is the incidence of overeducation increasing in time? McGuiness (2006) claims that there are no indications that the extent of overeducation has become more important since the time the economists started to quantify it.

2.2.2 Determinants of Overeducation

An important share of overeducation literature deals with the determinants of overeducation. What are thus the characteristics that may have impact on the probability of being overschooled?

The first factor is gender. Gender inequality in the labour market is a well-known fact, but researches in this area usually focus on the male-female earnings differentials. Frank (1978) suggests that this inequality is a more complicated problem. He claims that women are also more likely to be overeducated, as their job-search is restricted by the choices of their partners who probably contribute higher share of household's income (Leuven & Oosterbeek 2011) (Mysíková 2014).

This brings us to the next determinant. Probability of mismatch may also be affected by spatial factors, as a job-searching process is usually geographically

limited. In the context of overeducation, spatial factors were firstly discussed by Frank (1978) who examined these limitations for married women. Later on, the research was extended by McGoldrick & J.Robst (1996) and Buchel & van Ham (2003).

Another characteristic believed to be linked with overeducation is ethnicity. Study of OECD countries in 2007 demonstrated that in all countries in question except New Zealand, immigrants are more likely to be overeducated (OECD 2007). There are few other studies dealing with immigrant overeducation, Leuven & Oosterbeek (2011) attribute it to the fact that a comparison of educational attainment in different countries is still problematic.

It is worth mentioning that overeducation is commonly believed to decrease with the age of worker. Majority of workers change their working position during their lives (Sicherman & Galor 1990). As we already mentioned, the career mobility theory believes in progressive upgrade in career of overeducated workers. Also, Leuven & Oosterbeek (2011) suggest that with the on-the-job training, workers earn additional human capital that helps them to proceed in their career.

Besides personal characteristics, there are other factors that may influence the risk of overeducation. For instance, Karakaya *et al.* (2007) explored work-related factors including size of establishment, form of economic control over establishment and type of contract. Their research provides evidence that fully state-owned firms and fixed-term contract can reduce probability of overeducation. They did not find any significant relation between firm size and overeducation, but an earlier analysis of Battu *et al.* (1999) leads to a conclusion that working in a large firm can increase the risk of overeducation.

The other interesting conclusion of Battu *et al.* (1999) is that overeducation significantly varies across different fields of study. Their conclusion, consistent with results of Ortiz & Kucel (2008) is that certain disciplines such as medicine, mathematics, engineering or law can to some extent prevent overeducation and increase the probability of proper match. On the other hand, the risk of overeducation is higher for humanistic disciplines. Labour market theories provide different explanations of this phenomenon. Human capital theory would assume more "productive" skills relevant to the job market for graduates from these disciplines. Job competition model and Job market signalling theory lead to a conclusion that given the higher selectivity of certain disciplines, different credentials can signal different ability level (Barone & Ortiz 2011).

Relatively few researchers explored macroeconomic determinants of overed-

ucation. Their findings are not surprising. Liu *et al.* (2012) identified a countercyclical trend of overeducation based on Norwegian panel data. Following the intuitive reasoning, they link it to the unemployment:

> "(...) a typical recession, with a rise in unemployment rate by three percentage points, implies an initial increase of about 30% in the probability of mismatch and a 9% downgrading in the average quality of their matches. The effects of initial labor market conditions on mismatch decline over time but remain highly persistent over early careers, suggesting that some graduates never switch back to the "right" industry."

One possible explanation is that worsened situation in the labour market associated with a higher risk of unemployment forces individuals to accept jobs for which they are overqualified. A more detailed discussion of macroeconomic situation and overeducation is provided by Morano *et al.* (2014)

2.2.3 Consequences of Overeducation

The academic concern for overeducation is based on the knowledge of its possible consequences. First of all, if education is considered as an investment to human capital, an individual considering additional year of schooling weights its costs against present value of an additional year of education. The idea can be expressed by following relation:

$$\sum_{t=1}^{T-s} \frac{W_s - W_{s-1}}{(1+r_s)^t} = W_{s-1} + c_s$$

,where W_s stands for wage resulting from s years of schooling, c_s for cost of s^{th} year of education, r the annual interest rate and wage W_{s-1} represents the opportunity costs of not working during the s^{th} year (McGuiness 2006).

The problem is that future returns to education are in fact uncertain. When the overeducation is present, expectations about future career attainments are not fulfilled.[1] This includes lower earnings than expected. Many studies tried to discover whether the overeducation is connected with a pay penalty. In other words, whether the overeducated workers earn significantly less than equally educated "matched" individuals.

The existence of pay penalty of overeducation has been confirmed by McGuiness (2006) who aggregated the results of 22 different studies that use different

[1] This follows directly from the alternative definition of overeducation formulated by Tsang & Levin (1985) which we stated in the introduction.

models to estimate wage effects of mismatches. As expected, most studies concluded that overeducated individuals earn lower wages than the individuals with the same education but corresponding job requirements. Nevertheless, surplus years of education are in general still connected with positive returns. Therefore if we compare the overeducated with their co-workers whose education matches job requirements, their wages would be higher.

The pay penalty is not the only negative effect of overeducation that has been observed so far. If a worker's skill level exceeds job requirements, the worker tends to be less satisfied with his job. This may be attributed to the boredom linked with less challenging nature of the work (Sánchez-Sánchez & McGuinness 2013). Consequently, worker's productivity is lower, which implies additional costs of overeducation for the firms.(Tsang & Levin 1985) The result of econometric analysis of Fleming & Kler (2014) using Australian data is in accordance with this reasoning. Sánchez-Sánchez & McGuinness (2013) went a little further and used the model which separates overeducation from excess of specific job-related skills. Even after this corrective measure, the overeducation has been found to be associated with lower job satifsfaction.

Chapter 3

Focus on the Czech Republic

3.1 Czech Educational System

The system of educational institutions plays a central role in the process of labour force formation. For this reason we will discuss its design and output. The system of education in the Czech Republic includes preschools, elementary schools, secondary schools, tertiary educational institutions, elementary art schools and language schools authorized to organize state language examination (Act n. 561/2004, §7).

Compulsory education consists of 9 years of schooling. Primary and lower secondary schooling is provided by elementary schools and consists of two stages corresponding to levels 1 and 2 according to the International standard classification of education(ISCED)[1]. In the second stage, pupils can continue to complete the compulsory education at the elementary school, or to alternatively proceed to six or eight-year secondary general schools, or in the corresponding level of the eight-year educational program at a conservatory (MEYS 2012).

In the Czech Republic, there are several types of institutions providing upper secondary education. Firstly, there are four-year programs completed by a school-leaving examination corresponding to ISCED level 3A, namely secondary general schools, lyceums, secondary technical schools and conservatories. Secondly, there are secondary vocational schools completed after 2-3 years by attaining an apprenticeship certificate (ISCED 3C). Also, there are programs which consist of 1-2 years of general and vocational education completed without the certificate or school-leaving examination (ISCED 2C/3C)

[1]ISCED is a statistical framework used to organize information about education levels maintained by UNESCO

(MEYS 2012). This highly differentiated structure allows students to choose whether they will receive general schooling in order to be prepared for further education or specific qualification for the range of occupations.

Higher education institutions of university and non-university type provide tertiary education up to ISCED levels 5A and 6. Czech universities are providing mostly three years bachelor, two years master or engineer programs and following 3-4 years doctoral programs[2]. There are three types of universities: public, private and state (police and military) universities (Act n. 561/2004, §7). Tertiary professional schools and conservatories provide an alternative with 2-3 year education completed by absolutorium (ISCED 5B).

Post-secondary nontertiary education is provided by follow-up courses. These programs offer subsidiary education to those who completed secondary vocational schools with an apprenticeship certificate and want to get more qualification. Alternatively, some programs offer additional education to students with state examination. (Eurydice 2010)

Majority of programs offered by public universities are tuition-free, financed out of taxation. University education is thus relatively easily accessible. Since 1992, there is a "per capita" funding system of educational institutions. Ministry of Education, Youth and Sports sets a level of non-capital expenditure on one pupil/student (Eurydice 2010) (MEYS 2012). This model of allocating resources within educational system has been criticized for the risk of creating negative incentives for educational institutions. Financing based on the number of students could artificially raise the entry rate, which would consequently reduce the quality of education (Čermáková *et al.* 1994). However, the increase in entry rates is regulated and cannot exceed the limit set according to the previous year.

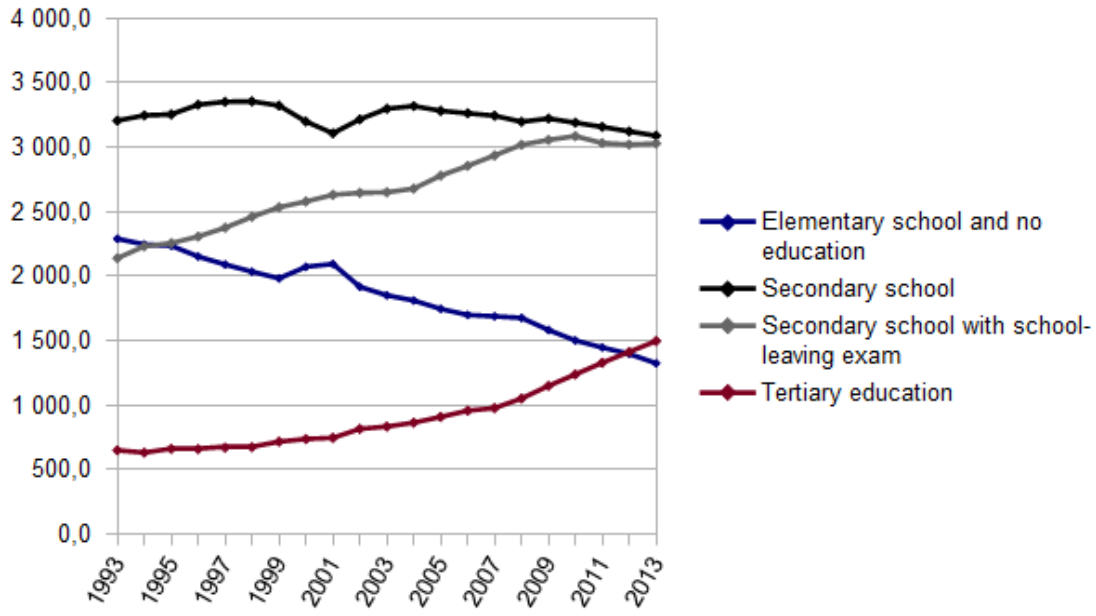
Development of the educational attainment in the Czech Republic follows similar patterns as in other OECD countries. However, in some aspects it is very specific. To some extent, this is a legacy of educational policies in the socialist period, when the admission rates were strictly determined by central planning authorities. During the transition, these limitations were removed for high schools, but legislative obstacles for establishing colleges remained (Filer *et al.* 1999). The result is that even two decades after the beginning of transition, the educational attainment of the Czech population differs from the other OECD countries. In 2012, the percentage of 25-64 year-olds whose

[2]Certain disciplines have different structure, such as medical and pharmaceutical schools or civil engineering

highest level of education was upper-secondary or post-secondary non-tertiary were 73.09%, which is considerably higher than the OECD average 45%. On the contrary, in 2012, only 19.28% of 25-64 year-olds in the Czech Republic were tertiary educated, compared to 30% average of all OECD countries (OECD 2014a).

As we already mentioned, many developed economies have to deal with unprecedent educational expansion. Situation in the Czech Republic is similar, but these changes occured with a delay. Figure 1 presents the ascending trend in educational attainment of the population older than 15 years between 1993 and 2013 in the Czech Republic. We can see that the number of individuals with elementary or no education has decreased, the number of individuals with secondary education was relatively steady, but the number of individuals with higher educational attainment has increased considerably (CZSO 2014). As suggested by (Groot & van den Brink 2000), the shift in educational attainment may be also illustrated by comparing educational attainments of different age groups. In case of Czech Republic, in 2012, the proportion of tertiary educated in population of 55-64 years old was 12.63% compared to 27.83% of 25-34 years old. The difference was slightly higher than the OECD average of 15 percentage points (OECD 2014a).

Figure 3.1: Educational Attainment of the Population older than 15, *in thousands*

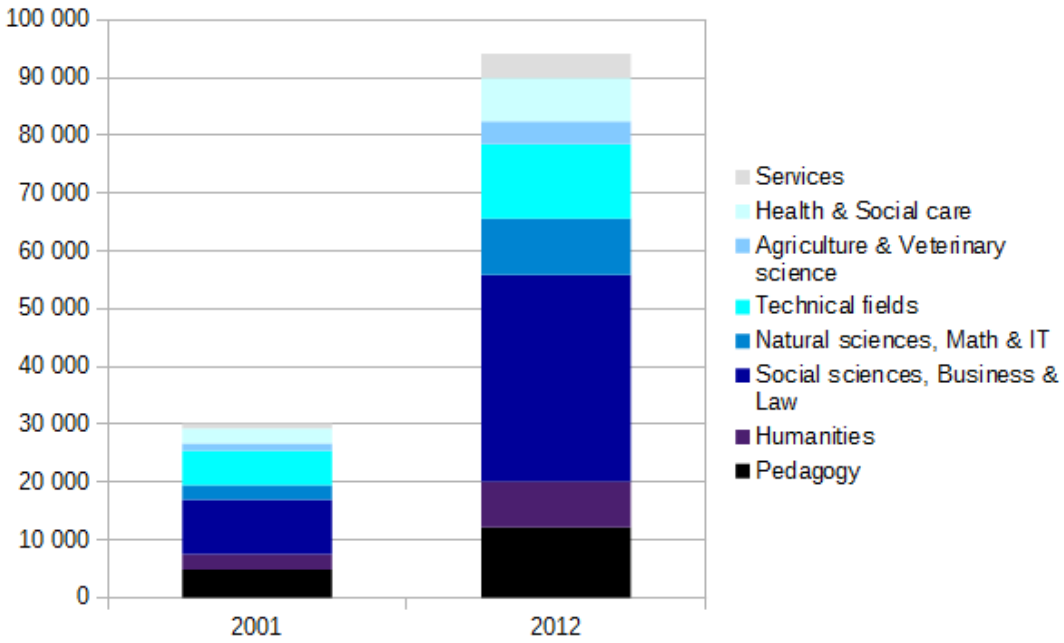


Source: Author based on CZSO

Clearly, even the individuals with the same level of education are not "perfect substitutes" as every discipline provides them with different skills. As the supply of college graduates increases, the question arises: what is the horizontal composition of college graduates? As we can see in Figure 2, in 2012, the highest number of graduates completed their studies in social sciences, business and law, followed by technical fields, pedagogical studies and natural sciences. From 2001 to 2012, there was an extreme increase in the number of graduates in social sciences, business and law, by more than 277%. On the contrary, the smallest, though still considerable relative increase in the number of graduates, by 113%, has been observed in technical fields. The number of graduates from other disciplines has increased as well (CZSO 2012b).

As for the population whose highest level of education is upper secondary, the vast majority of its representatives completed vocational upper secondary education. If we express it as a percentage of all 25-64 year-olds, in 2012, 72.99% of population of this age completed upper secondary vocational school, while only 0.9% of them completed upper secondary or post-secondary non-tertiary school with general orientation (OECD 2014a). This well reflects the fact that the general upper secondary schools aim to prepare for further education.

Figure 3.2: Change in Number of Graduates by Field of Study



Source: Author based on CZSO

In 2010, European Commission proposed the new strategy Europe 2020,

aiming for "smart, sustainable, inclusive growth"[3]. In accordance with this European strategy, Czech Republic implements The Educational Policy Strategy of the Czech Republic for 2020. Among other things, its targets include reducing the early school-leaving drop-out rate and increasing the number of tertiary educated in population of 30-34 year-olds to 32% at minimum (MEYS 2014).

The question thus stands whether the labour market is able to absorb realised and planned output of the educational system. There is a concern about possible oversupply of college graduates which may produce inefficiencies. This may be reflected in both social and private returns to education as well as in the number of college graduates with "non-college" jobs. So far, analysis of the relationship between the number of college graduates and the fraction of them working in non-college occupations has not found any positive correlation (Gebicka 2010). But one must keep in mind that this does not necessarily mean that there is genuinely no such relation.

3.2 Czech Labour Market

As a post-communist Central European country, Czech Republic is very specific when it comes to the evolution of labour market. Before the fall of the iron curtain, Czechoslovak labour market was highly regulated and determined by the full employment and low productivity of labour. But after 1989, inevitable changes associated with the transition to a market economy took place. A shift from the excess in demand for labour to the excess in labour supply, as well as the opening of the labour market were necessary. The increase in unemployment has been compensated by simultaneous increase in labour productivity. Growing productivity can be attributed to the technological lag at the beginning of the new era, and the fact that since then, the situation in Western and Central European countries started to equalize (Kotýnková 2006).

During the transition process, sectoral structure of the employed population changed substantially. In accordance with the trend in EU-27, there was a decline in primary sector[4] and in secondary sector[5] in favour of tertiary sector. In 2010, 4.5% of all employed worked in primary sector, which is approximately

[3] See "Europe 2020: Commission proposes new economic strategy", European Commission. Retrieved 5 March 2010.

[4] Primary sector comprise agriculture, fishery and forestry

[5] Secondary sector comprise manufacturing and construction

by 4 percentage points less than in 1995. Also, 34% of working population was involved in secondary sector, which is by 5 percentage points less than in 1995, but relatively high compared to the other EU countries (Lepič & Koucký 2012). In general, the more developed the country is, the lower fraction of labour force is working in primary and secondary sector, and the higher fraction is working in tertiary sector[6] (Kotýnková 2006).

To illustrate current situation in the Czech labour market, we proceed with a brief overview of the key labour market indicators including activity, employment and unemployment. In the third quarter of 2014, labour force participation rate in the Czech Republic was 59.4%. Among the population aged 15 to 64 years, 73.7% were working or actively searching for a job. Moreover, the employment rate in the third quarter of 2014 reached 55.9 % which corresponds to 69.3% of the population aged 15 to 64 years (CZSO 2015). Probably the most attentively observed indicator of labour market is the unemployment rate. Since 2008, the average rate of unemployment in the Czech Republic has been increasing. Beginning on 4.11% in 2008, it reached approximately 7.70% in 2014. Not suprisingly, the last year's lowest unemployment rate was in Prague, approximately 5.27% (CZSO 2015).

But how does the unemployment vary with the educational level? In 2012, the unemployment rate of 25-64 year-olds with the highest level of education below upper secondary education was 25.46 percent. From all of the OECD countries, it was the second highest number. However, it is much lower for the individuals with higher educational attainment. Approximately 5.69% of 25-64 year-olds with upper secondary or post-secondary non-tertiary education and 2.65% of tertiary educated of this age were unemployed in 2012 (OECD 2014a).

What might also be interesting for our analysis is the structure of employed population according to the educational level. In 2010, 39% of workers had secondary school without school-leaving examination as their highest educational level, which makes them the largest group. Similar fraction of employed population, 38% had completed secondary school with school-leaving examination. Workers with university education formed 18% workers, on the contrary, 5% of workers had only completed elementary school at the most. Residual 5% of workers had completed follow-up courses and tertiary professional schools (Lepič & Koucký 2012).

But perhaps more importantly, how are the workers with different level of qualification assigned to different jobs? Czech Statistical Office analyzed

[6]The sector of services

the situation from 2006 to 2010 in order to illustrate how the education of workers corresponds to qualification requirements of their occupation. Since the analysis was made before the ISCO classification of occupation has been adopted, an older classification KZAM was used.[7] Generally speaking, higher education is expected for categories 1 (Managers, Legislators) and 2 (Scientists and Professionals). Categories 3 to 8 (Qualified technicians and craftsmen, Clerical support workers, Services and sales workers, Plant and machine operators) usually require upper secondary education and category 9 (Helping and Non-qualified workers) are mainly intended for workers with elementary school. As we can see in Table 1, in 2006-2010, on average more than 34 percent of workers with tertiary education worked on less demanding positions (KZAM 3-9). Only 2.5 percent of workers with secondary school with school-leaving examination and 10.8 percent of the employees with secondary school without school-leaving examination had the occupation from category 9 which require less qualification (CZSO 2010).

Table 3.1: Assignment of Workers to the Occupations, by Education Level

	Lower education required	**Requirements match workers education**	**Higher education required**
Tertiary education	(KZAM 3-9) 34.3%	(KZAM 1-2) 65.7%	-
Secondary education, school leaving exam	(KZAM 9) 2.5%	(KZAM 3-8) 82.9%	(KZAM 1-2) 14.6%
Secondary education	(KZAM 9) 10.8%	(KZAM 4-8) 75.7%	(KZAM 1-3) 13.6%

Source: Author based on CZSO

Another labour market outcome of high importance is the level of wages. What might be interesting for us is not only the absolute level of average and minimum wage, but also their latest development as well as wage differentials between different groups of workers. In January 2015, minimum wage was adjusted by the government at 9 200 CZK per month, which corresponds to less than 48% of the EU average. It is though an increase of 700 CZK compared to the previous level of 8 500 CZK set in 2013. Before the increase in 2013, minimum wage level was stable for more than five years, which may be expected given the global economic and financial crisis (MoLSA 2015).

[7]For our computations, we use ISCO classification which is quite similar to KZAM classification. However, several differences may be found. For detailed description of these differences, see the hanbook on the website of the Czech Statistical Office available at `https://www.czso.cz/csu/czso/klasifikace_zamestnani_-cz_isco-`

As we can expect, there is an increasing trend in average nominal wage as well. In the fourth quarter of 2014, it reached 27 200 CZK[8] which is an increase of approximately 2.29 percent relative to the value in previous year's fourth quarter. As for the wage differentials between various groups of workers, the latest data provided by CZSO are from 2012. In this year, the absolute difference in average nominal wage of men and women was equal to 6233 CZK. If we express this pay gap in form of female to male monthly wage ratio, we get that women earned only about 78.4 percent of male wages.

Moreover, wages vary greatly across occupations. Among all CZ-ISCO[9] types of occupation, the highest wages were paid to categories 1 (Managers) and 2 (Professionals). On average, it was 58 343 CZK and 36 592 CZK, with female to male wage ratios 71.3% and 73%, respectively. On the contrary, the least paid occupations were the Elementary occupations (Helping and non-qualified workers) with average monthly wage of 14 627 CZK and female to male ratio 89.9% (CZSO 2012a).

Similarly, workers with different education earn different wages. Table 2 shows the earnings of differently educated workers in 2011 relative to the benchmark group of workers with secondary education. In the first column, we can see these differences between workers aged 25 to 64 years. Second column focuses on younger, 25-34 year-old workers. We can conclude that the situation in the Czech labour market confirms that the education in general leads to a wage premium, but its importance moderately declines with age (OECD 2014a).

Table 3.2: Relative Wages by Educational Attainment
Secondary Education = 100%

	Age of workers	
Education	**25 - 64**	**25 - 34**
Below upper secondary education	73	78
Secondary education	100	100
Tertiary education of type B	117	116
Tertiary education of type A	113	130

Source: Author based on OECD

[8]Czech statistical office published these informations about average wages in 2014 as preliminary with the expected correction to final values

[9]Extended version of ISCO-08 classification

It is worth mentioning that the Czech labour market seems relatively non-flexible. To assess the flexibility of labour market, number of "flexible" or "non-standard" forms of employment such as part-time, short term and contract work are monitored (Benjamin *et al.* 2002) (Kotýnková 2006). Only about 4.3% of total employment in 2012 was formed by work on a part-time contract. The OECD average was 16.9 percent. The share of temporary workers on dependent employment was 8.8%, which is slightly lower tan OECD average, 11.8% (OECD 2014b). A positive aspect of the low share of part-time employment is that there are fewer individuals working on part-time wishing to work more. These workers are also known as underemployed (Benjamin *et al.* 2002). Underemployment statistics are covered by Eurostat among relatively new supplementary labour market indicators[10]. In the third quarter of 2014, only 0.7% of Czech workers were underemployed compared to 4% average in EU-28 countries.

Another important aspect of labour market is a mobility of labour. As we already discussed in Chapter 2, the job-searching process is more or less geographically limited. But higher geographic mobility can be beneficial to both individuals and society. Not only the workers are able to choose between more job offers, but the diffusion of human capital enhances productive capacity of firms and prosperity of poor regions. As Benjamin *et al.* (2002) explain in Labour Market Economics textbook:

> "(...)economic theory predicts that the forces of competition would serve to reduce pure regional wage differentials (...) Those forces of competition were the movement of capital from high- to low-wage areas and the movement of labour from low- to high-wage areas."

Despite these positive effects, mobility of labour in the Czech Republic is still relatively low. A report on Geographic mobility in the European Union elaborated by Bonin *et al.* (2008) shows that several types of labour mobility that are being observed are low in case of Czech Republic.

[10]These are: underemployed part-time workers, jobless persons seeking a job but not immediately available for work and jobless persons available for work but not seeking it.

3.3 Previous Evidence of Overeducation in the Czech Republic

There are still few researches dedicated to overeducation the Czech labour market. Some studies such as (Ryška & Zelenka 2011) or (Gebicka 2010) explored overeducation among college graduates, estimating the number of graduates working in "non-college occupations".

With the data from 2005, based on survey among individuals who graduated in 2000, Barone & Ortiz (2011) included Czech Republic to their comparative analysis of European countries. Their computations lead to a conclusion that there were very few overeducated workers, the estimated incidence varied from 1.5% to 7.1% depending on the measure used. Barone and Ortiz explained that this exceptionally low estimate could be driven by the fact that the number of graduates in the Czech population was severly restricted at the time the data were collected. Interestingly, they also found that humanistic fields hepled to prevent overeducation, which is in sharp contradiction with common belief.

Recently, a more detailed study of educational mismatch has been made by Mysíková (2014). This study examined both overeducation and undereducation and focused on the determinants and wage effect of mismatches. Three datasets has been used, one of them, Programme for International Assessment of Adult Competencies(PIAAC), allowed to determine the number of self-declared mismatches. The estimated incidence of overeducation varied greatly depending on the method of measurement and dataset used. From about 6% using normative method and SILC[11] dataset (2012) to 25.4% using self-assessment method and PIAAC dataset (2011/2012). This study showed, above all, that results can be to a great extent dependent on the data source and method. Moreover, overeducation is no longer regarded as negligible, as the estimated extent is considerably higher than in the case of Barone and Ortiz (2010).

Our contribution is then twofold. We conduct an analysis of the impact of overeducation on the job satisfaction with focus on the Czech Republic, by using Czech data exclusively.[12] Then we extend and update past evidence about the incidence of overeducation, its determinants and its effect on earnings (using recent data and variations in commonly used models).

[11] SILC - Statistics on Income and living conditions. For further informations, see next section.

[12] So far, only Sánchez-Sánchez & McGuinness (2013) analysed the wage and satisfaction effect of overeducation with data from 13 countries, including Czech Republic. But this study analyses these relations for all countries at once, they only use a dummy for each country.

3.4 Hypothesis Statements

After having studied the overeducation literature and the Czech environment and its specifics, we can now proceed to a further specification of our research questions. So what drives the overeducation in the Czech labour market? Are the widely discussed factors important in our case? We construct the following hypotheses:

Hypothesis 1: Women are more overeducated than men.
Hypothesis 2: Immigrants are more likely to be overeducated.
Hypothesis 3: Spatial factors have impact on the probability of being overeducated.

Also, can we observe any negative consequences of overschooling in case of Czech workers? Considering the economic theory behind the overeducation, certain relationships can be assumed:

Hypothesis 4: Overeducated workers earn less than "matched" workers with the same education.
Hypothesis 5: The returns to years of overschooling are significantly lower than the returns to years of required schooling.
Hypothesis 6: Overeducated workers tend to be less satisfied with their job.

Chapter 4

Data and Methodology

4.1 Description of the Data

For the empirical part of this thesis, we use data from the European Union Statistics on Income and Living Conditions (SILC) dataset. This survey coordinated by the Eurostat reflects socio-demographic characteristics of individuals and households, their financial, health, working and housing conditions. More precisely, we work with its part for the Czech Republic, collected via the Czech Statistical Office, a cross-section capturing the situation at the time of collection in May 2013. We use only the data for individuals. This sample is formed by 19 105 observations.

For such a complex survey, sophisticated sampling designs are implemented to improve the quality of estimation. In case of SILC for the Czech Republic, data are collected by means of stratified multi-stage sampling. Considering the sampling design, we have to adjust our estimates to the whole population using the weighting coefficient PKOEF. This coefficient assigns statistical weights to individuals according to different factors including age, size of municipality, social group or income (Bartošová & Bına 2008).

The dataset contains high number of variables useful for the analysis of overeducation. It allows us to apply and compare both objective measures of overeducation, statistical and normative. Also, a great advantage is that we have information about wages and the module "well-being" composed of variables indicating a degree of satisfaction with different aspects of life, including work.

Moreover, we have information about the type of occupation according to the International Standard Classification of Occupations (ISCO). As we can see

in Table 3, ISCO distinguishes between 10 major groups (1-digit ISCO). These major groups are subsequently divided to sub-major groups (2-digit ISCO), then to minor levels and unit groups (3 and 4-digit ISCO). In case of the Czech Republic the classification used is CZ-ISCO. It is an extended version of ISCO as it has one additional level and the definitions are slightly changed to be in accordance with the Czech labour market. For our computations, we use first level (1-digit ISCO) with 10 major groups and second level (2-digit ISCO) with 43 sub-major groups, depending on the method of measurement.[1]

One inconvenience of this dataset is that it does not provide detailed information about years of schooling, only 10 educational levels: pre-school or non-completed first stage of elementary school (level 0), first and second stage of the elementary school (levels 1 and 2), secondary school without school-leaving examination (level 3), secondary school with school leaving examination (level 4), follow-up courses (level 5), tertiary professional schools (level 6), bachelor degree (level 7), master or engineer degree (level 8) and doctoral education (level 9). Since we need this for our computations, we derive the number of years of completed education based on the standard length of schooling in the Czech Republic corresponding to each educational level. See Appendix 1 for all the steps of this imputation.

For our analysis, we select the individuals in prime age, from 25 to 54 years old. This selection is motivated by the fact that the youngest workers are very specific and the patterns observed for this group might not apply to older workers, therefore including them might distort the overal results. For instance, in order to gain the first work experience, overeducation can be a temporary option for young workers. Also, a significant number of the individuals younger than 25 is still studying and thus the sample of workers in this age would by definition have lower education level than older workers. On the other hand, the oldest workers are gradually preparing for retirement, also, their health condition changes, which can considerably influence their career choices.

4.2 Selected Measures of Overeducation

To measure the incidence of overeducation and undereducation, we select full-time employees in prime age whose ISCO type of work is known. This reduces

[1] For an overview of ISCO sub-major groups, see `http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_172572.pdf`

the estimation sample to 5080 individuals.[2] As we discussed in Chapter 2, several measurement methods are used to estimate the incidende of overeducation. We apply both objective methods, statistical and normative.

The statistical method determines required education based on the actual situation on the labour market and realized matches. Its advantage is that it well reflects current situation in the labour market. Also, it is very clearly defined and available with a wide variety of datasets. But it has been criticized for ambiguous and seemingly arbitrary choice of the range of one standard deviation above and below the estimated required education (Dolton & Vignoles 2000).

But how do we apply the statistical method on our data? First we compare the number of years of education completed by an individual with the sample mean for his or her occupation type. For each one of the 43 2-digit ISCO groups separately, we estimate mean number of years of schooling. To illustrate the differences in educational attainment between different occupation types, we present descriptive statistics for the variable years of education by 1-digit ISCO categories in Table 4.1.

Table 4.1: Descriptive Statistics of the Variable Years of Education by ISCO Major Groups

1-digit ISCO category	**Number of observations**	**Mean**	**Median**	**Standard deviation**	**Min**	**Max**
Managers & legislators	246	15.11	15	2.69	10	21
Professionals	729	16.29	17	2.3	10	21
Qualified technicians	984	13.58	13	2.09	8	21
Clerical support workers	559	12.99	13	1.77	9	18
Services & sales workers	710	11.71	11	1.6	8	18
Agricultural workers	55	11.38	11	2.03	8	20
Qualified Craftsmen	846	11.20	11	1.16	8	18
Plant & machine operators	661	11.07	11	1.21	8	18
Non-qualified workers	261	10.79	11	1.25	8	16

Individuals within the range between one standard deviation above and below this value are classified as matched. If the individual's acquired level of education is higher than the mean number of years for his or her occupation type plus one standard deviation, we count them as overeducated. Similarly, if his or her educational attainment is lower than the mean for his or her occupation type minus one standard deviation, we count them as undereducated.

[2] 16 individuals that are employed and at the same time we do not know the ISCO type of their occupation, are excluded from this analysis, as we cannot determine the category they belong to.

Alternatively, we compare individual's years of completed schooling with the median number for his or her occupation type. For every ISCO sub-major group, we compute the median number of years of schooling. Again, the range of one standard deviation above and below the median determines which workers are matched, overeducated and undereducated.

The other method used, the normative method, is not based on actual educational attainment of all workers. The required education level is determined *ex ante*, based on job analysis. It is preferred by many authors for its ambition to be the most objective of all measures of educational mismatches. Nevertheless, it has its drawbacks, too. It is relatively static, which means that it does not easily adapt to recent development in the labour market, which affects job requirements (Hartog 2000).

Our dataset allows us to replicate the approach used in the OECD report (OECD 2007), which directly assumes levels of qualification needed for different types of occupations. ISCO major groups are divided into 3 broader groups: jobs requiring low, intermediate and high skills. This is illustrated in Table 4.2. Levels of education are rescaled to broader categories as well, as we can see in Table 4.3. Workers with high qualification are matched if their job requires high skills and overeducated if their job requires low or intermediate skills. Workers with intermediate qualification are overeducated in high-skill jobs, matched in jobs requiring intermediate qualification and undereducated in low-skill jobs. Workers with low qualification are matched in low-skill occupations and undereducated otherwise.

Table 4.2: Normative Measure: Skill Requirements by 1-digit occupational categories

1-digit ISCO category	**Required skill level**
Managers & legislators	high
Professionals	high
Qualified technicians	high
Clerical support workers	intermediate
Services & sales workers	intermediate
Agricultural workers	intermediate
Qualified Craftsmen	intermediate
Plant & machine operators	intermediate
Non-qualified workers	low

Source:Author based on OECD

Table 4.3: Normative Measure: Skill Levels corresponding to Education Levels

		Required Skill level		
ISCED level	**Corresponding levels in our dataset**	**low**	**medium**	**high**
Pre-primary education	0	x		
Primary education	1	x		
Lower secondary	2	x		
Upper secondary	3, 4		x	
Post-secondary non tertiary	5		x	
Tertiary first stage	6, 7			x
Tertiary second stage	8, 9			x

Source: Author based on OECD

One must be careful when applying this method. Considering the fact that the requirements of different occupation types were not studied for every country separately, the differences between countries have been neglected (OECD 2007).

The estimated incidence of overeducation resulting from an application of methods described above is presented in Chapter 5.

4.3 Determinants of Overeducation: Multinomial Logit

After estimation of the incidence of overeducation, we attempt to identify the determinants of overeducation. In other words, we explore the probability of being overeducated and the factors that might have impact on this probability. For this purpose, several models can be applied. Most studies estimate binary outcome models (Leuven & Oosterbeek 2011).

We have chosen the extended approach, accounting for the undereducation option. Thus the dependent variable takes three possible values. Individuals in prime age who are full-time employed, are divided into 3 disjoint groups: overeducated, matched and undereducated. For this purpose, multi-

nomial models accounting for three possible outcomes can be used. For an example, see Mysíková (2014) or Kiker *et al.* (1997).

We thus estimate the multinomial logistic regression, as this model is well-suited for non-ordered discrete outcome variables. This model works with any number S of alternative outcomes with index $s = 1, 2, ..., S$. For any observation i, the probability of the outcome s is given by the following relationship:

$$P_{is} = P(y_i = s) = \frac{exp(x_i^T \beta_s)}{\sum_{t=1}^{T} exp(x_i^T \beta_t)}$$

where $s = 1, ..., S$, x_i is a vector of regressors and β_s a vector of coefficients corresponding to the outcome s.

To enable identification of the model, one of the possible outcomes has to be set as a base outcome, or reference state. If the outcome $s = 1$ is chosen as a base outcome, the coefficients are redefined as: $\beta_s^* = \beta_s - \beta_1$ and thus $\beta_1 = 0$. Then the probabilities change to:

$$P_{is} = \frac{exp(x_i^T \beta_s)}{1+\sum_{t=2}^{T} exp(x_i^T \beta_t)} \qquad \text{for } s \neq 1,$$

$$P_{i1} = P(y_i = s) = \frac{1}{1+\sum_{t=2}^{T} exp(x_i^T \beta_t)}$$

Applying this model, a method of estimation is the maximum likelihood estimation. The vector of estimated coefficients maximizes loglikelihood function given by the formula:

$$logL = \sum_i logL_i = \sum_i \sum_s y_{is} log P_{is}$$

where y_{is} is a dummy variable equal to 1 if the outcome variable for the observation i is equal to s and 0 otherwise (Cramer 2003).

In our case, $S = 3$. Being matched has been chosen as a base outcome. Thus the estimated coefficients have to be interpreted relative to this state. As the interpretation of coefficients is not straight-forward, to determine direct effect on the probability of overeducation, the average marginal effects of each independent variable for the outcome overeducated are estimated.

The multinomial logistic regression assumes that the outcomes are disjoint and exhaustive, that is, for every observation, the independent variable has a single value. Moreover, the model assumes the Independence from irrelevant

alternatives (IIA). This means that the ratio of probabilities of any two outcomes depends only on the parameters of the outcomes in question, regardless of the other outcomes. Consequently, if some other alternative is added or removed, relative odds remain unchanged. For a further discussion and examples, see Cramer (2003) or Long & Freese (2006). There are several possibilities to test the IIA assumption. We apply the Small and Hsiao test. The idea is to examine changes in parameter estimates after a removal of an "irrelevant" alternative. The null hypothesis assumes no systematic change of estimates, but an improvement of the efficiency (Roman 2004).

We proceed with a brief discussion of the specification and choice of regressors. Following the educational mismatch literature, we examine the effect of gender, spatial factors and ethnicity. We are especially interested if female workers and workers whose nationality is not Czech are more likely to be overeducated. Also, a dummy variable for Prague can show if there is any systematic difference in probability of overeducation for residents of the capital. Besides these factors, we included different demographic characteristics, family arrangements, working environment, and another education and work-related variables. As Leuven & Oosterbeek (2011) point out, researchers in this field rarely reveal their motivation for a choice of particular control variable. Therefore we have to rely on intuition based on observed patterns in the labour market.

4.4 Wage effects of Overeducation: Mincer Model Extensions

Specifications used for the analysis of pay penalty of overeducation are based on the original Mincer model (1974) specified by the relatonship:

$$logW = \beta_1 S + \beta_2 Ex + \beta_3 Ex^2 + \epsilon \qquad (1)$$

where W represents wage, S years of acquired schooling, and Ex the experience (Heckman *et al.* 2003). An important feature of this specification is that it provides a useful tool to analyze the rate of return to additional year of education. The interpretation is following. The two individuals i and j, whose characteristics are identical except that one possess one more year of schooling, will have different wages according to the relationship:

$$\begin{aligned} logW_i - logW_j &= \beta_1 \\ logW_i - logW_j = log(W_i/W_j) \quad &\Longrightarrow \quad \tfrac{W_i - W_j}{W_j} = exp(\beta_1) - 1 \end{aligned}$$

The last expression represents proportional difference in earnings resulting from an additional year of education, or alternatively, the rate of return to one year of education (Bazen 2011).

For purposes of measuring wage effects of mismatches, the original Mincer is extended. The variable indicating years of acquired education is replaced by variables indicating years of schooling required and years of schooling below or above required level:

$$logW = \gamma X + \beta_1 S^r + \beta_2 S^o + \beta_3 S^u + \beta_4 Ex + \beta_5 Ex^2 + \epsilon \qquad (2)$$

where X is a vector of characteristics correlated with wage, S^r, S^o, S^u denote years of schooling required, years of surplus schooling and years of deficit schooling, respectively. As the years of actual schooling S are divided into these three components, S^o and S^u are defined the following way:

$$\begin{aligned} S^o &= max(0, S - Sr) \\ S^u &= max(0, Sr - S) \end{aligned}$$

For overeducated individuals $S_o > 0$ and $S_u = 0$, for undereducated individuals $S_o = 0$ and $S_u > 0$. In case of matched individuals $S_u = S_o = 0$ and number of years of schooling required S_r is equal to the number of actual years of schooling (Hartog 2000). The model (2), sometimes called ORU specification (which stands for Over, Required and Undereducation), was introduced by Duncan & Hoffman (1982).

Alternatively, instead of years of schooling, dummy variables indicating the mismatches can be used:

$$logW = \gamma X + \delta_1 D^O + \delta_2 D^U + \beta_1 Ex + \beta_2 Ex^2 + \epsilon \qquad (3)$$

where D^O is a dummy variable indicating that an individual is overeducated, D^U denotes the presence of undereducation and vector X can alliteratively contain actual years of schooling (McGuiness 2006) (Leuven & Oosterbeek 2011). This specification was first used by Verdugo & Verdugo (1989).

Both specifications are linear in parameters, which leads to the use of the Ordinary least squares estimation (OLS). However, they are very different when it comes to interpretation. The ORU specification (2) allows to separate the

returns to required education, returns to surplus education and pay penalty of deficit education. We can compare overeducated and undereducated workers with the individuals having the same job, or the job where the required education is the same. The Verdugo and Verdugo specification (3) compares overeducated and undereducated workers with equally educated individuals, whose job matches their educational attainment. Also, it only controls for actual years of education and thus it does not separate returns to required education from the returns to surplus education (Leuven & Oosterbeek 2011) (Mysíková 2014). As the interpretations differ, we refer to both of these commonly used specifications. For the Verdugo and Verdugo specification, two sets of dummy variables are defined, based on the normative and statistical measure. For the ORU specification, only the statistical measure of incidence can be used, because the normative measure does not include the estimation of required years of schooling.

In any case, the analysis of wage effects of overeducation might be tricky. The estimates are potentially biased, for several reasons. First of all, there might be a problem with nonrandom selection of the estimation sample. If this is the case, one of the CLM assumptions is violated, therefore a bias can be present. A frequently used method in labour economics is a two-step estimation introduced by Heckman (1979). It consists of estimation of the selection process by a probit model in the first step, and following inclusion of the Inverse Mill's ratio into the equation of interest in the second step. This allows to account for the selection process and thus corrects the sample selection bias (Greene 2008).

In our case, the situation is similar to Heckman's example of female labour supply.[3] Two types of wage should be considered. A *market wage* represents how the person's effort is valued in the labour market and *reservation wage* is a minimum necessary for a person to participate in the labour force. The problem is that we can only observe market wages for the individuals whose reservation wage is lower than the observed one. The so-called self-selection is then present, which implies non-randomness.

With survey data, a two-step Heckman selection is no longer appropriate, as the weights are used. The "standard" OLS estimation in the second step would not yield consistent estimates.[4] However, for this purpose we can use an

[3]See Greene (2008), p.782

[4]For more informations, see STATA survey data manual and manual for the Heckman selection model.

extension of Heckman selection model which applies Maximum Likelihood Estimation. This model comprises more complicated computations, thus we will not describe them here and we refer the interested reader to Nawata (1994) for detailes of this approach. Generally, the idea is similar to the original Heckman model and the estimation should yield constistent and unbiased estimates.

Apart from the non-random selection of the estimation sample, there is another possible obstacle. The variable indicating years of completed education is potentially endogenous. Moreover, variable indicating years of required education might be endogenous as well. It is due to the problematic control for the heterogeneity of workers, such as unobserved skill level having impact on both wage and educational choices.

Probably the most powerful corrective tool is provided by fixed effects techniques. Unfortunately it requires panel data, thus it is beyond the scope of this paper. For an example, see Bauer (2002). This study points to an upward bias of the estimates resulting from widely used OLS estimation. This could lead us to a conclusion that OLS overestimates the effect of overeducation on wages. However, Dolton & Silles (2008) went even further and presented measurement-error-corrected panel estimates. These were not significantly different from OLS estimates.

Another possibility is to use instrumental variables. For instance, in their analysis of overeducation in Sweden, Korpi & Tåhlin (2009) used sibship size, place of residence in the childhood, disruption and economic problems in the family as the instruments for years of schooling. But as these instruments are weak, the improvement compared to the ordinary least squares estimation is questionable .

Third technique, called propensity score matching, has been applied to the overeducation research by McGuiness (2007). It is also subject to criticism, because matching assumes that selection to different education levels is fully based on observed factors. As Leuven and Oosterbeek (2011) concluded: “His claim that this approach addresses omitted variable problems is in our view not realistic. At best, his results show that previous findings obtained from OLS are not attributable to common support problems.”

4.5 Overeducation and Job Satisfaction: Linear Regression

We proceed with exploring another possible consequence of overeducation: lower satisfaction with the job. For this part, only the individuals whose job satisfaction is known are selected. This reduces our estimation sample to 4911 observations. In our dataset, the variable indicating job satisfaction takes values from 0 (Not at all satisfied) to 10 (Completely satisfied). This is a discrete choice, plus the outcomes are ordered.

Therefore the most natural way to model job satisfaction would be to apply an ordinal regression. An important difference compared to linear regression model is that the ordered probability models account for possible ordinal nature of the dependent variable. In other words, numerical values of the differences between categories are unimportant, only the order matters. As a consequence, unlike linear regression, ordered logit and probit allow "distances" between categories to be unequal (Treiman 2009). In case of self-assessed satisfaction and well-being, there is a reasonable suspicion that these variables are ordinal. For instance, we can hardly compare utility scales between different individuals.[5]

Nevertheless, ordinal regression models pose a serious restriction. Parallel regressions assumption, in case of ordered logit also known as proportional odds assumption, states that the probability of response variable being in a given category or lower relative to higher categories is the same for all values of response variable. This implies that if we expressed the ordinal model as a set of binary regressions, coefficients would not differ (Agresti & Finlay 2009). As there are indications that this crucial assumption is not met in our case[6], we follow the example of Fleming & Kler (2014) and proceed with linear regression. As Ferrer-i Carbonell & Frijters (2004) point out, treating satisfaction variable as cardinal makes little difference and does not cause any significant change in estimated relationships.

The estimated model can be written as follows:

$$JS = \gamma X + \delta_1 D^O + \delta_2 D^U + \beta_1 logW + \beta_2 Ex + \beta_3 S + \epsilon$$

where JS denotes job satisfaction, W wage, Ex experience, S years of

[5]For a detailed discussion of ordinal nature of these variables, consult Powdthavee (2007)

[6] With sampling weights, Brant and LR test are not available in STATA. But to chceck the assumption, generalized ordered logit model can be applied. In our case, if generalized ordered logit is applied, the coefficient estimates are significantly different for each outcome, even when the response variable is rescaled to broader categories.

schooling, D^O and D^U are dummies for overeducation and undereducation and X is a vector of personal characteristics and work related characteristics including various employee benefits.

There is a complication associated with using OLS estimation with a response variable of this type. The disturbances are most likely heteroskedastic and non-normally distributed. The estimated standard errors thus can be biased and confidence intervals and test statistics constructed based on these errors are no longer valid (Wooldridge 2012). Fortunately, this concern should not be relevant for our analysis. With sampling weights, OLS estimation is by default heteroskedasticity robust. Also, our sample is very large, it contains several thousands of observations. Therefore the model accounts for heteroskedasticity of unknown form and the estimated standard errors are still useful for statistical inference.

Chapter 5

Results

In this chapter, we present and discuss the results from our analysis of overeducation obtained by an application of methods described in Chapter 4.

5.1 Incidence of Overeducation

Table 5.1 summarizes the estimated percentage of overeducated and undereducated individuals in prime age. The statistical measure, based on realized matches, is applied using both mean and median years of education for the range of occupations. Along with an estimate of total incidence of overeducation and undereducation, percentages of overeducated and undereducated female workers and male workers are presented. We classify individuals as overeducated or undereducated based on the mean or median estimated for both genders at once. Then the incidence of overeducation and undereducation for female workers is calculated as a share of women classified as undereducated or undereducated on total number of female full-time employees, respectively. For male overeducation and undereducation analogically.

As we can see, different measures lead to quite different results. With normative measure, estimated share of overeducation is relatively low, only 6.2%. Share of undereducation, on the other hand, is extraordinarily high, about 22.9%, which is nearly four times higher than the estimated share of overeducation. Statistical measure results in more balanced estimates of mismatches. If sample mean is used, 15.6% of workers are counted as overeducated and 15.7% as undereducated. If required education is determined according to median, resulting shares are the following : 15.6% overeducation and 11.1% undereducation. So the statistical measure yields very similar estimates of

Table 5.1: Incidence of Overeducation and Undereducation

	Statistical measure (mean)	**Statistical measure (median)**	**Normative measure**
		OVEREDUCATION	
Total	0.156 (0.006)	0.155 (0.006)	0.062 (0.004)
Female workers	0.149 (0.007)	0.152 (0.008)	0.083 (0.006)
Male workers	0.163 (0.008)	0.159 (0.007)	0.043 (0.004)
		UNDEREDUCATION	
Total	0.157 (0.006)	0.111 (0.005)	0.229 (0.007)
Female workers	0.152 (0.008)	0.125 (0.008)	0.249 (0.105)
Male workers	0.161 (0.008)	0.098 (0.007)	0.212 (0.009)

Notes: N=5080. Full-time employees in prime age with known ISCO type of occupation were selected. Standard errors in parentheses.

overeducation for mean and median, but the estimate of total undereducation is lower if median is used.

Interestingly, if we look at gender differences, the two measures produce opposite results. With normative measure, share of overeducation is about 4 percentage points higher for women. Percentage of the undereducated is also higher for female workers. With statistical measure this no longer holds, more men are defined as overeducated. In any event, differences between female and male overeducation and undereducation are not large, up to 4%.[1]

5.2 Determinants of Overeducation

The analysis of determinants of overeducation is exerced separately for normative and statistical measure, since different individuals are identified as overeducated and undereducated. As the incidence of overeducation estimated by means of statistical measure is very similar for mean and median approach, we proceed only for the mean approach.The reason for choosing the mean approach is that it is more frequently used in the literature. For both measures, we selected two specifications that best fit our data. Table 5.2 presents the estimated coefficients and standard errors for an outcome "Overeducated" as

[1]These gender differences are also statistically significant. Results from a t-test indicate that all of these differences are significantly different from zero.

well as the values of selected measures of fit in each case. As we pointed out in Chapter 4, the coefficients do not directly represent the effect of a change in regressors on the probability of being overeducated. They must be interpreted relative to the base outcome "Matched". To simplify the interpretation and to better illustrate the impact of examined factors on the risk of overeducation, we present the estimated average marginal effects (along with their standard errors) in Table 5.3. Note that for binary variables, marginal effects show the estimated difference in probability of outcome "Overeducated" relative to the reference category "Matched" driven by the change of the regressor from zero to one.

How was the performance of each model validated? Firstly, in each case, a comparison of the Wald statistics with the Chi-squared distribution lead to a conclusion that the null hypothesis of joint insignificance of all coefficients is strongly rejected. In other words, selected specifications fit our data significantly better than the model with no regressors. If we look at the McFadden pseudo R-squared, it confirms the superiority of the models compared to the intercept-only model.[2] The performance of a model can be also evaluated through a percentage correctly predicted, also known as hit rate. For selected specifications, this number varies from 72.87% to 74.53%. For each specification, the Small and Hsiao test for IIA assumption has been executed. Since the null hypothesis has not been rejected, we can conclude that no violation has been identified. The results of these tests are reported in Appendix 2.

The experience in years has been included as a control in all specifiactions. In case of statistical measure of overeducation, we controlled for the education in years, and we included its square as well. The motivation is to account for possibly nonconstant effect of additional years of education on the probability of overeducation. Dummies for education levels were included for the normative measure. In this case, the reference category are workers with less than secondary education with a school-leaving examination. Workers with secondary education with a school- leaving exam were found to be more overeducated relative to the base category, this no longer holds for tertiary educated workers. It supports the inclusion of the squared term (years of completed education squared) as there are indications that the effect of additional education on the probability of overeducation starts to decrease at some point.

[2]In McFadden's words, for multinomial logit, R-squared between 0.2 and 0.4 indicates an "excellent fit", see (McFadden 1977)

Table 5.2: Multinomial logit: Estimated Coefficients of the Outcome "Overeducated" and their Standard Errors

	Statistical measure		**Normative measure**	
	(1)	(2)	(3)	(4)
Female	-0.381**	-	0.606**	-
	(0.134)		(0.179)	
Parent	0.065	0.050	-0.203	-0.197
	(0.119)	(0.118)	(0.145)	(0.144)
Married	0.062	-	-0.072	-
	(0.120)		(0.144)	
Married Female	-	-0.376†	-	0.552*
		(0.209)		(0.248)
Single Female	-	-0.249	-	0.523**
		(0.158)		(0.202)
Married Male	-	0.228	-	-0.219
		(0.157)		(0.224)
Nonczech	-0.161	-0.161	0.780*	0.773*
	(0.293)	(0.292)	(0.326)	(0.326)
Prague	-1.026**	-1.017**	-0.646*	-0.650*
	(0.184)	(0.184)	(0.272)	(0.272)
Managerial position	-0.444**	-0.451**	-1.022**	-1.017**
	(0.119)	(0.119)	(0.255)	(0.255)
Large Firm	-0.385**	-0.381**	0.167	0.163
	(0.124)	(0.124)	(0.154)	(0.155)
Years of education	3.768**	3.760**	-	-
	(0.440)	(0.439)		
Years of education squared	-0.112**	-0.112**	-	-
	(0.015)	(0.015)		
Experience	0.027**	0.028**	0.000	-0.000
	(0.006)	(0.006)	(0.008)	(0.008)
Tertiary education	-	-	0.222	0.228
			(0.169)	(0.169)
Secondary education, school leaving exam	-	-	-1.021**	-1.023**
			(0.196)	(0.196)
Head of the household	0.077	-0.477	-0.271	-0.197
	(0.134)	(0.162)	(0.182)	(0.196)
Sick	-0.044	-0.020	1.024**	1.024**
	(0.433)	(0.432)	(0.364)	(0.364)
Constant	-31.781**	-31.716**	-2.256**	-2.244**
	(3.159)	(3.151)	(0.234)	(0.233)
Mc Fadden R-squared	*0.260*	*0.261*	*0.211*	*0.211*
Wald Chi-squared (22/24)	*2022.16*	*2022.46*	*376712.07*	*172283.99*
Percent correctly predicted	*72.87%*	*72.91%*	*74.53%*	*74.53%*

*Notes: N=5080. Weighted. Robust standard errors in parentheses. ** significance at 1% level, *significance at 5% level, †significance at 10% level.*

Now, how can we interpret output from the estimation? Which factors can be linked to a higher probability of being overeducated? And what are the other interesting findings?

- Women were found to be significantly more overeducated than men in case of normative measure, yet the statistical measure yields the opposite result.

- To examine the effect of marital status separately for men and women, we included interaction terms in specifications 2 and 4. According to the normative measure, women are more likely to be overeducated than men and the risk is slightly higher if they are married. But there is no significant difference between single and married men. If the statistical measure is applied, the result is different. At 10% level of significance, there are indications that both married and single women are less likely to be overeducated than single men, and the probability of overeducation is lower for married women.

- We did not find any significant impact of the parental status on the probability of overeducation, the same holds if an individual is the head of the household.[3]

- Prague dummy is statistically significant across all of the specifications and have negative marginal effects on probability of overeducation. This indicates that residents of the capital are less likely to be overeducated. Spatial factors thus might matter for the probability of overeducation.

- Workers with ethnicity other than Czech are significantly more likely to be overeducated in case of normative measure, in case of statistical measure, no significant relation has been identified.

- Managerial positions were found to be linked with a lower probability of overeducation, across all of the specifications.

- With the normative measure workers who assessed their general health as "bad" or "very bad" were found to have considerably higher probability of being overeducated. In case of statistical measure, there is no significant relationship.

[3]We also explored interactions of gender and parental status. As no significant patterns were identified, these are not included in any of the final specifications.

Table 5.3: Average Marginal Effects after Multinomial Logit, for the Outcome "Overeducated"

	Statistical measure		**Normative measure**	
	(1)	(2)	(3)	(4)
Female	-0.408**	-	0.031**	-
	(0.014)		(0.010)	
Parent	0.008	0.006	-0.011	-0.011
	(0.012)	(0.012)	(0.008)	(0.008)
Married	0.006	-	-0.005	-
	(0.124)		(0.008)	
Married Female	-	-0.040†	-	0.028*
		(0.021)		(0.012)
Single Female	-	-0.027†	-	0.027*
		(0.016)		(0.011)
Married Male	-	0.023	-	-0.013
		(0.016)		(0.012)
Nonczech	-0.016	-0.016	0.042**	0.041**
	(0.030)	(0.030)	(0.018)	(0.018)
Prague	-0.108**	-0.107**	-0.041**	-0.041**
	(0.018)	(0.018)	(0.015)	(0.015)
Managerial position	-0.052**	-0.052**	-0.065**	-0.065**
	(0.012)	(0.012)	(0.014)	(0.014)
Large Firm	-0.040**	-0.040**	0.009	0.009
	(0.013)	(0.013)	(0.008)	(0.009)
Years of education	0.404**	0.403**	-	-
	(0.046)	(0.046)		
Years of education squared	-0.011**	-0.012**	-	-
	(0.002)	(0.002)		
Experience	0.003**	0.003**	-0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)
Tertiary education	-	-	-0.256**	-0.177**
			(0.023)	(0.016)
Secondary education, school leaving exam	-	-	-0.068**	-0.068**
			(0.011)	(0.011)
Head of the household	0.008	-0.005	-0.016	-0.011
	(0.014)	(0.017)	(0.010)	(0.011)
Sick	-0.004	-0.001	0.057**	0.057**
	(0.044)	(0.044)	(0.020)	(0.020)

*Notes: **singificance at 1% level, *significance at 5% level, †significance at 10% level.*

- Firm size might play a role in the assignment of workers to less demanding positions. If the statistical measure of mismatches is used, large firms (with 50 workers at minimum) are possibly linked with a lower probability

of overeducation.

In conclusion, the probability of overeducation of Czech workers might be influenced by several factors. Based on our results, we cannot say that women are more overeducated than men. Although normative measure indicates that the expected higher female overeducation might be true, the statistical measure leads to the opposite result. Similarly, the nationality other than Czech could be linked with higher risk of overeducation according to normative definition, but this claim is not supported if we apply the other method. On the other hand, there is an evidence (consistent across different measures and specifications) that residents of the capital are less likely to be overeducated. The other factors possibly having influence on the risk of overeducation include managerial positions, general health and firm size.

5.3 Wage Effects of Overeducation

Firstly, we estimate the Verdugo and Verdugo specification, which uses dummy variables for educational mismatches. But instead of using linear regression, an extended Heckman selection model is used. Therefore the method of estimation is Maximum Likelihood Estimation, and it is exerced in two steps. Table 5.4 presents the second stage results, estimated coefficients and standard errors. The interpretation of coefficients is similar to linear regression estimated by OLS, they represent marginal effect of regressors on wage.

To check for possible multicollinearity problem, cross-correlation tables are constructed and reported in Appendix 3. They confirm that there are no indications of multicollinearity. Looking at the Wald statistics, in each case the null hypothesis of the joint insignificance of the regressors is strongly rejected. Apart from the coefficients, model estimates the number λ (lambda).This number is an estimate of the coefficient of Inverse Mill's Ratio and summarizes the effect of selection.[4] It can be also expressed as a product of estimated correlation between the error terms in wage equation and selection equation (ρ), and estimated standard error of the residuals in the wage equation (σ):

$$\lambda = \rho.\sigma$$

Its significance (along with the significance of ρ) indicates the presence of a selection problem and supports the use of the selection model.

[4]In selection models, the Inverse Mill's Ratio can be interpreted as "non-selection hazard".

Table 5.4: Wage Effects of Overeducation: Verdugo and Verdugo Specification

	Statistical measure	**Normative measure**
	(1)	(2)
Overeducated	-0.049†	-0.031
	(0.029)	(0.052)
Undereducated	0.103**	0.174**
	(0.026)	(0.040)
Experience	0.036**	0.034**
	(0.004)	(0.034)
Experience squared	-0.001**	-0.001**
	(0.000)	(0.000)
Years of education	0.052**	0.055**
	(0.008)	(0.008)
Lambda	*-0.523***	*-0.521***
	(0.020)	*(0.020)*
Wald Chi-squared	*1982.36*	*1912.82*

Notes: Weighted. Independent variables included in selection equation comprises interactions for gender and parental status, dummies for marital status and head of the houlsehold, age, years of education. The second stage regression includes dummies for ISCO major groups, Prague residents, nonczech ethnicity, gender, managerial position, unlimited contract, parental status and general health. N=7189 for the first stage and N=5343 for the second stage. Robust standard errors in parentheses.
***significance on 1% level *significance on 5% level,† significance on 10% level,*

As we can see, the results slightly differ between the two measures, although they have the same signs and order of magnitude. In case of normative measure, coefficient of the overeducation dummy is not statistically significant on conventional levels. However, using statistical measure, the coefficient of the overeducation dummy is significant at 10% level of significance. This coefficient can be interpreted the following way. *Ceteris paribus*, the overeducation results in pay penalty of approximately 4.9%. Overall, our analysis yields some evidence to support the hypothesis that overeducated workers might earn less than their matched counterparts. But this evidence is relatively weak, as the use of another measure does not support the hypothesis. On the other hand, for both measures, there are indications that undereducated workers earn systematically more than identically educated matched workers. We estimated that on average, this difference is approximately 10% in case of statistical measure and about 17% in case of normative measure.

A different view is provided by implementation of ORU specification. The

Table 5.5: Wage Effects of Overeducation: ORU Specification

	Statistical measure
Years of education required	0.114**
	(0.009)
Years of overeducation	0.018**
	(0.007)
Years of undereducation	-0.038**
	(0.008)
Experience	0.034**
	(0.004)
Experience squared	-0.001**
	(0.000)
Lambda	*-0.513** (0.020)*
Wald Chi-squared (21)	*1986.00*

Notes: Weighted. Independent variables included in selection equation are: interactions for gender and parental status, gender and marital status, dummies for head of the houlsehold, age, years of education. The second stage regression includes dummies for ISCO major groups, Prague residents, nonczech ethnicity, gender, managerial position, unlimited contract, parental status and general health.
N=7189 for the fisrt stage and N=5343 for the second stage.
Robust standard errors in parentheses.
**significance on 5% level, **significance on 1% level*

effect on earnings is estimated by dividing the years of actual completed education into three components: years of education required, years of overeducation and years of undereducation. Cross-correlation table reported in Appendix 3 indicates that multicollinearity problem is not present. Table 5.5 presents the estimated coefficients as well as their standard errors.

Again, comparison of the Wald statistics to chi-squared distribution indicates that the hypothesis of joint insignificance of all the independent variables is strongly rejected. Significance of lambda points out to nonzero correlation between the error terms in wage equation and selection equation. The use of selection model therefore appears to be appropriate.

As we discussed in Chapter 4, estimated coefficients of the years of education can be interpreted as returns to additional year of education. Therefore

our results are not very surprising. All three components of education are statistically significant at 1% level. The estimated returs to required education are the highest, in this case, an additional year of education results in approximately 11% increase in wage. Estimated wage effects of schooling above the required level are much lower, one year of "surplus" education leads to about 1.8 percent increase in wage. On the other hand, negative sign of the estimate of undereducation coefficient indicates that one deficit year of education decreases wage by approximately 3.8 percent.

Overall, the results of our analysis of wage effects of overeducation are in accordance with the theoretical debates and results from the past studies. Overeducation might be linked to lower earnings and the returns to schooling are significantly lower after surpassing required education level.

5.4 Overeducation and Job Satisfaction

Again, the analysis of the effect of overeducation on the job satisfaction is conducted separately for each measure of overeducation. Results are presented in Table 5.6. In each case, besides wage, experience and years of schooling, we control for work-related characteristics including managerial positions and type of contract. Moreover, 8 dummies for employee benefits[5] are included in the model. Variable "Fired" indicates that the worker was obliged by employer to quit his previous job. Other personal characteristics include gender and relation to the rest of the household.

We constructed cross-correlation tables to check for possible multicollinearity problem. These tables are reported in Appendix 3 and indicate that there is no sign of serious multicollinearity. As we discussed in Chapter 4, there is no need to check for heteroskedasticity of unobservables, as the estimation is already heteroskedasticity robust.

Reported F-statistics imply that for both regressions, the hypothesis of joint insignificance of all variables is rejected at 1% level of significance. Looking at the R-squared, we can see that little variance of job satisfaction is explained by dependent variables. More precisely, R-squared is approximately 0.101 in case of normative measure and 0.096 in case of statistical measure. This is expected, given the nature of our data. Many personal and work-related characteristics important for the explanation of job satisfaction, such as personality traits or

[5] These are company car, insurance, language courses, mobile phone, meal vouchers, transport subsidies, subsidies for sport and discount on company's products and services.

corporate culture, are not captured. However, this problem is not uncommon in the literature and the identified relationships between variables may still be illustrative.

What is thus the estimated effect of overeducation on job satisfaction? In case of normative measure of overeducation, the estimated coefficient of overeducation dummy is negative and statistically significant at 1% level. The interpretation is simple. On average, overeducated individuals report lower levels of job satisfaction than the reference group of workers whose qualification matches job requirements. This provides strong evidence that with this measure of overeducation, negative consequence in form of dissatisfaction is possible.

Table 5.6: Overeducation and Job satisfaction: OLS Coefficient Estimates and Robust Standard Errors

	Normative measure (1)	**Statistical measure** (2)
Overeducated	-0.781**	-0.117
	(0.217)	(0.165)
Undereducated	0.507**	0.221
	(0.138)	(0.169)
Log wage	0.702**	0.836**
	(0.150)	(0.148)
Female	2.434**	2.497**
	(0.139)	(0.138)
Managerial	0.142	0.214
	(0.157)	(0.156)
Permanent	0.101	0.0928
	(0.191)	(0.193)
Experience	0.0115†	0.0140*
	(0.007)	(0.007)
Years of	0.104**	0.0920**
education	(0.0263)	(0.0299)
Fired	-0.293	-0.293
	(0.416)	(0.421)
Head of the	1.410**	1.426**
household	(0.137)	(0.138)
Constant	-7.521**	-9.095**
	(1.734)	(1.702)
R-squared	*0.101*	*0.096*
F-statistic	*30.28*	*28.30*
(18, 4982)		

Notes: N= 4911. Weighted. Robust standard errors in parentheses.
8 dummies for employee benefits were included.
*** significance on 1% level, * significance on 5% level, † significance on 10% level*

On the other hand, statistical measure does not confirm the negative effect of overschooling on job satisfaction. Although the estimated coefficient of overeducation is negative, it is not statistically significant at conventional levels.

To conclude, the results are not surprising. As we expected, there are some indications that overeducation can negatively influence worker's job satisfaction. With our models, this relationship is proven in case of normative measure, but this result is not confirmed using statistical measure. Ideally, the analysis should be repeated with data which are better suited for exploring well-being and satisfaction.

Chapter 6

Conclusion

The existence of overeducation may have certain implications relevant to both individuals and society. From the point of view of individuals, education can be regarded as an investment. As returns to additional years of education can be questionable, individuals may want to reconsider further investments to their human capital. Instead, it can be reasonable to enter the job market earlier. This allows to gather experience and to have access to on-the-job training, which is also likely to be related to higher earnings in the future.

Another potential concern is that overeducated workers tend to be less satisfied with their jobs which also decreases their utility. Consequently, overeducation can be costly to the firms, too. Lower job satisfaction and lack of motivation could result in lower productivity of overeducated workers. If this claim turns out to be true, employers should weight this risk against the advantages of hiring employees with surplus qualification.

Our analysis yields several interesting results. Using statistical measure, incidence of overeducation exceeds 15%, but with normative method it reaches only about 6% of full-time employees in prime age. Moreover, there are differences between identified determinants of overeducation. With the normative measure, female workers tend to be more overeducated, the same holds for workers with ethnicity other than Czech. Statistical measure does not support these results. However, residents of the capital were found to be less likely to be overeducated, using both measures.

With the statistical measure, overqualified individuals were found to earn less than their matched coworkers. Returns to "surplus" education are lower than returns to required education. Normative measure did not yield any plausible evidence of negative wage effects. But it does support the claim that

overeducated workers are less satisfied with their job.

Evidently, measurement of the overeducation and analysis of its effects and determinants can be a challenging task. Overeducation is still not clearly defined, different measures identify different workers as overeducated. Therefore, results greatly depend on the chosen definition and related measure. This decreases the comparability between different studies and complicates the understanding of this phenomenon. But considering the fact that each measurement method has its advantages and its drawbacks, it cannot be easily decided which of them is more accurate and better illustrates the real situation.

Furher research should concentrate on fixing the imperfections of the empirical methods. For instance, it would be very informative to conduct an extended analysis of the Czech overeducation with panel data, to cope with endogeneity using fixed effect techniques. Panel data would also enable an analysis of the overeducation dynamics to see whether this mismatch is persistent or temporary.

Despite all of the concerns, education cannot be reduced to its economic value. Its importance goes far beyond such function. From this perspective, Oscar Wilde was right and there is no overeducation.

Bibliography

Agresti, A. & B. Finlay (2009): *Statistical methods for the social sciences: With SPSS from A to Z: A brief step-by-step manual.* Boston: Allyn and Bacon.

Barone, C. & L. Ortiz (2011): "Overeducation among European university graduates: a comparative analysis of its incidence and the importance of higher education differentiation." *Higher Education* **61(3)**: pp. 325–337.

Bartošová, J. & V. Bina (2008): "Influence of the calibration weights on results obtained from czech SILC data." *COMPSTAT 2010 Book of Abstracts* p. 367.

Battu, H., C. R. Belfield, & P. J. Sloane (1999): "Overeducation among graduates: a cohort view." *Education economics* **7(1)**: pp. 21–38.

Bauer, T. K. (2002): "Educational mismatch and wages: a panel analysis." *Economics of Education Review* **21(3)**: pp. 221–229.

Bazen, S. (2011): *Econometric methods for labour economics.* Oxford University Press.

Becker, G. S. (1964): *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education.* New York: Columbia University Press.

Benjamin, D., M. Gunderson, & W. C. Riddell (2002): *Labour market economics: Theory, evidence and policy in Canada.* McGraw-Hill Ryerson.

Bonin, H., W. Eichhorst, C. Florman, M. O. Hansen, L. Skiöld, J. L. Stuhler, K. Tatsiramos, H. Thomasen, & K. F. Zimmermann (2008): "Report no. 19: Geographic mobility in the european union: Optimising its economic and social benefits." *Technical report*, Institute for the Study of Labor.

Buchel, F. & M. van Ham (2003): "Overeducation, Regional Labor Markets and Spatial Flexibility." *Journal of Urban Economics* **53**: pp. 482–493.

Buchel, F. & A. Mertens (2010): "Overeducation, Undereducation and the Theory of Career Mobility." *Max Planck Institute for Human Development Discussion Paper* **(195)**.

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

Card, D. (1999): "The Causal Effect of Education on Earnings." *Handbook of Labour Economics* **3**: pp. 1802–1859.

Čermáková, Z., D. Holda, & V. Urbánek (1994): "Changes in funding of higher education in the Czech Republic." *European Journal of Education* **27**: pp. 75–83.

Chevalier, A. (2003): "Measuring Over-education." *Economica* **70**: pp. 509–531.

Cramer, J. S. (2003): *Logit models from economics and other fields*. Cambridge University Press.

CZSO (2010): "Význam vzdělání pro trh práce v ČR." .

CZSO (2012a): "Průměrné hrubé měsíční mzdy zaměstnanců podle pohlaví a hlavních tříd klasifikace zaměstnání CZ-ISCO." .

CZSO (2012b): "Studenti a absolventi vysokých škol v ČR." .

CZSO (2014): "Vzdělání obyvatelstva České Republiky ve věku 15 a více let podle výsledků výběrového šetření pracovních sil." .

CZSO (2015): "Zaměstnanost a nezaměstnanost v ČR podle výsledků VŠPS-4. čtvrtletí 2014." .

Dolton, P. & A. Vignoles (2000): "The Incidence and Effects of Overeducation in the UK Graduate Labour Market ." *Economics of Education Review* **19**: pp. 179–198.

Dolton, P. J. & M. A. Silles (2008): "The effects of over-education on earnings in the graduate labour market." *Economics of Education Review* **27(2)**: pp. 125–139.

Duncan, G. J. & S. D. Hoffman (1982): "The incidence and wage effects of overeducation." *Economics of Education Review* **1(1)**: pp. 75–86.

Eurydice (2010): "Structures of Education and Training Systems in Europe: Czech Republic." .

Filer, R. K., Š. Jurajda, & J. Plánovský (1999): "Education and wages in the Czech and Slovak Republics during transition." *Labour Economics* **6(4)**: pp. 581–593.

Fleming, C. M. & P. Kler (2014): "Overeducation in the Labour Market: a Meta-analysis." *Economics Analysis and Policy* **44**: pp. 143–155.

Frank, R. H. (1978): "Why women earn less: The theory and estimation of differential overqualification." *American Economic Review* **68**: pp. 360–373.

Freeman, R. (1976): *Overeducated American*. New York: Academic Press.

Gebicka, B. (2010): "College degree supply and occupational allocation of graduates–the case of the czech republic." *CERGE-EI Working Paper* **(407)**: p. 20.

Greene, W. H. (2008): *Econometric analysis*. Granite Hill Publishers.

Groot, W. & H. van den Brink (2000): "Overeducation in the Labour Market: a Meta-analysis." *Economics of Education Review* **19**: pp. 149–158.

Hartog, J. (2000): "Over-education and Earnings: Where are we, where should we go?" *Economics of Education Review* **19**: pp. 131–147.

Heckman, J. J. (1979): "Sample selection bias as a specification error." *Econometrica: Journal of the econometric society* **47(1)**: pp. 153–161.

Heckman, J. J., L. J. Lochner, & P. E. Todd (2003): "Fifty years of Mincer earnings regressions." *NBER Working Paper* **(9732)**.

Karakaya, G., R. Plasman, & F. Rycx (2007): "Overeducation on the Belgian labour market: Evaluation and analysis of the explanatory factors through two types of approaches." *Compare* **37(4)**: pp. 513–532.

Kiker, B. F., M. C. Santos, & M. M. De Oliveira (1997): "Overeducation and undereducation: Evidence for Portugal." *Economics of Education Review* **16(2)**: pp. 111–125.

Korpi, T. & M. Tåhlin (2009): "Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000." *Labour Economics* **16(2)**: pp. 183–193.

Kotýnková, M. (2006): *Trh práce na přelomu tisíciletí*. Oeconomica.

Lepič, M. & J. Koucký (2012): "Kvalifikační potřeby trhu práce." *Národní ústav pro vzdělávání* .

Leuven, E. & H. Oosterbeek (2011): "Overeducation and mismatch in the labor market." *Handbook of the Economics of Education* **4**: pp. 283–326.

Liu, K., K. G. Salvanes, & E. Sorensen (2012): "Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession." *NHH Dept. of Economics Discussion Paper* **(16)**.

Long, J. S. & J. Freese (2006): *Regression models for categorical dependent variables using Stata*. Stata press.

McFadden, D. (1977): *Quantitative methods for analyzing travel behavior of individuals: some recent developments*. Institute of Transportation Studies, University of California.

McGoldrick, K. & J.Robst (1996): "Gender Differences in Overeducation:A test of the Theory of Differential Overqualification." *American Economic Review* **86**: pp. 280–284.

McGuiness, S. (2006): "Overeducation in the Labour Market." *Journal of Economic Surveys* **20(3)**: pp. 387–419.

MEYS (2004): " Zákon č. 561/2004 Sb: O předškolním, základním, středním, vyšším odborném a jiném vzdělávání (Školský zákon) ." .

MEYS (2012): "The Education system in the Czech Republic." .

MEYS (2014): "Indikátory strategie vzdělávací politiky ČR do roku 2020." .

MoLSA (2015): "Přehled o vývoji částek minimální mzdy." .

Morano, C. P. *et al.* (2014): "The determinants of overeducation: Evidence from the Italian labour market." *Investigaciones de Economía de la Educación* **9**: pp. 681–698.

Mysíková, M. (2014): "Educational Mismatch in the Czech Labour Market." *Investigaciones de Economía de la Educación* **9**: pp. 699–714.

Nawata, K. (1994): "Estimation of sample selection bias models by the maximum likelihood estimator and Heckman's two-step estimator." *Economics Letters* **45(1)**: pp. 33–40.

OECD (2007): *International Migration Outlook*. OECD Publishing.

OECD (2014a): *Education at a Glance: OECD Indicators*. OECD Publishing.

OECD (2014b): *Employment outlook*. OECD Publishing.

Ortiz, L. & A. Kucel (2008): "Do fields of study matter for over-education? The cases of Spain and Germany." *International Journal of Comparative Sociology* **49(4-5)**: pp. 305–327.

Powdthavee, N. (2007): "Economics of happiness: A review of literature and applications." *Chulalongkorn Journal of Economics* **19(1)**: pp. 51–73.

Roman, S. (2004): *Introduction to the mathematics of finance: from risk management to options pricing*. Springer Science & Business Media.

Ryška, R. & M. Zelenka (2011): "Charakteristiky vysokoškolského vzdělávání a profesní úspěch." *Orbis scholae* **5(1)**.

Sánchez-Sánchez, N. & S. McGuinness (2013): "Decomposing the impacts of overeducation and overskilling on earnings and job satisfaction: An analysis using REFLEX data." *Education Economics* **23**: pp. 419–432.

Sicherman, N. & O. Galor (1990): "A Theory of Career Mobility." *Journal of Political Economy* **98**: pp. 169–192.

Spence, M. (1973): "Job market signaling." *The quarterly journal of Economics* **87(3)**: pp. 355–374.

Thurrow, L. C. (1975): *Generating Inequality*. New York: Basic Books.

Treiman, D. J. (2009): *Quantitative Data Analysis: Doing Social Research to Test Ideas*, volume 27. John Wiley & Sons.

Tsang, M. C. & H. M. Levin (1985): "The Economics of Overeducation." *Economics of Education Review* **4(2)**: pp. 93–104.

Verdugo, R. R. & N. T. Verdugo (1989): "The impact of surplus schooling on earnings: Some additional findings." *Journal of Human Resources* **24**: pp. 629–643.

Wooldridge, J. (2012): *Introductory econometrics: A modern approach.* Cengage Learning.

Appendix A

Imputation of the Variable Years of Completed Education

Years of completed education were determined according to standard lenght of schooling corresponding to each education level.[1] There were several reforms having impact on years required to complete elementary school. If we restrict our analysis on prime-aged workers, only one reform is relevant and we get 2 categories: 25-35 year-olds who had 8 years of elementary school, for 35-54 year-olds it was 9 years. Because of this reform, years of education required to complete every education level higher than the first stage of the elementary school are computed separately for 25-35 year-olds and 35-54 year-olds. Note that VZD is a variable indicating the highest attained education level of an individual.

- Elementary school, first stage($VZD = 1$) : 5 years
- Elementary school, second stage($VZD = 2$) : 8/9 years (8 years for 25-35 year olds, 9 years for 35-54)
- Secondary school ($VZD = 3$): 10/11 years[2]
- Secondary school with school-leaving exam ($VZD = 4$): 12/13years
- Follow-up courses ($VZD = 5$): 13/14years
- Tertiary professional schools ($VZD = 6$): 15/16years[3]

[1] Source:`http://www.msmt.cz/vzdelavani/skolstvi-v-cr/system-vzdelavani-v-cr`

[2] Lenght of study of secondary schools without school-leaving exam is 1-3 years, for the sake of simplicity, we calculate with 2 years

[3] Some bachelor programs require 4 years to complete, but we neglect this possibility as it is not very common

- Bachelor degree ($VZD = 7$): 15/16 years
- Master/engineer($VZD = 8$): 17/18 years
- Doctoral programs($VZD = 9$): 20/21 years

Appendix B

Small and Hsiao Tests for the Independence of Irrelevant Alternatives

Figure B.1: Small- Hsiao Test, Specification 1

Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

Omitted	lnL(full)	lnL(omit)	chi2	df	P>chi2	evidence
1	-818.032	-813.424	9.217	12	0.684	for Ho
3	-732.581	-737.141	-9.121	12	1.000	for Ho

Figure B.2: Small-Hsiao Test, Specification 2

**** Small-Hsiao tests of IIA assumption (N=5080)

Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

Omitted	lnL(full)	lnL(omit)	chi2	df	P>chi2	evidence
1	-819.278	-817.254	4.049	13	0.991	for Ho
3	-744.323	-752.063	-15.481	13	1.000	for Ho

Figure B.3: Small-Hsiao Test, Specification 3

**** Small-Hsiao tests of IIA assumption (N=5080)

Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

Omitted	lnL(full)	lnL(omit)	chi2	df	P>chi2	evidence
1	-527.679	-524.108	7.142	13	0.895	for Ho
3	-1018.906	-1009.116	19.579	13	0.106	for Ho

Figure B.4: Small-Hsiao Test, Specification 4

**** Small-Hsiao tests of IIA assumption (N=5080)

Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

Omitted	lnL(full)	lnL(omit)	chi2	df	P>chi2	evidence
1	-530.199	-519.724	20.951	13	0.074	for Ho
3	-1006.450	-1001.799	9.301	13	0.750	for Ho

Appendix C

Cross-correlation Tables

Table C.1: Cross-correlation Table for Verdugo-Verdugo Model, Statistical Measure

Variables	OVERmean	UNDERmean	EXP	YRS	NONCZECH	MNGR	PERM	FEM	PAR	PRG	SICK
OVERmean	1.000										
UNDERmean	-0.122	1.000									
EXP	0.015	0.030	1.000								
YRS	0.342	-0.271	-0.101	1.000							
NONCZECH	-0.005	-0.007	-0.015	0.040	1.000						
MNGR	0.107	0.052	0.064	0.228	0.009	1.000					
PERM	0.179	0.174	0.198	0.091	-0.063	0.238	1.000				
FEM	-0.045	-0.025	-0.113	0.048	-0.005	-0.125	-0.062	1.000			
PAR	0.001	-0.049	0.116	0.008	-0.005	0.014	-0.002	0.152	1.000		
PRG	-0.021	-0.019	-0.055	0.182	0.116	0.020	0.004	-0.009	-0.075	1.000	
SICK	-0.050	-0.030	0.027	-0.099	0.015	-0.054	-0.169	0.024	-0.041	-0.023	1.000

Table C.2: Cross-correlation Table for Verdugo-Verdugo Model, Normative Measure

Variables	OVERmean	UNDERmean	EXP	YRS	NONCZECH	MNGR	PERM	FEM	PAR	PRG	SICK
OVERmean	1.000										
UNDERmean	-0.122	1.000									
EXP	0.015	0.030	1.000								
YRS	0.342	-0.271	-0.101	1.000							
NONCZECH	-0.005	-0.007	-0.015	0.040	1.000						
MNGR	0.107	0.052	0.064	0.228	0.009	1.000					
PERM	0.179	0.174	0.198	0.091	-0.063	0.238	1.000				
FEM	-0.045	-0.025	-0.113	0.048	-0.005	-0.125	-0.062	1.000			
PAR	0.001	-0.049	0.116	0.008	-0.005	0.014	-0.002	0.152	1.000		
PRG	-0.021	-0.019	-0.055	0.182	0.116	0.020	0.004	-0.009	-0.075	1.000	
SICK	-0.050	-0.030	0.027	-0.099	0.015	-0.054	-0.169	0.024	-0.041	-0.023	1.000

Table C.3: Cross-correlation Table for ORU Model

Variables	Sr	So	Su	EXP	FEM	PAR	NONCZECH	MNGR	PERM	PRG	SICK
Sr	1.000										
So	-0.602	1.000									
Su	0.474	-0.362	1.000								
EXP	0.020	0.355	-0.012	1.000							
FEM	-0.106	0.125	-0.035	-0.012	1.000						
PAR	0.409	-0.214	0.148	0.067	0.055	1.000					
NONCZECH	-0.330	-0.252	-0.151	-0.484	-0.025	-0.261	1.000				
MNGR	0.356	-0.208	0.171	0.021	-0.092	0.153	-0.102	1.000			
PERM	0.577	-0.571	0.345	0.028	-0.079	0.348	-0.277	0.349	1.000		
PRG	0.037	0.038	0.002	0.022	0.000	-0.038	0.011	0.019	0.008	1.000	
SICK	-0.197	0.255	-0.084	0.223	0.059	-0.069	-0.097	-0.072	-0.165	-0.008	1.000

Table C.4: Cross-correlation Table for Job Satisfaction Model, Statistical Measure

Variables	JS	OVERmean	UNDERmean	LWAGE	FEM	MNGR	PERM	EXP	YRS	HEAD	FIRED	CAR	TICK	PHONE	LANG	TRAN	SERV	SPORT	INS
JS	1.000																		
OVERmean	0.024	1.000																	
UNDERmean	-0.024	-0.182	1.000																
LWAGE	0.110	0.065	-0.089	1.000															
FEM	0.181	-0.030	-0.006	-0.335	1.000														
MNGR	0.078	0.052	0.000	0.373	-0.128	1.000													
PERM	0.041	-0.007	-0.001	0.162	-0.024	0.053	1.000												
EXP	0.063	-0.031	-0.013	-0.007	-0.001	0.032	0.132	1.000											
YRS	0.130	0.398	-0.363	0.424	0.049	0.249	0.003	-0.170	1.000										
HEAD	0.031	0.050	-0.037	0.316	-0.629	0.145	0.010	0.178	0.006	1.000									
FIRED	-0.025	-0.017	0.010	-0.088	0.013	-0.028	-0.187	-0.024	-0.041	0.006	1.000								
CAR	0.019	0.030	-0.005	0.263	-0.136	0.257	0.050	0.011	0.153	0.119	0.004	1.000							
TICK	0.078	0.022	-0.056	0.143	0.063	0.064	0.100	0.023	0.146	-0.016	-0.017	-0.019	1.000						
PHONE	0.046	0.071	-0.037	0.313	-0.133	0.260	0.056	-0.035	0.203	0.130	-0.013	0.387	0.032	1.000					
LANG	0.066	0.041	-0.029	0.170	-0.010	0.090	0.017	-0.057	0.159	0.022	-0.012	0.089	0.091	0.187	1.000				
TRAN	0.049	0.026	-0.027	0.136	-0.054	0.118	-0.002	-0.008	0.076	0.073	-0.010	0.284	0.000	0.223	0.092	1.000			
SERV	-0.016	0.032	-0.004	0.043	-0.001	0.039	0.024	-0.022	-0.007	0.031	0.013	0.076	0.012	0.113	0.023	0.063	1.000		
SPORT	0.070	0.019	-0.034	0.137	0.011	0.060	0.045	-0.001	0.095	0.033	-0.029	0.023	0.119	0.089	0.182	0.038	0.044	1.000	
INS	0.090	0.046	-0.055	0.224	-0.051	0.098	0.110	0.074	0.099	0.083	-0.035	0.081	0.191	0.154	0.130	0.068	0.075	0.145	1.000

Table C.5: Cross-correlation table for job satisfaction model, normative measure

Variables	JS	OVERnorm	UNDERnorm	LWAGE	FEM	MNGR	PERM	EXP	YRS	HEAD	FIRED	CAR	TICK	PHONE	LANG	TRAN	SERV	SPORT	INS
JS	1.000																		
OVERnorm	-0.052	1.000																	
UNDERnorm	0.078	-0.138	1.000																
LWAGE	0.110	-0.165	0.114	1.000															
FEM	0.181	0.083	0.047	-0.335	1.000														
MNGR	0.078	-0.079	0.167	0.373	-0.128	1.000													
PERM	0.041	-0.009	0.062	0.162	-0.024	0.053	1.000												
EXP	0.063	-0.030	0.093	-0.007	-0.001	0.032	0.132	1.000											
YRS	0.130	0.029	-0.149	0.424	0.049	0.249	0.003	-0.170	1.000										
HEAD	0.031	-0.081	0.002	0.316	-0.629	0.145	0.010	0.178	0.006	1.000									
FIRED	-0.025	0.001	-0.010	-0.088	0.013	-0.028	-0.187	-0.024	-0.041	0.006	1.000								
CAR	0.019	-0.043	0.109	0.263	-0.136	0.257	0.050	0.011	0.153	0.119	0.004	1.000							
TICK	0.078	-0.016	0.044	0.143	0.063	0.064	0.100	0.023	0.146	-0.016	-0.017	-0.019	1.000						
PHONE	0.046	-0.052	0.100	0.313	-0.133	0.260	0.056	-0.035	0.203	0.130	-0.013	0.387	0.032	1.000					
LANG	0.066	-0.033	0.037	0.170	-0.010	0.090	0.017	-0.057	0.159	0.022	-0.012	0.089	0.091	0.187	1.000				
TRAN	0.049	-0.014	0.039	0.136	-0.054	0.118	-0.002	-0.008	0.076	0.073	-0.010	0.284	0.000	0.223	0.092	1.000			
SERV	-0.016	-0.019	-0.002	0.043	-0.001	0.039	0.024	-0.022	-0.007	0.031	0.013	0.076	0.012	0.113	0.023	0.063	1.000		
SPORT	0.070	0.011	0.030	0.137	0.011	0.060	0.045	-0.001	0.095	0.033	-0.029	0.023	0.119	0.089	0.182	0.038	0.044	1.000	
INS	0.090	-0.023	0.039	0.224	-0.051	0.098	0.110	0.074	0.099	0.083	-0.035	0.081	0.191	0.154	0.130	0.068	0.075	0.145	1.000