Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Fiscal Gains from Extra Investment into Roma Education in the Czech Republic

Author: **Bc. Marie Kubíková**

Supervisor: **Petr Janský, 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, and the thesis has not been used to obtain a different or the same degree.

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

Prague, June 5, 2015

_____

Signature

Acknowledgements

Abstract

The thesis aims at quantifying the potential fiscal gains of an extra investment into Roma education. From the budgetary point of view, it would pay off to invest extra money into Roma children, if it increased future net contributions to the budget. The extra investment into one child could take values up to the average discounted value of future net benefits per person. The net benefits represent the increased net contributions due to a higher level of education. The contribution of the thesis is the estimation of different net contributions of Roma with various educational levels to the national budget in order to encourage the investment into Roma children. Secondly, we estimated the yearly loss of the national budget due to lower education among the Roma. And thirdly, we conducted a survey among the clients of the endowment fund Verda to find the relationship between their education, employment status and other variables.

Abstrakt

Cílem této práce je vyčíslit potenciální fiskální příjmy z dodatečných investic do vzdělání Romů. Z pohledu státního rozpočtu by se vyplatilo investovat více do romských dětí, kdyby to zvýšilo budoucí čisté příjmy státního rozpočtu. Investice do jednoho dítěte by mohla maximálně dosahovat diskontované hodnoty budoucích čistých benefitů z vyššího vzdělání státního rozpočtu za jednu osobu. Čisté benefity reprezentují vyšší příjmy státního rozpočtu z jednoho Roma způsobené vyšším vzděláním. Přínosem této práce je odhad různých čistých celoživotních příjmů státního rozpočtu z jednoho Roma podle jeho vzdělání. Tyto odhady by měli podpořit investice do vzdělání romských dětí. Dále odhadujeme o kolik státní rozpočet ročně přichází z důvodu nižší vzdělanosti Romů. Nakonec jsme provedli výzkum mezi stipendisty nadačního fondu Verda, kde jsme hledali vztahy mezi jejich vzděláním, zaměstnaností a dalšími proměnnými.

Contents

List of Tables

List of Figures

Acronyms

PIT Personal Income Tax

VAT Value Added Tax

Master Thesis Proposal

Author:	Bc. Marie Kubíková	Supervisor:	Petr Janský, Ph.D.
E-mail:	marie.kubikova1@gmail.com	E-mail:	jansky@fsv.cuni.cz
Phone:	737675142	Phone:	775022260
Specialization:	Finance, Financial Markets and Banking	Defense Planned:	January 2015

Proposed Topic:

Fiscal gains from extra investment into Roma education in the Czech Republic

Motivation:

Supporting education is seen as one possible solution to existing problems with the Roma minority; maybe the most important one as it goes straight to one of their sources rather than solving the consequences. Beside social reasons there might be economic ones. My interest is based on existing literature which shows that investment into disadvantaged children simply pays off (e.g. Heckman (2006) and (2010), Milcher and Zigová (2005)).

If we look at education as an investment, then it would be worth investing up to the amount that a person will contribute in the future. If we invest less, then the person will be beneficial for the national budget and if we invest more, it will not be worth from the economic point of view. I will use the threshold of completed secondary education with maturity examination ("maturita" in Czech), as it is a basic requirement for the majority of jobs and also opens the possibility to enter university. The proportion of people who completed secondary education is much lower among the Roma minority than is the average in Czech Republic. Therefore, an investment that would help an average Roma person complete secondary education will improve his or her situation on the labour market and consequently increase his or her wage and taxes paid and decrease the transfers obtained. Inspired

by Kertesi and Kézdi (2006) I would like to estimate the net budgetary benefit of completing secondary school with maturity examination. In other words I will estimate the expected extra contribution to the national budget of an educated Roma person when compared to an uneducated one. Supposing there is an investment that would ensure passing the maturity exam, then, in terms of national budget, we could invest up to that amount.

Naturally, a question arises, how net benefits differ among the Roma population. It is probable, that they cannot be the same for everybody (see for example Balestra and Backes-Gellner, 2013). Surely there are more talented students who are also more likely to earn more in the future and therefore contribute more to the national budget. I would like to estimate the distribution of net benefits with respect to the national budget.

Furthermore, I would like to look closely at the endowment fund Verda (http://verda.me/), which is a real example of the theory above. Verda has been giving scholarships to selected Roma students since 2001 and has already supported more than 400 of them. In cooperation with Verda I will contact the supported students and evaluate the impact of the scholarship as an extra investment. Main questions of interest will be their school completion and whether they are employed or not. I would like to compare their real experience to my theoretical findings. The main issue will be, whether a completed secondary education really increases the probability of being employed, which is a basic assumption of all theoretical papers about returns to education.

Hypotheses:

1. Hypothesis #1: Investing extra money into Roma education pays off in fiscal terms.

2. Hypothesis #2: Investment into Roma education yields varying fiscal benefits.

3. Hypothesis #3: Completed secondary education has increased the probability of being employed among clients of the endowment fund Verda.

Methodology:

I would like to estimate the difference between an educated and an uneducated Roma person from the view of the national budget. I will employ the methodology

of Kertesi and Kézdi (2006), which focuses on several channels through which every person either contributes to the national budget or obtains transfers from it. These are "personal income tax on income earned from registered full-time employment, social security contributions paid by employers and employees on earned income, unemployment benefits, means-tested welfare benefits, earning from public employment projects, value added and excise tax on consumption, and incarceration costs." (Kertesi and Kézdi, 2006). The flows in each of these channels will be discounted to the age of 4 and summed together. Due to unavailability of detailed Roma data and also due to comparability with the results of Kertesi and Kézdi (2006) I will use general data for Czech population from the Czech Statistical Office adjusted for Roma differences.

To answer the second hypothesis I will assume that if there were no socio-economical barriers, the distribution of education and occupations would be the same among Roma as is the overall distribution in the Czech Republic. Based on that I will compute the expected net benefits to the national budget according to this distribution. I will use the same source of data and logic of computation as when answering the first hypothesis.

To evaluate the activity of Verda I will conduct a survey among its former clients. Its first part will be creating a questionnaire and distributing it. The exact form and distribution of questionnaires will be consulted with Verda. The main questions of interest will be the employment status, occupation, the highest level of education attained and the effect of the extra scholarship. The second part will be evaluating the questionnaires. To answer the third hypothesis I will compare former students of Verda with respect to their employment status. The main issue here will be ensuring sufficiently large sample (which is already bounded by the total amount of clients) and controlling for impact of other factors.

Expected Contribution:

I will estimate the exact additional sum of money that an average Roma person would contribute to the national budget if he or she would complete secondary school with maturity examination compared to the one who did not. This might be quite an important figure when developing both Roma and education policies. Furthermore, it can support the effort of non-profit organizations as for example Verda, and possibly serve as an argument when asking for grants or attracting sponsors.

Furthermore, I will also estimate the expected distribution of net benefits among the Roma population, which might also be of interest for both policy makers and non-profit organizations focused on the Roma minority.

Finally, I will evaluate the performance of Verda, that takes practical steps to increase the probability of completing secondary school, and is a real example of implementing the theory by Kertesi and Kézdi (2006). I will also compute what is the contribution of Verda clients to the national budget.

Outline:

1. Introduction

2. Literature background

3. Description of data used (both datasets, the one with general data and the one obtained from the survey)

4. Methodology (of testing the three hypotheses)

5. Results

6. Conclusion

Core Bibliography:

Balestra, S., Backes-Gellner, U. (2013): „Heterogeneous Returns to Education Over the Wage Distribution: Who Profits the Most?", Leading House Working Paper No. 91

Chobanov, D., Kostadinova, S., Angelov, G. (2007): „Expected Long-Term Budgetary Benefits to Roma Education in Bulgaria", Institute for Market Economics

Heckman, J. J. (2006): "Investing in Disadvantaged Young Children is an Economically Efficient Policy", presented at the Committee for Economic Development/The Pew Charitable Trusts/PNC Financial Services Group Forum on "Building the Economic Case for Investments in Preschool"

Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A., Yavitz, A.: „The Rate of Return to the Highscope Perry Preschool Program", Journal of Public Economics,

2010, 94(1-2), pp. 114-28

Kertesi, G., Kézdi, G. (2006): „Expected Long-Term Budgetary Benefits to Roma Education in Hungary", Roma Education Fund

Milcher, S., Zigová, K. (2005): „Evidence of Returns to Education Among Roma in Central and Eastern Europe and Their Policy Implications", Managing Global Transitions, Volume 3, Number 1

1 Introduction

The Roma in the Czech Republic, as well as in other European countries, have on average lower income, higher probability of unemployment, lower educational level, the Roma families are more dependent on welfare benefits than other families, their children have lower school attendance at compulsory school age (FRA-UNDP, 2012). The Roma are more often incarcerated (Bakalář, 2004). Even though the Roma minority lives in Europe for centuries, they did not assimilate much to the majority society. At the same time, the Roma population is on average younger and it is almost certain that the proportion of the Roma in the Czech Republic will continue growing (UNDP, 2002). This also means that the problems between the Roma and the majority will most probably intensify.

All these facts increase a generally bad feeling against the Roma. This has many consequences. There have been ethnically motivated attacks. Due to prejudices the Roma have much more difficult conditions to succeed at work. It is a sort of vicious circle, from which it is difficult to escape. The children growing up in Roma families are influenced by their background. If their parents are uneducated and unemployed, they are likely to repeat what they learned when they became adults. The family often does not support education, but the less they go to school, the less the school can influence them. One possible way how to break this circle is promoting education. At school, you can learn different patterns of behaviour, you can gain motivation and develop abilities for further studies and for work (Milcher and Zigová, 2005). Many non-profit organisations (and to some extent also the state schools) focus on this. However, they do not have sufficient financial resources and competences. Our goal is to prove that it would pay off to invest more into the education of the Roma.

The main purpose of the thesis is to quantify potential gains from investment into Roma education. We will first replicate a study by Kertesi and Kézdi (2006), which estimates the gains from education from the individual's point of view: What additional money would one person bring if he or she had better education. Secondly, we will extend this computation to be more detailed. And thirdly, we will compute the total yearly loss of the national budget due to lower education among the Roma. This will be computed as a gap between the current situation and a potential situation in which the distribution of highest achieved education among Roma would be the same as it is among all Czech citizens. The fourth part will focus on clients of the endowment fund Verda. We conducted a small survey among them during the autumn

2014. The main question of interest was the impact of education on their employment. We wanted to find support for one of our basic assumptions used in the preceding computation, which is, that a higher level of education means a higher probability of employment among the Roma.

We do not focus on how to increase the Roma education but we believe together with many other people (see for example Milcher and Zigová (2005) that it is one of the best possible ways how to improve the social situation of the Roma minority and we want to support it from the financial point of view: That it simply pays off. It might seem that our motivation is to get as much money as possible from the Roma. This is definitely not true. We only present a very narrow view of the problem. There are many other dimensions, such as humanity or respect for another culture, that are not in our field. We keep these in mind, even though we do not focus on them in the thesis.

After this introduction, a literature review follows in chapter 2, which focuses generally on education of disadvantaged children, on Roma specifics and on education of the Roma in the Czech Republic. Chapter 3 describes the multiple data sources we use in our analyses. Chapter 4 then explains the methodology we use. Chapter 5 summarizes our findings and chapter 6 presents the conclusion of our thesis.

2   Literature review

2.1 Short historical background of the Roma in Europe

The Roma are the largest ethnic minority in Europe and basically in all the countries where they live they belong to the poorest people (FRA-UNDP, 2012). They came from India about 700 years ago and spread across many countries and since then they have mostly lived outside the rest of society with a nomadic style of life. The Roma never owned land or real estate, they worked for somebody or sold their products or services (Kertesi and Kézdi, 2010). The agricultural based nations built their institutions and norms differently from those of the Roma and here is one of the possible roots of the Roma focusing more on today's consumption rather than on accumulation of property (UNDP, 2002). During the Second World War they faced Roma Holocaust and during the Communist regime they were subject to a planned assimilation. The Roma were resettled into locations inhabited by the majority population, they had to be employed and to send their children to schools (UNDP, 2002). The economic transformation after the end of the Communist regime affected the Roma unfavourably. They had often no property and no legal rights to live in their houses and therefore many of them migrated to rural areas and settled in ghettos (UNDP, 2002). The decline in Roma employment in these times can be among others also explained by the decrease of demand for low skilled labour (Kertesi and Kézdi, 2010). Other possible reasons for Roma unemployment are described later in this section.

2.2 Return to education

The main objective of the thesis is to find out to what extent it is profitable for the national budget of the Czech Republic to invest into Roma education. One of the basic assumptions used is that education increases earnings and the probability of being employed. Otherwise, it would not make sense to support education from the fiscal point of view. Generally, the relationship between education and earnings is well described and estimated by many researchers. Hundreds of studies showed that more educated individuals have better jobs, higher earnings and are less likely to be unemployed. But as Card (1999) denotes, it is more difficult to detect whether it is only a correlation or whether education has a causal effect on earnings and it is also

very likely that the return to education is not equal in the population, but differs with many determinants, such as family background or ability.

One of very popular models is the human capital earnings function introduced by Jacob Mincer where the log of earnings is explained by education and quadratic experience term. This model is often further enlarged with other variables such as race, gender, ethnicity, number of children, location and many others in order to explain more of the variation in earnings (Polachek 2007). Psacharopoulos and Patrinos (2004) compute the mincerian return to education coefficient for almost a hundred countries all around the world and during many years and bring evidence of a positive impact of education on earnings in all the countries involved. The basic version of mincerian earnings function supposes that every year of education has the same return. Opposed to that, many authors describe a so called "sheepskin effect", which states that not the number of years of education, but the credential earned is what matters (see for example Flores-Lagunes and Light 2007)[1].

Education evidently does not influence only how much one is earning, but also the risk of becoming unemployed and the chance of re-employment. Evidence from US male workers in the 1970s and 1980s shows that higher education decreases the risk of becoming unemployed and with lower significance also decreases the duration of unemployment. Interesting are the sources of the observed patterns. Lower turnover of the more educated workers is found to be caused by their greater attachment to the firm due to more on-the-job training. Shorter duration of unemployment of higher educated workers is caused by lower relative costs of searching for a new job when employed and higher ability to search and process information about new possible jobs. Moreover, the firms search more intensively for more skilled employees. (Mincer 1991)

Riddel and Song (2011) highlight the adaptability to changes and shocks in the labour market as another important factor through which education decreases the probability of becoming unemployed and increases re-employment chances. Rapid change is nowadays a characteristic feature of the labour market due to technological progress and globalization and more educated people are capable to better adjust to new circumstances. Riddel and Song (2011) further present clear evidence from the US market about the causal effect of education on re-unemployment.

---

1   We use this logic in our computations, since we sort education according to the highest completed level.

2.3 Return to education of disadvantaged children

Even more interesting for us is the research concerning disadvantaged children and the return to their education. It is not natural to suppose that investment into disadvantaged children pays off[2]. James J. Heckman in many of his papers proves that it does and recognizes three main channels through which the society can benefit from it: through the reduction of crime, the increased workforce productivity and the lower teenage pregnancy (Heckman 2006). For better understanding of the papers by James J. Heckman it is important to note that he identifies two types of skills: cognitive and non cognitive. He underlines the significance of non cognitive skills (e.g. motivation, attention, health, self-confidence) and remarks that the majority of achievement test scores are focused only on the cognitive skills and therefore omit a very important determinant of school and life success. If a society wants to support the education of disadvantaged children it should take into account both types of skills. Another finding of James J. Heckman is that the returns to investment into human capital are decreasing through the life of an individual and are more efficient for the pre-school children. Moreover, any later investment is more efficient when it is applied to children with preceding early-childhood intervention. (Heckman 2008)

A great and often cited evidence comes from the High/Scope Perry Preschool Program, which was an experiment starting in the 1960s in the United States. A group of disadvantaged children was randomly chosen to receive a preschool programme and then followed together with a control group until the age of 40. The programme group had significantly better educational level (the difference was especially large for women), better results in various tests (i.e. cognitive skills), higher earnings and also higher employment rate and better results in many other indicators, particularly increased taxes, lower crime and welfare benefits. The discounted economic return per participant was about sixteen times the per participant cost of the preschool investment (Schweinhart et al. 2005). Another experiment with a similar design and also similar results was the Abecedarian Project that took place ten years later (Campbell et al. 2002).

---

2 For example Herrnstein and Murray (1994) claim that the main factor of income and employment is intelligence (not education), which is genetically determined. As a consequence, higher education would not help much to those who are now socially excluded. This paper was very controversial, however it had its supporters.

2.4 Education among the Roma

Roma children in the Czech Republic are to a large extent a typical example of disadvantaged children mentioned above. However besides the social disadvantage, they seem to face an ethnic barrier (FRA-UNDP, 2012). What part of this barrier comes from discrimination, different family background or other sources is an important research question.

Kertesi and Kézdi (2014) have decomposed the gap in test scores between Roma and non-Roma students in Hungary. They have used a detailed family background index (e.g. information about the education and employment of parents, family income, housing and location) and have found that the variation in test scores can be largely explained by these variables. Their sample includes Roma and non-Roma students from the same schools and the same social background so that they can control for this. They have concluded that the achievement gap is to a large extent caused by social differences and that ethnicity itself does not play an important role. Another their finding is that Roma tend to attend schools where the rest of non-Roma students have low test scores. In other worlds the Roma are over-represented in low quality schools. This is a similar problem as described by O'Higgins and Brüggemann (2013) that in the Czech Republic an excessive amount of Roma children attend special schools for mentally disabled without being mentally disabled. This might be a hidden means of discrimination because attending a special school heavily decreases the probability of higher education and consequently employment. The discrimination might not be limited only to refusing an employee because of ethnicity.

Milcher and Zigová (2005) computed returns to education for Roma in several European countries using an extended mincerian model and found positive returns, which are higher for the more advanced economies. One of their observation is that Roma parents are often not motivated to support the education of their children because they are themselves usually also not educated, they live in less developed areas with a high unemployment and they are used to receiving social contributions. The low motivation of parents cause that their children have lower chances to succeed both at school and at the labour market. The way out from this vicious circle could be educational intervention at early age. (Milcher and Zigová, 2005) Nevertheless, the Roma have a much lower return to education than the majority as found by Trentini (2014) or O'Higgins (2009). Among possible reasons is a lower quality of their education and the labour market discrimination (Trentini, 2014). Furthermore, evidence that education is negatively correlated with crime not surprisingly exist (see for example Machin et el., 2010).

2.5 Employment among the Roma

Another Roma issue is the unemployment. The fact that employment is generally lower among the Roma population than the national average can be either caused mainly by lower education (and therefore increased education would increase employment) or low education and low employment might have another common source (e.g. discrimination, insufficient motivation) and therefore solely increasing education would not help to decrease the unemployment. An important remark to evaluation of Roma unemployment figures is given by O'Higgins and Ivanov (2006). They say that the information about unemployed Roma is often overestimated, as many of the Roma although formally unemployed have some income from the shadow economy.

Kertesi and Kézdi (2010) focus on employment of Roma in Hungary that after the end of the Communist regime decreased dramatically and remained low. They estimated that the post-communist employment gap between Roma and non-Roma ranges between 36 and 40 percentage points. The main reason found (explaining about one third of the employment gap) is lower education and geographic location, the rest remains unexplained by the used variables and might include unobserved differences in skills or labour market discrimination. The authors have found indirect evidence for the latter.

Milcher (2011) decomposes differences in earnings between Roma and non-Roma in several European countries into income related characteristics (education, experience etc.) and a discrimination effect (differences in returns to the income related characteristics) and finds that a large share is due to human capital differences and only a very small part might be explained by labour market discrimination (and only in some countries).

2.6 Costs and benefits of investment into Roma education

The majority of literature agrees that the worse education and employment of Roma is mainly due to their social background and not due to their ethnicity. However, the Roma tend not to assimilate with the majority and therefore live in a continuous cycle, as Heckman (2008) writes: "The accident of birth is a major source of inequality". If there was a possibility how to help the Roma children to get higher education, it would surely be worth as illustrated in the preceding paragraphs. Kertesi and Kézdi (2006) analyse the costs and benefits of Roma education in Hungary to find the exact additional amount that an educated person would bring to the national budget and consequently the maximum amount of investment that could

be made so that it would be profitable from the fiscal point of view. They found that on average an educated Roma brings 70 thousands Euro more than an uneducated one during his lifetime discounted to the age of 4.Their study was followed by Chobanov et al. (2007) with estimated 82 thousands Euro for Bulgaria. In our first hypothesis we will try to replicate their computations using Czech data. A very similar study by the World Bank (2010) also compares the benefits of a completed secondary education to that of elementary education for Roma in several countries including the Czech Republic. It estimates that a secondary educated Roma earns on average 110% more than an elementary educated one. This report further estimates that the total yearly costs of Roma exclusion from the labour market in the Czech Republic are 260 million Euro. This figure is interesting for us as it partly corresponds to our estimate of yearly loss due to lower education of the Roma. The report by the World Bank however used different computation and more importantly, they also used different data (data about Roma living in marginalized localities) (World Bank, 2010).

2.7 The summary of Roma education in the Czech Republic

The Roma have a lower education than the Czech average is, as several surveys show (see for example the results of the Czech Census 2011 or the Roma Data from the United Nations Development Program). Our whole thesis is motivated by this fact, but in this section we are interested in observed patterns of Roma education and also initiatives and possibilities for its improvement that already exists in the Czech Republic.

2.7.1 High proportions of the Roma in special schools

One easily identifiable problem of the Roma in Czech schools is that they more often than their non-Roma schoolmates end up in special schools. It became a political issue when 18 Roma students complained to the European Court of Human Rights in 2000 for being assigned to special schools. According to them this happened because of their ethnicity and not because they had lower mental abilities. The case "D.H. and others versus the Czech Republic" (named after one of the students, Denisa Holubová) was finally decided in 2007 when the court agreed that the students were discriminated. The court considered information from schools in Ostrava that 56% of the students in special schools were Roma and that the Roma students made only 2,26% in elementary schools in this region. This meant that a Roma child was 27 times more likely to attend special school than a non-Roma one (The European Court of Human Rights, 2007).

The latest (and with respect to its extent also the most precise) inquiry by Czech School Inspection[3] in September 2013 found that while the percentage or Roma students in elementary schools was on average 10%, in special schools it was 28,2% (Czech School Inspection 2013). Evidently this inquiry had better results than the one that only questioned the schools in Ostrava, which only shows that in regions where there is a larger proportion of the Roma, the problem of placing their children into special schools is deepened. The research of the Public Defender of Rights from 2012 also supported these numbers, as it found that the proportion of Roma children in special schools was 32%-35% according to the methodology used, while the percentage of the Roma in the population was only 1,4%-2,8% (depending on the number of Roma being from 150 thousand to 300 thousand) (Public Defender of Rights, 2012).

2.7.2 Non-governmental activities

There is no doubt that in many cases the student's background plays an important role. If the parents do not motivate and support the student, it is easier for him or her to end up in special schools. Closely linked to it is a low awareness about special schools and their consequences for life of Roma parents. For example in the case "D.H. and others versus the Czech Republic", the parents of all the 18 students agreed with the placement of their children into special schools or even in a few cases asked for it (The European Court of Human Rights, 2007).

Very influential can be non-governmental activities. For example, since 2013 a non-profit organisation Slovo 21 has organised an initiative "Mum, Dad, I want to go to school"[4] consisting of workshops in many cities, trying to improve the parent's consciousness about the importance of attending school but evading special schools. In the Czech Republic and especially in some regions there is lack of capacity of nursery schools, 24% of children from the age of 3 to the minimum compulsory age school, that is up to 6, do not attend any formal child care, which is one of the highest percentages in Europe[5]. About 52% of the children from socially disadvantaged families do not attend pre-school education at all (Czech School Inspection, 2010). We should therefore mention projects that are focused on the promotion of nursery schooling, for example "Let's go to nursery school[6]", organised by 9 non-profit organizations (among them People in Need) that emphasizes the importance of nursery schools for further school success. The attendance of nursery schools basically equalizes the starting positions of the children in the first class of

---

3   In Czech: Česká školní inspekce

4   In Czech: Mami, tati, já chci do školy

5   Eurostat: Formal childcare by age group and duration (2012)

6   In Czech: Pojďte do školky

elementary school. It can help the child to get used to the school environment earlier than at the beginning of the elementary school, in some cases it also can help the child to learn Czech. Some organizations also offer a sort of a club for pre-school children that do not attend a state nursery school to substitute it (for example Cheiront T, People in Need). When children come from a disadvantaged background, to send them directly to school might be inadequate. Some non-profit organisations such as People in Need, Slovo 21, Roma centre Drom in Brno and others therefore offer free individual or group remedial classes, usually led by volunteers.

2.7.3 Governmental activities

The Czech Republic as well as 6 other European countries formed the "Decade of Roma Inclusion 2005 – 2015". It is a political commitment to eliminate the discrimination of Roma and to lower the gap between the Roma and non-Roma in four areas: housing, employment, health and education.

We are interested in concrete commitments concerning education, such as financial support of schools with inclusive education, which means not dividing children with special needs, but on the other hand keeping all the children in one class with possible use of teaching assistants and then financing their wages or financial support of socially excluded students[7]. One part of the Decade is the Roma educational fund which apart from scholarships mentioned later financially supports policies or programmes that improve education of the Roma.

The Czech Ministry of Education Youth and Sports supports socially disadvantaged Roma students through the programme "The Support of socially disadvantaged Roma students of secondary and short-cycle tertiary education"[8]. The students are selected by their schools that apply for the subsidy, and, if approved, also obtain the money. The amount ranges from 4000 to 8000 CZK per one school term per student, the higher level of education means the higher level of the scholarship. During 2013, 1 280 students got the support which amounted to 5 607 000 CZK in total (The office of the Government of the Czech Republic, 2014). The subsidy is to be used for covering costs of school cafeteria, school accommodation, expenses connected to travelling to school, school supplies and activities of the school (The Ministry of Education Youth and Sports in the Czech Republic, 2014).

---

7 More   information   available   at: http://www.vlada.cz/cz/ppov/zalezitosti-romske-komunity/dokumenty/informace-o-naplnovani-dekady-romske-inkluze-20052015-v-roce-2012-113606/ (accessed 4.11. 2014)

8 In Czech: Podpora sociálně znevýhodněných romských žáků středních a studentů vyšších odborných škol

2.7.4 Study grants and scholarships

Roma students also have the possibility to apply for study grants and scholarships. Here they have to make an effort, which means they ave to be motivated already before applying for the scholarship. On the other hand this can be an advantage, as it selects those who are more likely to successfully finish the school. First we should mention the Roma Educational Fund that was created as a part of the "Decade of Roma Inclusion" in 2005. It offers scholarships to about 700 tertiary students yearly from 16 countries in middle and east Europe, among others also the Czech Republic. The scholarships to Czech students are provided mainly through the Roma Memorial University Scholarship Programme. The amount is 80 EUR per month plus up to 1200 EUR per one school year to cover school fee expenses. The students are chosen with respect to their former school results, motivation, an essay about some Roma topic and references about him or her. Last year 41 Czech students obtained the scholarship and in the preceding years the numbers were similar. Another programme that Czech Roma students can use is the Roma International Scholar Programme, which provides a partial support for students studying outside their home country. In the last two years, only 3 Czech students obtained this scholarship.

Another fund we want to mention is the endowment fund Verda, which was founded in 2001 and has given scholarships to about 500 secondary and tertiary students since. It is a Czech fund financed by private donors, scholarships usually range between 8000 and 13000 CZK per one school term. The amount given is dependent on the educational level and also on previous school results. The better are the results and the higher is the educational level, the higher is the scholarship. The students are chosen by their former school results and entrance interview. The scholarships are paid twice a year, always after the study period, because its successful completion is required. In cooperation with this fund we have conducted the survey which is the source of data for testing the third hypothesis.

3   Data description

The relevance of the results of the first two hypotheses relies heavily on the choice of suitable data and their adjustment so that they capture the reality as well as possible. The first two hypotheses require many variables that could be mostly derived from publicly available data sources. On the other hand, the third hypothesis is tested on data that we collected by ourselves. In the following sections the sources and derivation of the data used are outlined.

3.1   Data for the first hypothesis

The very basic framework for our data is that we use only information about the age from 16 to 65, the data for individuals under 16 are mostly not available and the data for those older than 65 have nearly no impact due to discounting (this is further explained in chapter Methodology). If the data are reported for age groups (e.g. five years groups in case of the incarcerated people) we assume that the distribution of the given variable is equal within each group. Then, consistently with Kertesi and Kézdi (2006) we use 5 educational categories in our analysis:

0: no education

A: elementary education

B: secondary school without maturity exam

C: secondary school with maturity exam

D: tertiary education

As the educational categories in the datasets we use have not been unified, we had to transform the different ones into our 5 categories (see Appendix D).

For each variable we are interested in the average value and its probability in a given age and educational category, for example: what is the average income tax of a 30 years old person with elementary education and with what probability does this representative person pay it (is employed). The first hypothesis also requires some detailed data on Roma population, not available to us, as they were not available to Kertesi and Kézdi (2006). In order to estimate them, we will employ a similar as Kertesi and Kézdi (2006), which at the same time increase their mutual

comparability. The detailed procedure is described in chapter Methodology, the data used are described below in section Roma data.

The following paragraphs describe all the data sources used. First we will describe data on national level, when three major data sources are used: the Population and Housing Census 2011, TAXBEN model based on The Statistics on Income and Living Conditions (SILC) and The Household Budget Survey.

3.1.1 Population and Housing Census 2011

The Census was organized by The Czech Statistical Office and it is the largest statistical survey in the country. The respondents were obliged to answer and to answer truly[9] all the obligatory questions. We use the Census to get the most general data about education, age and economic activity. The economically active population is further divided into employed and unemployed.

3.1.2 TAXBEN model using SILC database

We use the TAXBEN model recently developed by Libor Dušek, Klára Kalíšková and Daniel Münich which is based on SILC 2011 and Czech laws about taxes and social benefits from 2013. SILC 2011 is a representative sample of 8 866 households that include 20 629 individuals, collected by the Czech Statistical Office as a part of the EU-SILC database.

The TAXBEN model among other things simulates taxes and welfare benefits on individual level and therefore gives us the data about income taxes (the actually paid income taxes, after accounting for tax deductible items), social and health insurance contributions (both paid by the employee and the employer), unemployment benefits, children benefits, housing benefits, subsistence benefit[10] and supplementary housing benefit[11] by age and education. For more details about the TAXBEN model see Dušek et al. (2013a) or Dušek et al. (2013b).

3.1.3 The Household Budget Survey

Data on consumption taxes and VAT were taken from The Household Budget Survey, a representative sample of 3000 households collected by The Czech Statistical Office. This survey contains data on consumption at the household level. With the kind permission of Petr Janský we use the household data on taxes that he computed for of one of his papers (Janský, 2014).

Each household in the dataset has its head, then his or her partner (called wife) and then there might be also another adults. In order to assign the taxes to individuals,

---

9   Under a penalty of 10,000 CZK

10 In Czech: příspěvek na životbytí

11 In Czech: doplatek na bydlení

we first divided the amount of taxes paid by the amount of adults in the household. This assumes that only the adults decide about the consumption of the household, not the children (Kertesi and Kézdi 2006). The number of children appears here in increased consumption (more children, more consumption). Then the corresponding part of the taxes paid is assigned to the two individuals for which we had information about age and education (these are the head of household and his/her partner). And using these data we compute the average consumption tax and VAT paid by individuals with respect to age and education.

3.1.4 Other sources

Besides these major sources described above, we have used information from the Prison Service of the Czech Republic (VSČR, 2013) that includes statistics about the incarcerated by age and education, but not both of them. We have therefore simulated the multivariate distribution of incarcerated by first assigning the population probabilities to educational categories and within each educational category further according to the overall age distribution of prisoners. Moreover, as the age is reported for five years age groups, we assume equal distribution within them. The average daily costs of the Prison Service of the Czech Republic per one incarcerated person are reported as 1319 CZK (VSČR, 2013). We have multiplied this figure by 365 to get the average yearly costs.

Secondly, in order to estimate the yearly costs of secondary and tertiary education per one student, we have only focused on the variable part which the state contributes to schools according to the number of their students, as explained in chapter Methodology. In case of secondary schools, there are so called norms per students that differ by age. These norms define the sum that the state gives to its regions, per one student[12]. The regions then further redistribute the funding to their schools. For our purposes we will take the 2014 norm for students 15-18 years old that covers the secondary school students which is 58 313 CZK according to the web page[13] of Ministry of Education Youth and Sports.

In case of tertiary education, funding to schools is allocated according to several indicators. The one, which is variable with respect to the number of students is the so called "Indicator A" and is computed as the number of normative students times basic norm per student, which is 26 323[14] CZK per year in 2014 (MŠMT,

---

12 http://www.msmt.cz/vzdelavani/skolstvi-v-cr/ekonomika-skolstvi/principy-rozpisu-rozpoctu-a-rozpis-rozpoctu-primych-vydaju-5 (accessed 29. 8. 2014)

13 http://www.msmt.cz/vzdelavani/skolstvi-v-cr/ekonomika-skolstvi/republikove-normativy-skol-a-skolskych-zarizeni-zrizovanych-5 (accessed 26.6.2014)

14 Taken from the ministry web page http://www.msmt.cz/vzdelavani/vysoke-skolstvi/rozpis-rozpoctu-vysokych-skol-na-rok-2014 (accessed 26.6. 2014)

2014). The amount of normative students is a weighted measure with respect to the demandingness of each school but for our purposes we will take it simply as cost per student. We do not use the funding per student as it differs for each school. The use of normative students is furthermore consistent with the logic of lower bound as the number per student is for the majority of schools higher than the norm per normative student. The norms per student mentioned above are for public schools only, which are the most often type of school in the Czech Republic and therefore we omit the private schools as irrelevant.

3.1.5 Roma data

The population census mentioned above included among others an optional open question about nationality. It was possible to indicate one or two nationalities. 5 135 identified themselves only as Roma, 7 818 respondents reported two nationalities, one of them being Roma. However, it is obvious, that the total of 12 953 respondents is a completely unrealistic figure when compared to the official estimate of 150-300 thousand Roma in the Czech Republic (The office of the Government of the Czech Republic, 2013). Another evidence on the number of Roma came from the Czech censuses in 1970 and 1980 (which were the only ones when the Roma nationality was recorded by the census inspectors and therefore regarded as the most reliable census results concerning the Roma nationality), which report 60 279 and 88 587 Roma respondents respectively (Socioklub, 1999). As the Roma population is growing since then, we can assume that 150-300 thousand is a reliable estimate.

Although the total number of Roma according to the Census is surely unreliable, we will use its probability distribution of Roma education, as it hopefully represents the reality better than using the national distribution. Furthermore, comparison with other source of data, Roma Data from United Nations Development Program (UNDP) for the Czech Republic supports this (see Table 3.1). We do not use the UNDP data, even thought it might at first seem to be more appropriate (larger sample, better sampling methodology), because it only covers Roma people that live in regions with higher than average proportion of the Roma and are therefore representative only for Roma living in those areas (FRA-UNDP, 2012). We assume that this might neglect a significant part of better educated Roma (evidence of that can be seen in Table 3.1). The census data we use are those of the largest sample possible (i.e. all those who declared to be Roma either as the only or one of two nationalities), the total 12 953 people.

Table 3.1: Roma educational distribution

	0	A	B	C	D
National (Census 2011)	1%	15%	35%	35%	14%
Roma 5,125 (Census 2011)	7%	55%	14%	1%	8%
Roma 12,953 (Census 2011)	12%	57%	16%	8%	7%
Roma (UNDP)	5%	65%	30%[15]		0%

To derive the Roma representative data we further need to know the percentage of the Roma among the unemployed, among those, who obtain individual welfare benefits and those who are incarcerated. The Ministry of Labour and Social Affairs of the Czech Republic keeps records of the number of Roma unemployed who are registered as searching for job and this was 47 854 in 2012 (The office of the Government of the Czech Republic, 2013). We will assume that the number of unemployed and the number of unemployed registered as searching for job is the same for the Roma. This might not be true, but surely the latter figure is equal or smaller and therefore it again corresponds to the lower bound principle. The proportion of Roma obtaining individual welfare benefits is not available. We will assume that these figures are the same as national ones (taken from the TAXBEN model) which is consistent with the lower bound principle (we assume that the Roma receive generally more welfare benefits than is the national average). The nationality of the incarcerated is not recorded nowadays, but it used to be in the early 1990s. In 1992, 16% percent of incarcerated were Roma (Bakalář, 2004). We have no reason to suppose that the figure differs significantly nowadays.

3.2   Data for the second hypothesis

The second hypothesis has similar logic as the first one and also uses the same data with two exceptions. The first one are the educational categories. Now more detailed version is used whenever possible. When the data include only broader categories, we assume that the distribution is equal within the narrower ones included in each broad category. The detailed categories used are:

0: no education

1: elementary education

2: secondary education without maturity exam

3: secondary education with maturity exam

---

15 30% in B and C together

4: post-secondary non-tertiary education[16]

5: short-cycle tertiary education[17]

6: tertiary education (bachelor degree)

7: tertiary education (master degree)

8: tertiary education (Ph.D. and more)

Secondly, we also employ the different Roma age structure as reported by the Census, that was irrelevant for the first hypothesis. We use the largest sample available i.e. all those who reported to be only Roma or Roma and another nationality. Even though the Census is not a perfect source of data, we believe that the age structure it reports for the Roma again reflects the reality better than the national one and this is further confirmed by comparison to UNDP data. The main difference from the national age distribution is that the Roma population is on average much younger (see Table 5.6). We will further use this data for the computation of the total yearly loss of the state budget due to lower education among the Roma.

3.3 Data for the third hypothesis

In cooperation with the endowment fun Verda we have conducted a survey among its former clients (June 2014 - December 2014). The amount of Verda´s former and current clients is nearly 500 and the beginning of the activity of Verda goes back to 2001, which makes it difficult to contact many clients who have changed all ways to contact them (as email address, telephone number and home address. We have sent 29 emails and addressed 73 people on Facebook, that seemed to be the clients we were searching for. Some people did not reply, and some answered only partially. The total number of returned and filled questionnaires is 27. Its questions relate to education, the amount of scholarship obtained from Verda, employment, social benefits, family background and respondent's subjective opinion on the impact of scholarship and education on the quality of life. The full version of the questionnaire can be found in Appendix C.

Unfortunately, there are some imperfections of these data. Among the most significant ones are: the limited size of the sample, no counter factual group and possible sampling bias caused by the fact, that students that are more diligent are more likely to answer and also those who are successful in life could be more likely

---

16 In Czech: nástavbové studium, pomaturitní kurzy

17 In Czech: vyšší odborné studium

to answer than those who can be ashamed for not having a job or not finishing school successfully. We have been aware of these limitations and we have considered them when interpreting the results. Addressing these problems would definitely increase the quality of the data. We will later suggest some solutions how to improve them in Chapter 5.

4   Methodology

The methodology differs for individual hypotheses, however the first two hypotheses use a similar logic (4.1 and 4.2), then we make a slightly modified computation in part 4.3 and the third hypothesis on the other hand is completely different. The first two hypotheses use a large variety of data and also the computation is composed of many parts. A detailed description of the methodology follows.

4.1   The first hypothesis: Investing extra money into Roma education pays off in fiscal terms

The methodology used for solving the first hypothesis is inspired by Kertesi and Kézdi (2006). We will intentionally follow their procedure to make our results comparable to theirs and to Chobanov et al. (2007) who replicated the study in Bulgaria. Only minor changes have been done due to differences in data availability and educational systems, and they are always mentioned in the text. The basic framework used is the cost-benefit analysis, we look at what a person brings to the national budget, what he or she receives from it, and how these figures differ for Roma in the Czech Republic.

To be explicit about what exactly the hypothesis means, the terms used should be clarified at first. *Investing extra money* in our setting is defined as using extra money beyond the usual costs of state schools to help *a young Roma* successfully finish secondary school with maturity exam. We assume that such an investment exists and that without it a young Roma would obtain only lower education. The investment can occur at any time from the age of four[18], any later investment is discounted to this age. However, the exact form of the investment is behind the scope of this text. *A young Roma* is an average Roma child at the age of four. And third, an investment *pays off in fiscal terms,* if it has positive net benefits for the state budget for an average young Roma. The exact figure of net benefits is an additional result that will be obtained.

---

18 In sensitivity analysis we will also allow for a lower age

The basic financial channels that exist between the state budget and its citizens (and that are at the same time quantifiable and available to us) can be summarized as follows:

1.   Personal income tax (paid by full-time employees)

2.   Social security contributions and health insurance paid by both, employers and employees

3.   Welfare benefits (we use only these: children benefits, housing benefits, subsistence benefit and supplementary housing benefit)

4.   Unemployment benefit

5.   Value added tax and consumption taxes

6.   Incarceration costs

7.   Costs of education

The first, second and fifth channel represent contributions to the state budget and therefore have a positive sign. The rest of the channels have a negative sign, as they represent the flows from the state budget. Apart from Kertesi and Kézdi (2006) we do not include public employment projects, as to our best knowledge no sufficiently detailed data is available. We further assume that anyway this part would not have any impact and that not including it will not bias our results, as Kertesi's and Kézdi's (2006) found that public employment project represent almost 0% of the total contribution.

We do not include welfare benefits such as maternity leave[19] (which is actually a contribution from health insurance) and parental leave[20], because everyone taking care of a child gets these independently on his or her social situation and education. We also do not include birth grant[21], even though it is a benefit only for people with income below certain level and therefore might differ between people with different education, because this benefit is obtained only for the first child (in the current situation) and the data are very poor.

---

19 In Czech: mateřská dovolená

20 In Czech: rodičovský příspěvek

21 In Czech: porodné

For each of these accounts we take average flow to or from the national budget for a person through his or her life (one value for each age from 16 to 65 years) conditional on his or her education. We use five educational categories:

0. No education

A. Elementary education

B. Secondary education without maturity exam

A. Secondary education with maturity exam

D. Tertiary education

The total contribution of an individual is the sum of the flows between him or her and the national budget in the 7 channels mentioned above. Let X_a be the contribution of an individual X with education *a*. And let *b*=1, …, B to be the individual channels. Then we can write:

$$X_a = \sum_1^B X_{ab}$$

Each channel's contribution, X_{ab}, is the sum of yearly flows from 16 to 65 years. The flows are discounted to the age of four when according to (Kertesi and Kézdi, 2006) the investment should start to be the most effective. At this age the child is supposed to start with education. When we sum the channels we get one figure, X_a, for each educational category *a,* all being discounted to the age of four. The final figure for an individual X with education *a* is then the following:

$$X_a = \sum_1^B X_{ab} = \sum_1^B \sum_{t=16}^{65} X_{abt} / (1+r)^{(t-4)}$$

From this formula the reason why we do not include years 66 and more is understandable. It is because the discount at that time is very strong. At the age of 66, using the interest rate 0.2 and discounted to age 4, all the accounts are multiplied by $1/(1+0.2)^{(66-4)}$ which is about 0.3.

The flows in each channel *b,* for each education *a* in time *t* X_{abt} are typically constructed as a multiple of two figures. The first one, $P_{at}(b)$, is the probability that the person of given age and education participates in this channel. For example, in case of personal income tax it is the probability that the person is employed, or in case of incarceration costs it is the probability that the individual is

incarcerated. The second figure is the expected value of the flow for a person of given age and education, $E_{at}(b)$. For example, the yearly average personal income tax paid by employees of given age and education or the yearly costs of incarceration. The total contributions of an individual with education a:

$$X_a = \sum_1^B X_{ab} = \sum_1^B \sum_{t=16}^{65} X_{abt}/(1+r)^{(t-4)} = \sum_1^B \sum_{t=16}^{65} P_{at}(b) * E_{at}(b)/(1+r)^{(t-4)}$$

Finally, we are interested in the net benefits of maturity exam, this means the difference between the contributions of a person with maturity exam and that with education level A and B. We must not forget that a part of students who pass the maturity exam also continue to tertiary education. The benefits of maturity exam are therefore equal to a weighted sum of total benefits of education C and D. The weight is the probability that an individual who passes maturity exam continues to and successfully finishes tertiary education. We assume that the probability is for everyone the same. To determine it, we compare the Czech national proportion of people with education levels C and D. The conditional probability of education D can be written as:

$$\Pi_D = d/(c+d) \quad ,$$

where d is the number of people with education D and c is the amount of people with education C.

Then the total contribution of secondary school with maturity exam is:

$$X_{maturity} = \Pi_D * X_D + (1-\Pi_D) * X_C$$

And the net contribution of maturity exam can be either

$$Net\,benefits_A = X_{maturity} - X_A$$

or

$$Net\,benefits_B = X_{maturity} - X_B$$

depending on whether we expect the individual to have elementary or secondary education without maturity exam in case no investment occurred.

In case the net benefits are a positive number, we can conclude that hypothesis 1 is true: from the fiscal point of view, it pays off to invest extra money into the Roma education. The investment can take values up to the amount of the net benefits. Or more precisely, net benefits represent the maximum value of any investment through the individual's life discounted to age 4.

The only exception to this computation is the seventh channel, educational costs. Here we take the average duration of education for each educational level and

its yearly costs. We also know in which years of life the education takes place and discount the costs accordingly. The summary of the computation of individual channels can be found in Appendix B.

4.1.1 Roma adjustment and discrimination factors

Unfortunately, we have very limited data about the Czech Roma, basically only the distribution of age and education from the Census, and therefore we use several constant adjustment factors to the national data to make them more Roma representative. The construction of individual adjustment factors differs, as we use various forms of information. This is a slight difference from Kertesi and Kézdi (2006) who only use a single procedure which is described in the following paragraph.

First, we use the information from the Ministry of Labour and Social Affairs about the number of unemployed Roma, R_u , we know that the total number of Roma, denoted R, is assumed to be between 150 and 300 thousand. We use the mid point, 225 thousand, for our main computation and allow it to change in the sensitivity analysis later. Then we compute the probability that a Roma is unemployed as R_u / R . Then we assume, that the probability, that a Roma with education a at the age of t is unemployed is the national probability times a constant adjustment factor d:

$$P_{at}(Unemployed\,|\,Roma)=d*P_{at}(Unemployed\,|\,Czech)$$

We take P_a as the overall probability that a person is unemployed, given he or she has education a. Then we can say that the probability that a Roma is unemployed is the weighted sum of probabilities according to education. Using further the adjustment factor d, we can write:

$$R_u / R = \sum_a [(R_a / R)*P_a(Unemployed\,|\,Roma)]=$$

$$=\sum_a [(R_a / R)*d*P_a(Unemployed\,|\,Czech)]=$$

$$=d*\sum_a [(R_a / R)*P_a(Unemployed\,|\,Czech)]$$

so we can write d as:

$$d=(R_u / R)/\sum_a [(R_a / R)*P_a(Unemployed\,|\,Czech)]$$

The adjustment factor d is basically the ratio of a probability that a Roma is unemployed, and a hypothetical probability that a Roma is unemployed, given the Roma educational structure and using the national probabilities.

Second adjustment factor relates to the probability that a Roma person is in prison. Now, we do not know the number of incarcerated Roma, but we have the

information what percentage of all incarcerated are Roma. We also know the total amount of incarcerated and from this we can compute the number of incarcerated Roma and continue in the same way as when computing the unemployment adjustment factor.

Consistently with Kertesi and Kézdi, we use a discrimination factor 15% for both wage and employment probability. We assume that the Roma earn on average only 85% of what does the others earn just due to their ethnicity. We also multiply the national employment probabilities by 0,85, to reflect the fact that some employers refuse the Roma applicants. As the proportion of 15% has been chosen arbitrarily, we allow it to differ in sensitivity analysis to see what would happen if the discrimination was different.

Opposed to Kertesi and Kézdi, we add no adjustment to the probability of obtaining welfare benefits because we have no reliable information about the number of Roma receiving welfare benefits or about the proportions of Roma among all the people receiving them. Nonetheless, assuming the Roma obtain welfare benefits with the same or higher probability, the omission of the adjustment factor corresponds to the lower bound principle, as it probably underestimates the net benefits.

4.1.2 Sensitivity analysis

In the computation of net benefits we use several elective fixed figures which are: discount rate, discount age, Roma wage and employment discrimination adjustment. However, these can be different or change quickly in the real world and therefore we are interested to what extent are our results robust to such changes. We find that by simply observing the impact of changing one of them while keeping the others fixed. We will allow the discount rate to take values from zero to 4%, the discount age to be from 0 to 10 years and the employment discrimination factor as well as the wage adjustment to be from 0% to 30%. We will observe the changes of the last two parameters with respect to each other. We also use the number of Roma for the computation of adjustment factors. We will allow it to change from 150 to 300 thousands.

4.2 The second hypothesis: Investing into Roma education yields varying fiscal benefits

In hypothesis one, we have computed the net benefits to the state budget of an average young Roma who finished secondary school with maturity exam. We have also assumed that without any additional intervention a young Roma would end up either with elementary or with secondary education without maturity exam. And we have further assumed that with an investment the young Roma would surely finish

secondary school with maturity exam and with some probability would enter and finish university. Even though this has brought interesting results, it has included one simplifying assumption, which is not realistic. It had assumed that all the people are the same, or at least they are all capable to achieve the intended level of education and that without the investment they would not achieve any higher education. Now we would like to allow for heterogeneity in cognitive skills. We will assume that skills are non-equally distributed among people and therefore not everyone is capable of finishing secondary or tertiary education even with a huge investment. This fact implies that investing into people might yield different benefits according to their mental abilities and the highest level of education they are able to achieve. We would like to quantify these benefits and compare them to support or contradict the second hypothesis, which states that investment into Roma education yields varying fiscal benefits. We will also allow for the fact, that without the investment, the person can achieve also one of the higher educational level. We will therefore compute the additional net benefits of each level of education as compared to the preceding one. This would quantify the maximum investment that could be done to increase the education by one level.

We will use the highest achieved education as a proxy for cognitive skills or generally "ability" in case of the national data. There surely exist better ones such as IQ, but to our best knowledge, there exist no Czech dataset that includes information about IQ and consumption, welfare benefits etc. We assume that the highest achieved education is the best proxy which is commonly available in the datasets. Therefore, the national distribution of highest achieved education will represent for us the distribution of ability in the population. We will further assume that the distribution of ability is the same among the Czechs (including Roma) and the Roma, even though the distribution of education actually differs. We assume that the Czech educational distribution is the best possible using fully the born abilities of Czech citizens, while the Roma on average do not use their born ability fully and so have lower education than they would be able to have given better educational opportunities.

In the same way as in the first hypothesis we will compute discounted net contributions to the national budget of an individual by summing the discounted costs and benefits in seven channels from the age of 16 to 65 years. The basic difference from the first hypothesis will be that we will use nine narrower educational categories, so that the computation will be more precise. These categories are listed in chapter Data. (For clarity of the text, from now on only their numbers will be mainly used.)

Even though all the datasets used include the categories 0 (no education) and 8 (tertiary education – PhD), we will not use these because there is only a small

proportion of people in these categories which causes lots of missing values which would cause very imprecise results.

The majority of data are available in the rest 7 categories, but for two channels, welfare benefits and unemployment benefits, we used the broader categories as in hypothesis one. The data for these channels are limited and consequently the division into the 7 categories causes many missing values and lower representativeness of the data (fewer people in each category). Therefore, we considered categories 3,4 and 5 to be homogeneous. In the same way we connected categories 6 and 7. This makes the difference between these two sets of categories to be underestimated.

Finally, we will obtain the net benefits of each educational level by comparing the net contribution to the state budget of categories 2 to 7 to category 1, which is the lowest category. Secondly, we will compare the contributions of categories 3 to 7 to category 2. And third, we will also compute the additional net benefits of each educational category with respect to the preceding one, as an additional net benefit of increasing the education by one level. Only categories 4, 5 and 6 are not assumed to follow each other, these are compared to category 3. If there are significant differences between the net benefits of individual educational levels, we will conclude, that hypothesis 2 is true. The investment into Roma education yields varying net benefits dependent on the individual's cognitive skills. If the differences between the net benefits are negligible, we will conclude that hypothesis 2 is not true.

We made two exceptions from the procedure explained above due to unavailability of data. The first is, that for the incarceration probabilities, we do not have the data for category 4, so we expected it is the same as for the third category. Secondly, we do not assign any extra costs to educational category 4, as in this category the education can be either paid by the student or not. This might cause its net benefits to be moderately overestimated.

In this part we use the same Roma adjustment and discrimination factors as for hypothesis one. We do not apply the sensitivity analysis here, as we use the same computational procedure and the same data as for the first hypothesis, even though in more detail. Therefore, we assume, that the sensitivity to changes in parameters would be the same.

4.3 The total yearly loss of the state budget due to lower education among the Roma

Additionally, we will compute the total amount the Czech national budget is loosing per year due to the fact that the Roma have lower education than the national

average is. Here we will use the same data as in the second hypothesis, only a slightly different computation. The goal is to find to what extent increasing the education of all Roma would be significant for the national budget.

In this part we will use the assumption that the distribution of cognitive skills is the same among all nations and so it is among the Czechs and the Roma. This implies that under equal conditions, the distribution of education among the Roma would be the same as among the Czech people. Furthermore, we will assume that the Czechs are as educated as possible, this means that everyone achieves the highest education he or she is capable of. This is surely not realistic, however, as we will see, it corresponds to the lower bound principle used throughout this thesis, which lowers the expected benefits of Roma education. The figure we want to compute is the amount of money the budget is loosing every year due to the lower education of Roma (lower than the national average is). Suppose that the distribution of education among Roma was the same as it is among all the Czech citizens, we are interested in the total gap between this potential situation and the current situation.

The main difference in the computation against the preceding part is that we do not discount the costs and benefits through an individual's life, on the contrary, we take all the Roma that currently live in the Czech Republic. We will multiply the number of Roma in each educational and age category by the probability of obtaining individual benefits, being incarcerated or paying individual taxes and by the expected value of these taxes if paid. Then we will sum up all the categories and all the benefits and costs to the national budget to obtain one figure. We will call it current costs or benefits of the Roma population.

Then we will assume that the distribution of education among the Roma is the same as is the national one but that the Roma keep their age structure. We will do this by distributing the Roma by age into the educational categories using the national average proportions attained education. Now we will do the same as before, but using this changed age-educational structure. Again, we will obtain one figure and we will call it "optimal" costs or benefits of the Roma population to the national budget.

The final figure is the difference between these two costs or benefits:

$$GAP = costs/benefits_{\text{optimal}} - costs/benefits_{\text{current}}$$

This GAP represents the loss of the national budget due to lower education of Roma. In the computation of GAP, we have still assumed that the Roma have different wage and employment opportunities (represented by the discrimination factor) and different probability of incarceration and unemployment (changed by the adjustment factors). These factors were the same in both situations, in the current state and the hypothetical situation.

Secondly, we will imagine a situation, where not only the Roma would be as educated as the Czechs, but also discrimination and any other differences between the Roma and non-Roma would not exist. We will compute GAP 2, which is the difference between the current state and the hypothetical situation. The computation is the same as for GAP, only in the hypothetical situation we leave out all the Roma adjustments (wage and employment discrimination, incarceration and unemployment adjustment factors).

$$GAP\,2 = costs/benefits_{\text{optimal, without discrimination}} - costs/benefits_{\text{current}}$$

The computation of both values, GAP and GAP 2, has a few similar specifics. First, the data on national education includes also the Roma, which lowers the "optimal" educational distribution. However, we neglect this fact, as the Roma represent a minority and moreover, only about 5 000 people have claimed to be Roma in the Census and the rest is not identifiable. Second, we do not include educational costs as it does not make sense in this case. Third, we compute the gaps only for the Roma between 16 and 65 years of age, therefore we assume that the Roma outside this age interval have zero net benefits to the national budget. The people under 16 years of age usually have the same educational costs, therefore there are no gaps. The people above 65 years of age most likely differ in their net contributions to the budget. This means that this restriction underestimates the total gap (in accordance with the lower bound). And last, we assumed that all the Roma are in educational categories 1-8 and due to a limited sample of people in category 8, we used for these the values from category 7.

This computation is from its logic heavily dependent on the number of Roma. As we do not know the exact number, we compute the gaps for several probable sizes of their population ranging of Roma from 150 to 300 thousand Roma.

4.4 Missing data

An additional methodological issue is the imputation of missing data. Due to a limited sample of SILC and The Household Budget Survey we have some cases, when there is no person of given age and education which makes it look like that all the people with given age and education do not pay corresponding taxes or receive benefits. The methodology described so far consists of summing up the lifelong financial flows between an individual and the state and such missing values might overestimate (in case of costs to the national budget) or underestimate (in case of benefits) the final figure of net contributions of an individual. Moreover, the missing values occur more often in categories A and D than in B and C which would bias the results. The imputation mechanism is the same for the first and the second hypothesis and for the computation of the total yearly loss.

4.4.1 Description of the missing data

To decide how to handle the missing data we first have to look at them and try to find out why they are missing (Gelman and Hill, 2006). The data are constructed from large samples by averaging the values for persons with the same age and education. We assume that the missing data arise randomly only due to a limited sample, as it does not make sense that the probability of obtaining benefits or consumption is discontinuous. We assume one exception of the outer data for the youngest or the oldest persons (see for example values for ages 16,17 and 65 in Table 4.1). The missing values here might be caused by the fact that these people simply do not obtain the given benefits of do not pay the given taxes at all. It is almost sure for example that a 16 years old person does not obtain unemployment benefits. Therefore, we take these boundary data as missing non-randomly and leave them empty. In the illustrative table below, we would impute all the missing values between ages 18 and 64 (e.g. 19).

Table 4.1: Illustrative table for missing data

Age	Expected value of PIT
16	0
17	0
18	30 927
19	0
...	...
63	25 023
64	26 148
65	0

Source: Author

4.4.2 Imputation of the missing data

We have several possibilities how to impute randomly missing values. If we could suppose that there is no trend with respect to age, we could simply compute an average or a median from the other values, or take the last value before the missing one and use this to fill the missing ones (see for example Gelman and Hill (2006). However, in our case it is very likely that there is a relationship between age and, for example, consumption, unemployment or wage (and consequently income tax and employment contributions). In this case we will use regression imputation, also called conditional mean imputation, which means regressing the given data on age and then fill the missing values by the predicted values (Gelman and Hill, 2006). The summary

of all the regressions can be found in Appendix A. In some cases we used the linear regression and in some the quadratic one. The decision was always based on the representativeness of the model as expressed by R squared and p-values but also on the logical explanation of the concrete tax or benefit. In the summary table in Appendix A it can be seen that the R-squared is often quite small. We would like to emphasise that we were not searching for the best explanation of variation in individual taxes of benefits. In such case many other variables than age would be surely taken into account. On the other hand we only wanted to find any dependency of the data on age, even if age is not the most important determinant of the data. Therefore, we use the regressions to predict the missing values even though the indicators like R-squared and p-values are not convincing, just because we assume that it is still better than to use the simple average.

Imputing values predicted by simple regression imputation does not respect the variation in the data. For that a more sophisticated method, random regression imputation, can be used which also includes the uncertainty by adding the prediction error (Gelman and Hill, 2006). However, our samples are quite small (less than about 49 observations) and therefore we suppose that the simpler method is more suitable. Furthermore, the reason why it usually matters that the regression imputation does not include variation is that it causes the variance of the data to be biased (Gelmann and Hill, 2006). But in our case, the variation is not the subject of computations.

In some cases we had to erase one or two outlayers but we were careful with this and did it only in really evident cases, where one outlier changed the prediction significantly and this change was against the logic of the tax or benefit. These outliers probably arose when there was a single observation for given age and education, which was unusually high.

Finally, we imputed the data in case of personal income tax, employment contributions, consumption taxes and unemployment benefits but we decided not to impute the welfare benefits data. The reason is that the number of missing values is much higher than in case of the other imputed data. This is obviously because the number of people obtaining the particular welfare benefit is much smaller than the number of people paying consumption of income taxes. Therefore, the limited sample is a bigger problem here. However, it does not seem to make sense to impute the data, as the distribution of missing values might also represent the distribution of probability that the person does not obtain the given welfare benefit at all.

4.5 The third hypothesis

We use descriptive statistics to describe the results of our survey. We are mainly interested in the information about the level of education and employment

status. To answer the third hypothesis we will compute the proportion of employed people for each educational level.

Additionally, we will introduce also other results of the survey, such as the relationship between parent's and children's education and employment, and the amount of respondents receiving welfare benefits.

Unfortunately, our sample is too small to use more sophisticated methods. A regression analysis for example would be very suitable in this case in order to reveal the relationships between education, employment, welfare benefits, the total sum of scholarship and study grants obtained and parent's situation.

5   Results

In this section we will present our results. The first part describes the findings related to the net benefits of Roma education, the second section shows the differences in these net benefits according to the highest level of education achieved and also computes the total yearly loss of the national budget due to lower education of the Roma, and the third one summarizes the results of our survey among the clients of the endowment fund Verda.

5.1   Net benefits of investing into Roma education

The first hypothesis asks whether it pays off to invest extra money into a young Roma. Using the methodology we described in the preceding chapter, we can conclude that it does. At first, we computed the contributions for the state budget for a young Roma (through all his or her life discounted to present) depending on his or her education to be:

Table 5.1: Net budgetary contributions of a Roma according to education

Educational category	Net contributions in CZK
A: elementary education	132 275
B: secondary education without maturity exam	1 112 626
C: secondary education with maturity exam	1 685 908
D: tertiary education	2 555 362

Source: Author's computation using the data presented in Chapter 3

We expect that once a Roma finishes secondary school with maturity exam (C), he or she has a 28,5% probability to finish also a tertiary education (D). This probability is the national conditional probability of D, as explained in chapter Methodology. Therefore, when comparing the difference between a Roma that has education C and a Roma with education A (elementary education) or B (secondary education without maturity exam), we have to account for the fact that nearly a third of the students who finish C also finish D. Therefore the expected contribution of a Roma finishes C is:

$$0,285*2\ 555\ 362 + (1-0,285)*1\ 685\ 908=1\ 933\ 702\ CZK$$

The net benefit is then the difference between this combined contribution and the contribution of person B or A respectively.

(Versus B) 1 933 702 - 1 112 626 = **821 076 CZK**

(Versus A) 1 933 702 - 132 275 = **1 801 427 CZK**

Assuming a hypothetical investment that causes a young Roma at the age of four that he will finish education C and assuming that without this investment the young Roma would end with elementary education (A), the net benefit of such an investment is 1 801 427 CZK. If we assume that without this investment the young Roma would end up with secondary education without maturity exam, then the net benefit of the investment is smaller, but still considerable: 821 076 CZK. This means that we could invest more than 1,8 million into one person, if we knew it would make him or her finish education C and that without this intervention he or she would have education A. And we could spend more than 800 thousands CZK on one person, if we knew it would make him or her finish a secondary school with maturity exam, if otherwise he or she would study and finish only a secondary school without this exam. These investments can take place at any time and be divided into smaller repeated investments. However, the computed figure represent sum of all the investments discounted to age 4 of the person.

To make our results comparable, they have to be expressed in euro. We use the current exchange rate by the Czech National Bank, which is 27,59 CZK per 1 EUR[22]. The following table includes our results and the results of the two studies we are referring to. Both our estimates are lower, however, we have to keep in mind, that the other studies are older. This means that both, the exchange rates and the price levels in these countries, have most probably changed since then. From this point of view, our results are of similar size and also the difference between the two estimates comparable.

---

22 Exchange rate on 12.12.2014, accessed online at
   https://www.cnb.cz/cs/financni_trhy/devizovy_trh/kurzy_devizoveho_trhu/denni_kurz.jsp

Table 5.2: Net benefits of maturity exam in the Czech Republic, Hungary and Bulgaria in EUR

	Czech Republic (Author's computation)	**Hungary** (taken from Kertesi and Kézdi (2006))	**Bulgaria** (taken from Chobanov et al. (2007)
Versus A (elementary education)	65 293	70 000	82 000
Versus B (secondary education without maturity exam)	29 760	55 000	36 000

The net benefits figures are composed of several parts, and now we will look at them to see what are their relative impacts. The percentage figures in Table 5.2 denote the relative part of the net benefits, i.e. the difference between the total contributions of persons with education C and A or C and B. This means they refer to the change of the individual parts of the computation, not to their absolute size. For example, although the consumption tax makes about 7% of the total contributions of a Roma individual with education C, the change of consumption tax, however makes only 1% of the difference between the total benefits of an individual A and C (see Table 5.2).

Table 5.3: Proportional impact of individual parts of the computation on the net benefits

	Versus A	Versus B
Education costs	-12%	-4%
Personal income tax	16%	25%
Employment contribution	53%	52%
Value added tax	11%	13%
Consumption taxes	1%	0%
Unemployment benefits	6%	6%
Welfare benefits	12%	3%
Incarceration costs	14%	6%
Total	100%	100%

Source: Author's computation using the data presented in Chapter 3

We can see in Table 5.3, that only education costs have a negative impact on net benefits. It is because these costs increase with increased education, in other cases, the contributions to the state budget increase with education and the costs decrease, the impact of both on net benefits is positive.

The main part of the net contributions create employment contributions and personal income tax. Incarceration costs seem to be also very important. This is surely because we used a rather high adjustment factor for Roma incarceration probability. However, it is also true that the probability of being incarcerated decreases rapidly with higher levels of education. Very interesting is the comparison of value added tax and consumption taxes. While the consumption of goods taxed by VAT seems to increase with higher levels of education (and consequently higher wages), the consumption of alcohol, tobacco and fuel which are taxed by consumption tax seem not to change with increased education. In fact, if we look at the total net contributions, it slightly decreases with higher levels of education. We can further see that welfare benefits decrease a lot with higher levels of education, on the other hand, unemployment benefits change much less.

For the sake of comparison, we have also computed the net benefits of maturity exam for a Czech person, i.e. without discrimination factors and Roma adjustments. The net benefits are slightly higher, 2 101 601 CZK, when compared to education A and 1 013 128 CZK when compared to education B. The reason why these net benefits are higher is that the Roma adjustment lowers some benefits

(personal income tax, employment contributions) and increases some costs (incarceration, unemployment benefits).

We have to mention here that the computation of contributions of an individual suffers from omitted dimensions. Shadow economics would be probably very significant one. However, adding it to the computation would probably increase the net benefits and therefore again is consistent with the lower bound principle. Pensions on the other hand, which we also omit, represent very significant costs of the state budget.

5.1.1 Sensitivity analysis

Now we are interested in the robustness of the results of hypothesis 1. In the computation we have used several changeable parameters: discount rate, discount age, and wage and employment discrimination factors. In the computation of adjustment factors, we have also used the number of Roma in the Czech Republic. We will observe to which extent do these parameters influence the results.

First we have allowed the discount rate to be 0%, 1%, 3% and 4% instead of the 2% used originally. As we can see in Figure 5.1 the discount rate influences the results heavily with the highest estimated discount rate the net benefits would be less than half of the originally estimated.
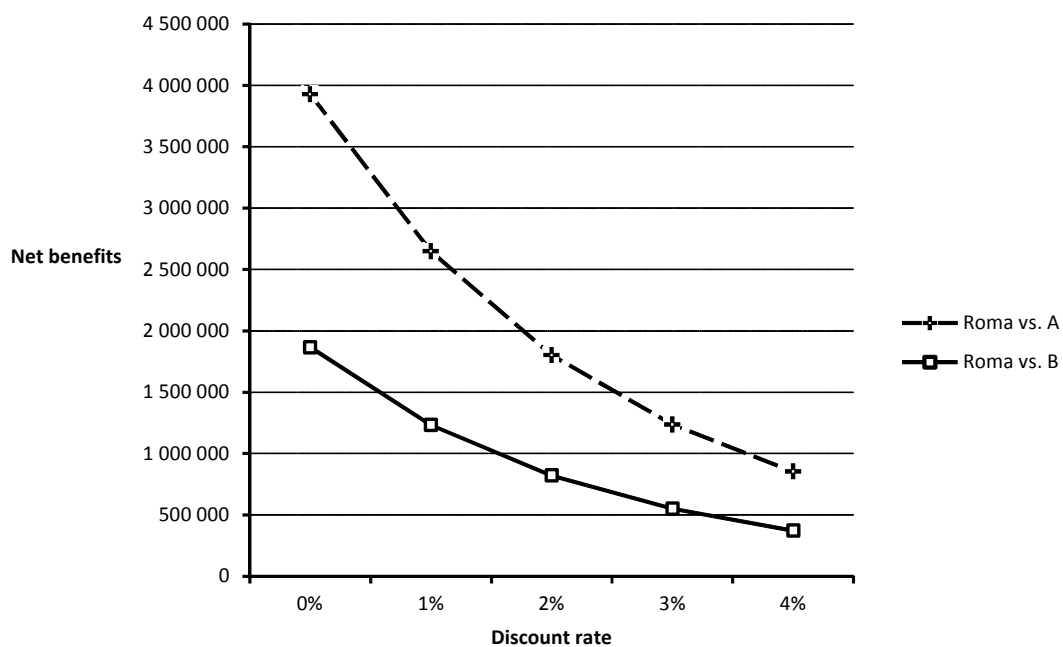
Figure 5.1: Sensitivity analysis of the discount rate

Source: Author's computation using the data presented in Chapter 3

Then we have changed the discount age from 0 to 10 years. We find this very important, as we believe that the investment should start as early as possible. We can

see in Figure 5.2 that the discount age has rather small impact on the figure of net benefits.
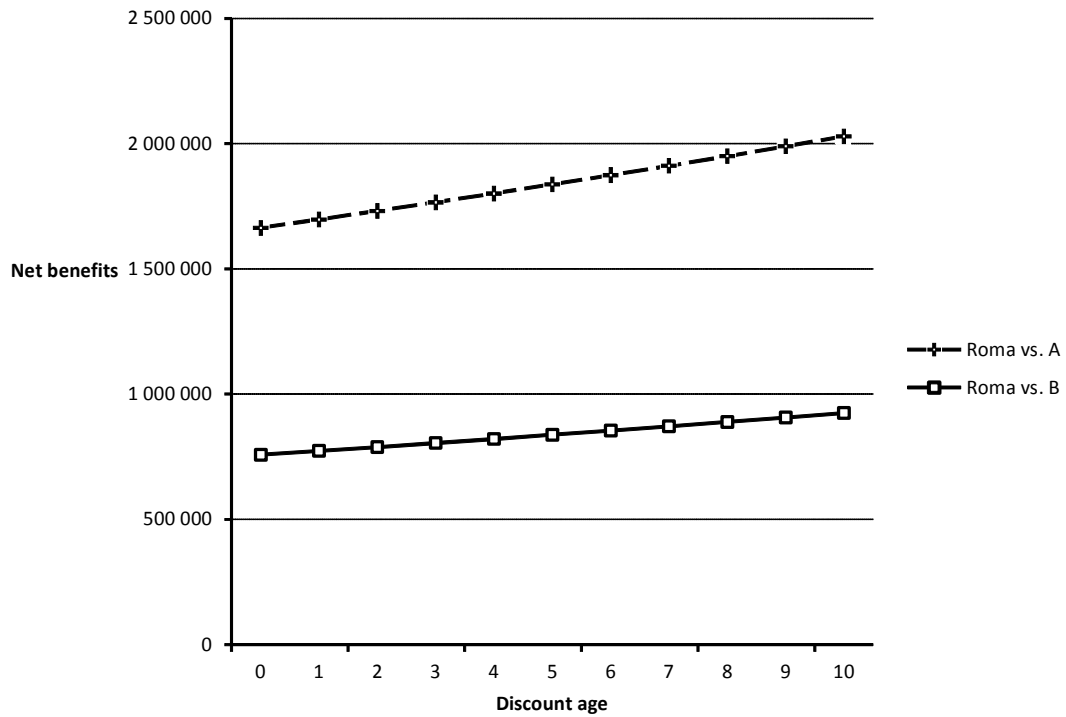
Figure 5.2: Sensitivity analysis of the discount age

Source: Author's computation using the data presented in Chapter 3

We have let the discrimination factor that lowers both, the employment probability and the wage to vary from 0% to 30%. The Figure 5.3 shows the comparison to a person with education A, for three levels of wage discrimination and three levels of employment discrimination. Not surprisingly, the discrimination factor lowers the net benefits.
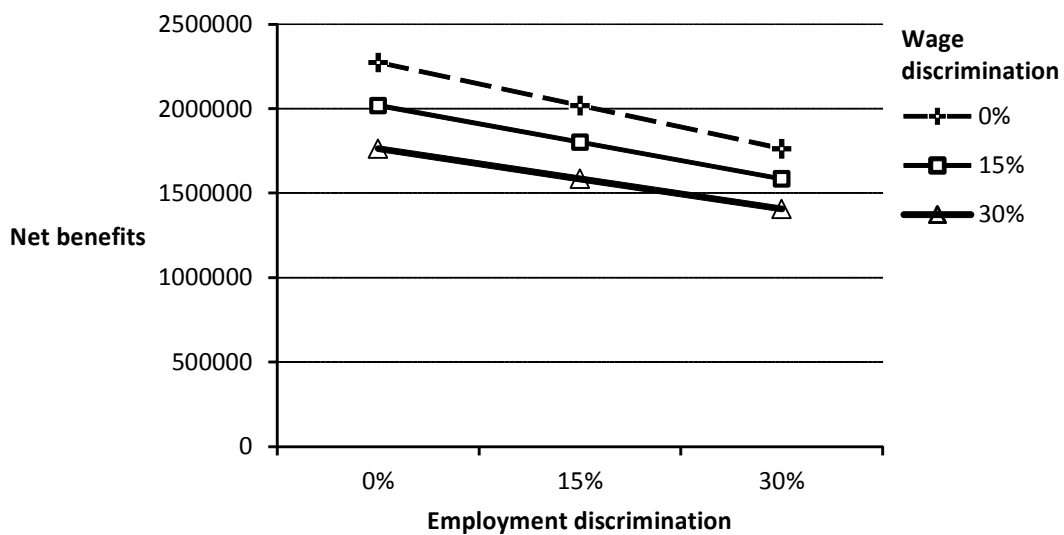
Figure 5.3: Sensitivity analysis of the discrimination factor. Compared to educational level A.

Source: Author's computation using the data presented in Chapter 3

Secondly, Figure 5.4 shows the net benefits of maturity exam as compared to person with education B, again for varying levels of wage and employment discrimination
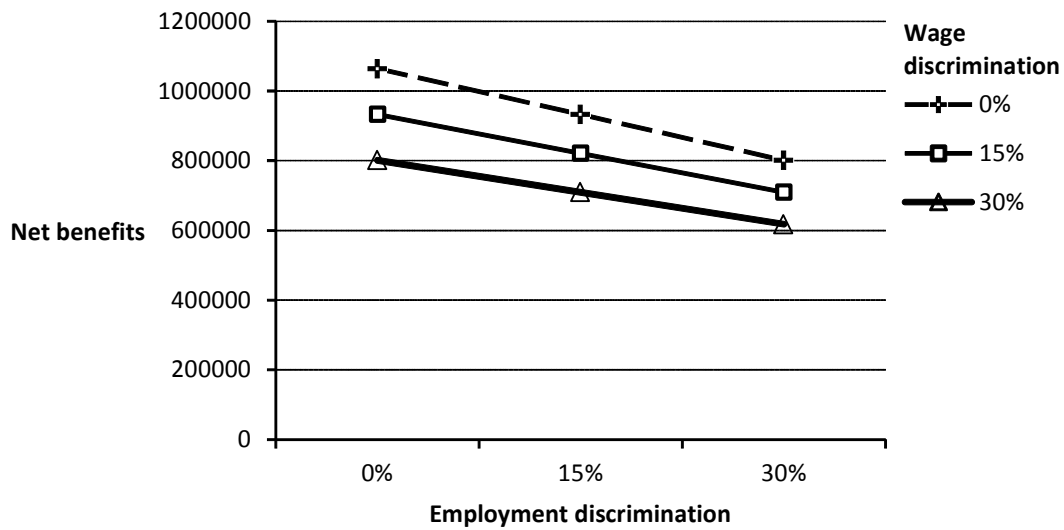
Figure 5.4: Sensitivity analysis of the discrimination factor. Compared to educational level B.

Source: Author's computation using the data presented in Chapter 3

We have let the number of Roma to change from 150 to 300 thousand. Increasing number of Roma causes lower probability of incarceration or unemployment. This decreases the expected costs of unemployment benefits and incarceration. As we can see in Figure 5.5, the net benefits are decreasing with higher numbers of Roma. The explanation is that the net benefits represent the difference between more and less educated people. The unemployment and incarceration costs are higher for more educated people, but with increasing number of Roma, this difference is lower (see the way how we compute the adjustment factors, explained in chapter Methodology). However, the number of Roma does not have a large effect on the net benefits due to the fact that it influences only two channels.
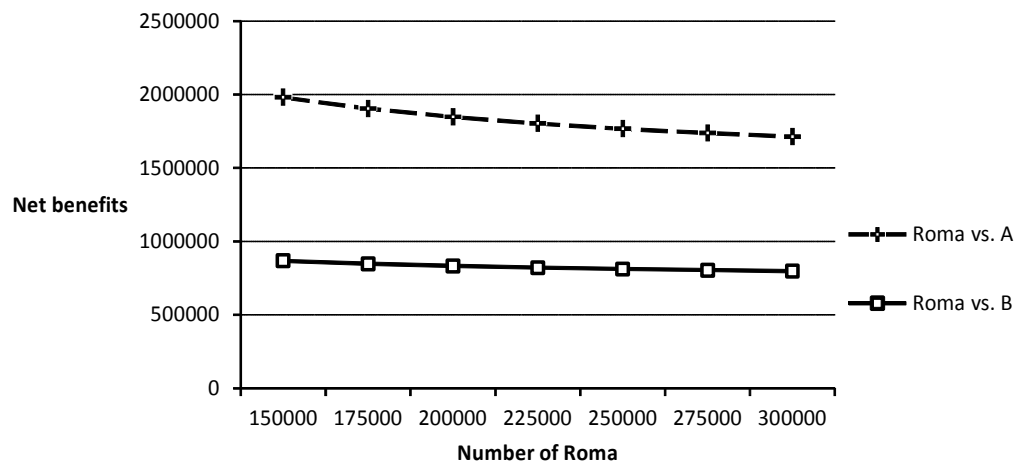
Figure 5.5: Sensitivity analysis of the size of Roma population

Source: Author's computation using the data presented in Chapter 3

5.2 Varying fiscal benefits

The hypothesis 2 states that investment into Roma education yields varying fiscal benefits. Based on assumptions made in chapter Methodology, we suppose that cognitive skills differ among the population. We take the highest possible level of education that can be achieved as a proxy for cognitive skills. If we took a single investment, it could pay off in case of one person (with high cognitive skills) but for another person it might not. We computed the exact contributions of individuals with seven different levels of education (as listed in chapter Methodology) to see, what are the differences between them.

Table 5.4: Net contributions of a Roma according to his or her education

Educational category	Net contributions in CZK
1: elementary education	209 158
2: secondary education with maturity exam	1 123 718
3: secondary educations with maturity exam	1 667 053
4: post-secondary non-tertiary education	1 723 617
5: short-cycle tertiary education	1 661 880
6: tertiary education (bachelor degree)	2 145 409
7: tertiary education (master degree)	2 617 261

Source: Author's computation using the data presented in Chapter 3

We then compute the net benefits as compared to elementary education (1) and to secondary education without maturity exam (2). These correspond to A and B in

Hypothesis 1. In Figure 5.6 we can see the results. The figures correspond to the results of the first hypothesis.

A surprising result is that there is almost no difference between categories 3, 4 and 5 and consequently these categories have also similar net benefits. From this, we can conclude that in hypothesis 1, joining categories 3, 4 and 5 together did not make any simplification. On the other hand, joining 6 and 7 together hides a significant difference between them.
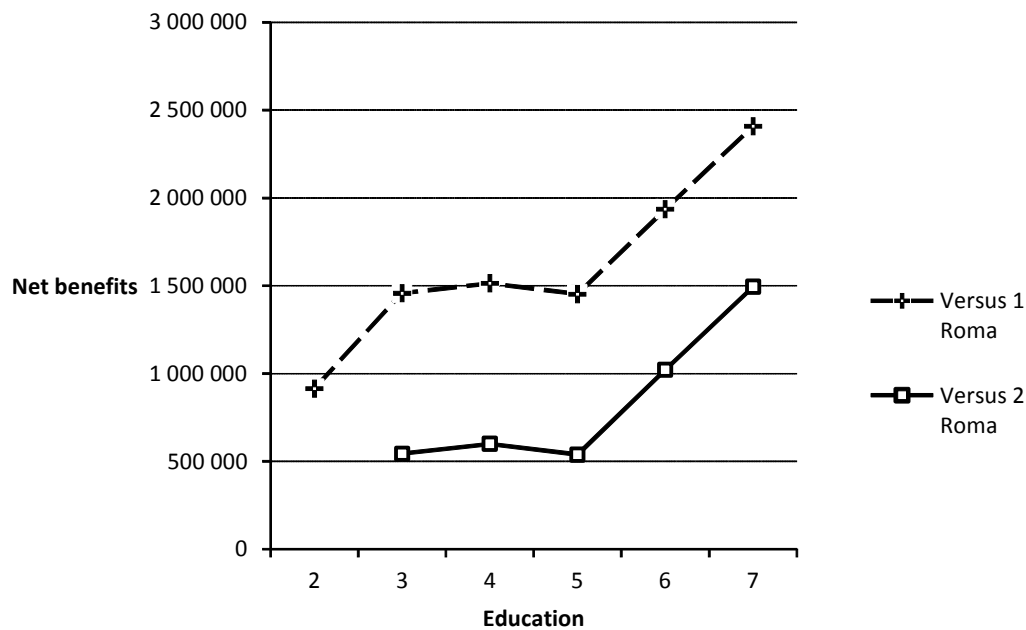
Figure 5.6: Net benefits of individual levels of education as compared to educational categories 1 and 2

Source: Author's computation using the data presented in Chapter 3

Figure 5.6 contains comparisons of total Roma fiscal benefits at individual educational levels to levels 1 or 2. This means that without any intervention the individual would end up with education 1 or 2 respectively, even though he or she could be capable of studying more. However, it is more realistic for an individual to end up at one level lower education than he or she would be able to if supported by an investment. This is why we further present the net benefits of each level as compared to the preceding one (except for levels 4, 5 and 6 that are al expected to follow level 3 and therefore are compared this level).

Table 5.5: Net contributions of education as compared to the preceding level in CZK

Education	Compared to	Net contribution respective to the preceding level
2	1	914 560
3	2	543 335
4	3	56 564
5	3	-5 173
6	3	478 356
7	6	471 852

Source: Author's computation using the data presented in Chapter 3

We can observe (in Figure 5.6), that the net benefits in case of educational levels 3,4 and 5 are nearly the same. Usually both levels 4 and 5 follow directly level 3, so it seems, that adding level 5 lowers the expected benefits compared to stopping at level 3 or 4. However, the differences are very small, co we conclude that these three categories have basically the same net benefits, and we cannot deduce any from the small differences. Moreover, we did not add any additional costs of education of category 4, because in this category, some students pay the education by themselves. This causes the net benefits to be higher than those of category 5, where the extra costs of education are accounted for.

From observing our data, the net benefits of category 8 would be higher or at least the same as that of category 7. However, due to a lot of missing observations, we decided, as further explained in chapter Methodology, not to compute the net benefits for this category.

We can conclude that the fiscal contributions vary significantly between the Roma. Those with the highest level of cognitive skills can bring through their lives more than 2.5 million CZK discounted to the age of 4. Those with the lowest level of cognitive skills bring only about 200 thousand CZK. If a single investment was applied to all Roma children, it would bring varying return.

5.3 The total yearly loss of the Czech national budget due to lower education among the Roma

Using the same data as for answering the second hypothesis, we computed two figures. The first one, called GAP, represents a yearly loss due to lower education of the Roma in the Czech Republic, keeping all the Roma adjustment (representing wage and employment discrimination, higher unemployment among the Roma,

higher proportion of the Roma being incarcerated). The second one, called GAP 2, specifies the same loss, but uses no Roma adjustment, assumes the Roma has the same chances on the job market and the same probability to be incarcerated and unemployed. Both these figures are by definition heavily dependent on the number of Roma in the Czech Republic. Or more specifically, on the number of Roma between 16 and 65, because only for these the net benefits are computed. In our computation we assume that the Roma under 16 and above 65 years of age have zero net benefits. We present all the results in more alternatives with respect to the size of the Roma population.

First we computed the percentage of the Roma between 16 and 65 years of age from the Census. As you can see in the following table, it is nearly the same as the national percentage. However, the percentage of people under 16 years of age indicates, that the proportion of Roma in economically active population will rise in the future.

Table 5.6: The age distribution of Czech and Roma population

Years of age	0-15	16-65	66 and more
Roma	29%	68%	3%
National	15%	70%	15%

Source: Census 2011

All the time, when we specify a number of Roma, we actually assume, that only 68% of this amount (see Table 5.6) have positive net contributions for the state budget.
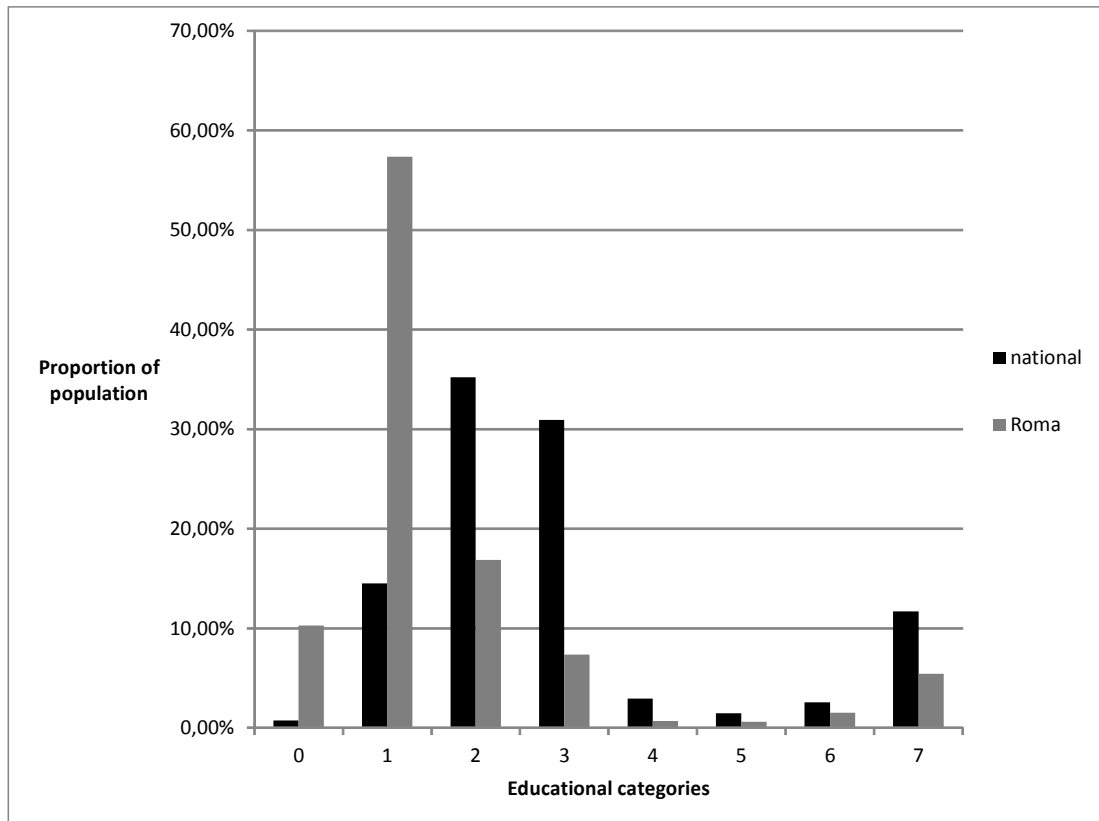
Figure 5.7: Czech and Roma educational distribution

Source: Census 2011

Then we have computed the yearly net benefits of all the Roma. First, we computed the net yearly contributions of the Roma population, this means the positive contributions minus the costs of all the Roma. Net contributions, denoted X in Figure 5.8, show the current situation. Second net contributions, Y, represents a hypothetical situation when the Roma would be as educated as Czechs. And the third net contributions, Z, represent a situation when the Roma would be as educated as the Czechs and there were no differences between Roma and non-Roma (in our case represented by Roma adjustments).
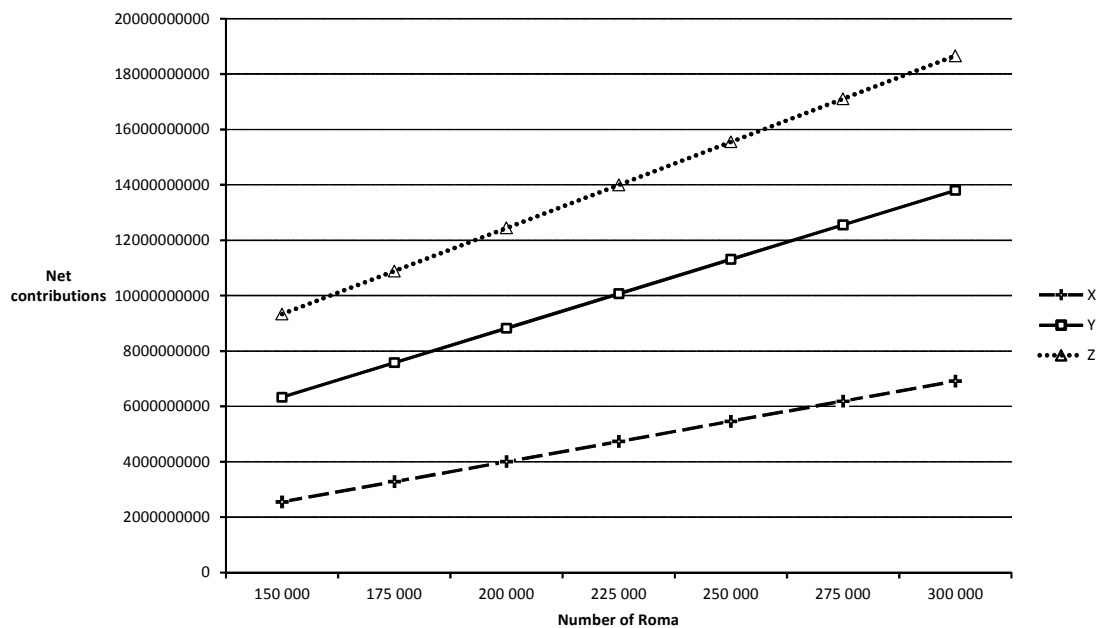
Figure 5.8: Net yearly contributions of the Roma population in the current state (X) and in two hypothetical situations (Y and Z)

Source: Author's computation using the data presented in Chapter 3

Finally, we have computed the differences between these three levels of net contributions for 7 different sizes of Roma population, which are listed in Table 5.7. The first values, GAP, represent the difference between X and Y, and the second values, GAP 2, represent the difference between Z and X. These gaps quantify the total loss that the state faces every year due to lower education of the Roma (GAP) and the lower education plus discrimination of the Roma and their higher likelihood to be unemployed or incarcerated (GAP 2).

Table 5.7: GAP and GAP 2, both in CZK

Number of Roma	GAP	GAP 2
150 000	3 781 332 997	6 776 621 367
175 000	4 298 199 825	7 604 542 813
200 000	4 815 066 653	8 432 464 259
225 000	5 331 933 480	9 260 385 705
250 000	5 848 800 308	10 088 307 151
275 000	6 365 667 136	10 916 228 597
300 000	6 882 533 964	11 744 150 043

Source: Author's computation using the data presented in Chapter 3

The gaps are composed of several parts, as in the computation of hypotheses 1 and 2. Their relative impact on the final gap has corresponding composition as can be found in Table 5.3. All the channels enter the gap as a positive number. In case of

contributions to the national budget it is always a gap between a lower and a higher contribution. In case of costs it is on the other hand a gap between a higher cost (a negative number) and lower cost, therefore again a positive gap.

The yearly income of the budget of the Czech Republic is nowadays more than 1000 milliards CZK. This means that the gap represents about 0,4-1,2% of it, depending on the number of Roma and magnitude of discrimination and Roma specifics. This cannot be seen as insignificant and, moreover, as the number of Roma is growing, we can suppose that also the GAP will increase.

5.4   The clients of the endowment fund Verda

Hypothesis 3 states, that completed secondary education increases the probability of being employed among the clients of endowment fund Verda. We can further extend this to a more general hypothesis, that higher levels of education increase the probability of being employed. Unfortunately, we have only obtained a very limited number of answers. Out of the 27 clients only 3 did not finish the school for which they obtained the scholarship. In the following table below you can see the summary of education achieved and employment status of the clients that participated in our survey. In case of unemployment, we denote in the brackets the number of clients from all the unemployed who answered that they have just finished school (they answered in the summer) and that they are searching for job.

Table 5.8: Results of the survey among clients of Verda

Education	Number of clients	Employed	Unemployed (out of these are still studying)	Studying	Parental leave
1: elementary	6			3	3
2: secondary without maturity	2	1	1		
3: secondary with maturity	14	3	5(2)	5	1
4: post-secondary non-tertiary	1				
5: short-cycle tertiary	2		1(1)	1	
6: tertiary (bachelor)	2	1		1	
7: tertiary (master)	0				
8: tertiary (Ph.D.)	0				
Total	27	5	7	10	4

Source: Author's survey among clients of the endowment fund Verda, 2014

Out of the 27 clients, 5 are employed and 7 unemployed. Out of the 8 clients without maturity exam only one person is employed (i.e. the probability of being employed is 1/8=0.125). Out of the 19 clients with maturity exam 4 are employed (i.e. the probability of being employed is 4/19=0.21). We can see, that the probability of being employed is higher among the clients with maturity exam, however, based on these observations, we cannot confirm that the third hypothesis is true because of the limited amount of answers. The situation is further made difficult by the fact that many of the respondents are currently on parental leave or still studying.

In the questionnaires we have also asked about the parent's education and employment status. We were interested, whether there is a correlation between the parent's and children's education and employment. In such case, it would be very likely that the scholarship would have a low effect and that the student's background would be the main factor influencing the level of education achieved. We have observed that 18 clients had higher education than their parents. In a few cases the parents had elementary education and the clients had secondary education with

maturity exam or higher. Some clients are still studying, so the difference will be probably even higher. On the contrary, we had only 5 answers, where the education of at least one of the client's parents was higher than that of the client and out of that 2 clients are still studying, therefore they might have higher education in the future. In other cases the students either choose not to answer or their education the same as the higher out of the education levels of the client's parents (see the summary in Table 5.9).

Table 5.9: The educational level of clients of Verda relative to the educational level of their parents (we consider the higher education of the two parents)

	Number of clients	Out of that still studying
Higher than parents	18	6
Same as parents	3	1
Lowe than parents	5	2

Source: Author's survey among clients of the endowment fund Verda, 2014

We also asked the clients about receiving welfare benefits. Unfortunately, due to sensitivity of the subject asked, many respondents refused to answer. 10 clients currently receive welfare benefits and 9 do not. Due to the limited sample we cannot deduce any dependency of receiving welfare benefit on the level of education (see Table 5.10).

Table 5.10: Verda clients receiving welfare benefits by education

Welfare benefits	Yes	No
1: elementary education	1	1
2: secondary education without maturity exam	1	1
3: secondary education with maturity exam	8	2
4: post-secondary non-tertiary education		1
5: short-cycle tertiary education		2
6: tertiary education (bachelor degree)		2
Total	10	9

Source: Author's survey among clients of the endowment fund Verda, 2014

One of questions asked was also the total amount of scholarship obtained. Here the problem was, that the respondents very often did not remember it exactly, so it would be better to receive this information directly from the endowment fund. Unfortunately, so far the endowment fund has no such available records.

We recognise two main weaknesses of our work regarding the endowment fund Verda. The first one has been an insufficient sample of clients. And secondly, because of bad data we have been unable to use more advanced methods.

To reduce the first weakness, we would suggest to make a much better records of clients with several means of contact such as e-mail, phone number, address. More contacts increase the probability that even after a few years at least one of them will be valid. Additionally, it would be helpful to contact the students regularly to create a panel dataset, that could be analysed using econometric techniques. However, the clients would have to agree with such a long-term monitoring. This might be a problem, since the clients would not have any motivation for it. We have faced the lack of motivation to answer our survey very often. Furthermore, also other questions could be added to the questionnaires, such as the number of siblings, the number of children, wage, the exact sum of welfare benefits obtained etc. To evaluate the impact of scholarships, the most suitable method would be to use randomized evaluation. This could be done quite simply by choosing randomly the students, who get the scholarship and those who do not. Having a control group would bring many new possibilities of analysis. In the current situation, basically all the applicants get the scholarship. In order to start with the randomized evaluation, either Verda could increase the of scholarship and give it only to a part of the applicants or try to increase the amount of applicants and stay with the current scholarships. Even though it would be a great opportunity for analysis, it might be against the aim of Verda to help as many students as possible. Nevertheless, it might bring valuable information for them.

Having better data, we could surely use better methods than descriptive statistics. The regression analysis could find relationships between the level of education and employment, between parent's and children's education randomized evaluation, regression analysis, relationship between scholarship and education, scholarship and employment, relationships between parent's and children's education with respect to receiving scholarships.

6   Conclusion

Our purpose was to quantify the potential gains from extra investment into the education of Roma. First we discounted the costs and benefits to the national budget of an individual through his or her life. We found, that a Roma with maturity exam brings to the national budget about 1.8 million CZK more than if he or she had only elementary school. And about 800 thousand CZK more than if the Roma had secondary school without maturity exam. This is the maximum sum that could be spent on one Roma at the age of 4 or later, if discounted, to ensure that he or her would pass the maturity exam. Up to that figure, the investment would be profitable. In comparison with similar studies in Hungary (Kertesi and Kézdi, 2006) and in Bulgaria (Chobanov et al., 2007) our results are slightly lower.

We further extended the computation to find how the fiscal contributions vary according to the level of education achieved. If we assume that there are levels of cognitive skills that can be represented by levels of education, then the maximum investment that would pay off differs for people with different level of cognitive skills. It is highly likely that in case a uniform investment would occur, the Roma with higher cognitive skills would bring more profit and those who cannot attain higher levels of education would bring a loss.

From the budgetary point of view, we computed the yearly loss of the national budget which is caused by the fact that the Roma have on average lower education than the majority in the Czech Republic. We found that if the Roma would be as educated as are the Czechs, the budget would get yearly 3.7-6.8 milliard CZK more. The exact estimate depends on the number of Roma in the Czech Republic, which is unknown. Furthermore, if there was no discrimination and no differences between the Roma and non-Roma, the budget would gain yearly 6.7-11.7 milliard CZK.

We find all these potential gains rather high and worth thinking about an investment. To be the most efficient, the investment should start at a very young age and focus on pre-school education, social workers in families, more individual education (for example, employment of pedagogical assistants at schools), remedial classes and financial support of education for children from poor families. It would be great to have more detailed data about the Roma population in the Czech Republic, then it would be possible to make more precise estimates, as we mostly used Czech data with additional Roma adjustment that might not represent the Roma population exactly. Specifically, we were lacking data about the size of the Roma population, its allocation in the Czech Republic, its education, employment, average

wage, etc. Having quality data about the Roma, all the Roma policies could be targeted more accurately.

The survey among the clients of the endowment fund Verda was unfortunately not very successful. Due to a low number of respondents, we did not get the answer to our hypothesis. However, we found some interesting facts about the relationship of parent's and children's education. It would be certainly worth to collect good-quality panel micro data on the Roma in the Czech Republic. Using such data it would be possible to observe the patterns of Roma employment (what motivates the Roma to work, how often they change their jobs, how qualified the Roma workforce is, what impact does their education have on their job etc.). Such data and the results that could be drawn from them would be certainly very beneficial for everyone working with the Roma and for researchers interested in this topic. All theoretical computations about the Roma such as ours would be more credible if they were based on such data. Non-profit organisations have high potential to be intermediaries of such a survey as they are in close contact to the Roma, they often care about some families for a long period and they also offer services to the Roma, that could be used as a motivation for participation in a survey.

Bibliography

Bakalář, P. (2004): "Psychologie Romů." Praha: Votobia

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J. & S. Miller-Johnson (2002): "Early Childhood Education: Young Adult Outcomes From the Abecedarian Project." Applied Developmental Science (1), pp. 42-57.

Card, D. (1999): "The Causal Effect of Education on Earnings." In Orley Ashenfelter and David Card, editors, Handbook of Labor Economics Volume 3A. Amsterdam: Elsevier, 1999.

Chobanov, D., Kostadinova, S. & G. Angelov (2007): "Expected long-term budgetary benefits to Roma education in Bulgaria." Institute for Market Economics

Czech School Inspection (2013): "Výsledky šetření počtu romských žáků v roce 2013", Available online at http://www.infoprovsechny.cz/request/983/response/2279/attach/5/vysledky%20setreni%20CSI%201.pdf (Accessed 4.11.2014)

Czech School Inspection (2010): "Výroční zpráva ČŠI za školní rok 2009/2010", Available online at http://www.csicr.cz/cz/DOKUMENTY/Vyrocni-zpravy/Vyrocni-zprava-CSI-za-skolni-rok-2009-2010 (Accessed 13.12.2014)

Dušek, L., Kalíšková, K. & D. Münich (2013a): "Distribution of Average, Marginal and Participation Tax Rates among Czech Taxpayers: Results from a TAXBEN Model." Czech Journal of Economics and Finance 63(6): pp. 474-504

Dušek, L., Kalíšková, K. & D. Münich (2013b): "Model TAXBEN pro hodnocení dopadů daňových změn." Národohospodářský ústav AVČR, v.v.i., available online at http://idea.cerge-ei.cz/files/danove-modely/certif_metod_taxben_komplet.pdf (Accessed 26.6.2014).

Flores-Lagunes, A. & A. Light (2007): "Interpreting Sheepskin Effects in the Returns to Education." ERS Working Paper 22, Education Research Section, Princeton University.

FRA-UNDP (2012): "The situation of Roma in the 11 EU Member States", European Agency of Fundamental Rights and UNDP. Available online at

http://fra.europa.eu/sites/default/files/fra_uploads/2099-FRA-2012-Roma-at-a-glance_EN.pdf (accessed 4.1.2015).

Gelman, A. and Hill, J. (2006): "Data Analysis Using Regression and Multilevel/Hierarchical Models", Cambridge University Press

Heckman, J. J. (2006): "Investing in Disadvantaged Young Children is an Economically Efficient Policy." Presented at the Committee for Economic Development/The Pew Charitable Trusts/PNC Financial Services Group Forum on "Building the Economic Case for Investments in Preschool"

Heckman, J. J. (2008): "The Case for Investing in Disadvantaged Young Children." CESifo DICE Report, Ifo Institute for Economic Research at the University of Munich, vol. 6(2), pages 3-8.

Kertesi, G. & G. Kézdi (2006): "Expected long-term budgetary benefits to Roma education in Hungary." Budapest Working Papers on the Labour Market 2006/5, Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences.

Kertesi, G. & G. Kézdi (2010): "Roma Employment in Hungary After the Post-Communist Transition." Budapest Working Papers on the Labour Market 1009, Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences.

Kertesi, G. & G. Kézdi (2014): "On the test score gap between Roma and non-Roma students in Hungary and its potential causes." Budapest Working Papers on the Labour Market 2014/1, Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences.

Janský, P. (2014): "Consumer Demand System Estimation and Value Added Tax Reforms in the Czech Republic" Czech Journal of Economics and Finance 64(3): pp. 246-273

Machin, S., Marie, O. & S. Vujić (2010): "The Crime Reducing Effect of Education." CEP Discussion Paper 979, Center for Economic Performance.

Milcher, S. & K. Zigová (2005): "Evidence of Returns to Education Among Roma in Central and Eastern Europe and Their Policy Implications." Managing Global Transitions 3(1): pp. 51-70

Milcher, S. (2011): "Decomposing Income Differentials between Roma and non-Roma in South East Europe." The Journal of the Romanian Regional Science Association 5(1): pp. 27-53

Mincer, J. (1991): "Education and Employment." NBER Working papers 3838, National Bureau of Economic Research.

MŠMT (Ministerstvo školství, mládeže a tělovýchovy) (2014): "Pravidla pro poskytování příspěvku a dotací veřejným vysokým školám." Available online at http://www.msmt.cz/vzdelavani/vysoke-skolstvi/pravidla-pro-poskytovani-prispevku-a-dotaci-verejnym-vysokym-1 (Accessed 26.6.2014)

O'Higgins, N. & N. Ivanov (2006): "Education and Employment Opportunities for the Roma." Comparative Economic Studies, 48(1): pp. 6-19

O'Higgins N. & C. Brüggemann (2013): "The Consequences of Cumulative Discrimination: How Special Schooling Influences Employment and Wages of Roma in the Czech Republic." IZA Discussion Paper 7668, Institute for the Study of Labor.

O'Higgins, N. (2009): ""It's not that I'm a racist, it's that they are Roma": Roma Discrimination and Returns to Education in South Eastern Europe." IZA Discussion Paper 4208, Institute for the Study of Labor.

People in Need (2014): http://www.clovekvtisni.cz/cs/socialni-prace/socialni-integrace/pojdte-do-skolky-vyzyvaji-neziskove-organizace (Accessed 4.11.2014)

Polachek, S. W. (2007): "Earnings Over the Life Cycle: The Mincer Earnings Function and its Applications." IZA Discussion Paper 3181, Institute for the Study of Labor.

Psacharopoulos, G. & H. A. Patrinos (2004): "Returns to Investment in Education: A Further Update." Education Economics, 12(2): pp.111-34.

Public defender of Rights (2012): "Výzkum veřejného ochránce práv k otázce etnického složení žáků bývalých zvláštních škol". Available online at http://www.ochrance.cz/fileadmin/user_upload/DISKRIMINACE/Vyzkum/Vyzkum_skoly-zprava.pdf (Accessed 4.11.2014)

Riddel, W. C. & X. Song (2011): "The Impact of Education on Unemployment Incidence and Re-employment Success: Evidence from the U.S. Labour Market." IZA Discussion Paper 5572, Institute for the Study of Labor.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). "Lifetime effects: The HighScope Perry Preschool study through age 40." Monographs of the HighScope Educational Research Foundation, 14. Ypsilanti, MI: HighScope Press.

Socioklub, kolektiv autorů (1999): Romové v České republice (1945-1998). Praha: Socioklub

The European Court of Human Rights (2007): "D.H. and Others v. the Czech Republic judgement" Available online at http://hudoc.echr.coe.int/sites/eng/pages/search.aspx?i=001-83256 (Accessed 4.11.2014)

The Ministry of Education Youth and Sports in the Czech Republic, (2014): "Vyhlášení Programu Podpora sociálně znevýhodněných romských žáků středních škol a studentů vyšších odborných škol na rok 2014". Available online at: http://www.msmt.cz/vzdelavani/socialni-programy/roma-ss-a-vos (Accessed 4.11.2014)

The office of the Government of the Czech Republic (2013): "Zpráva o stavu romské menšiny v České republice za rok 2012." Oddělení kanceláře Rady vlády ČR pro záležitosti romské menšiny a sekretariátu Rady vlády pro národnostní menšiny, Sekce pro lidská práva, Úřad vlády České republiky. Available online at http://www.vlada.cz/assets/ppov/zalezitosti-romske-komunity/dokumenty/III_Zprava_pripominky_1_11_2013.pdf (Accessed 26.6. 2014)

The office of the Government of the Czech Republic (2014): "Zpráva o stavu romské menšiny v České republice za rok 2013", Oddělení kanceláře Rady vlády ČR pro záležitosti romské menšiny a sekretariátu Rady vlády pro národnostní menšiny, Sekce pro lidská práva, Úřad vlády České republiky. Available online at: http://www.vlada.cz/assets/ppov/zalezitosti-romske-komunity/dokumenty/Zprava-o (Accessed 4.11.2014)

Trentini C. (2014): "Ethnic patterns of returns to education in Bulgaria." Economics of Transition, 22(1): pp. 105-137

UNDP (2002): "Avoiding the Dependency Trap". United Nations Development Programme

VSČR (Vězeňská služba České republiky) (2013): "Statistická ročenka." Generální ředitelství Vězeňské služby České republiky. Available online at http://www.vscr.cz/client_data/1/user_files/19/file/spr%C3%A1vn %C3%AD/statistiky/Statistick%C3%A9%20ro%C4%8Denky/Rocenka_2013.pdf (Accessed 30.8.2014)

World Bank (2010): "Roma Inclusion: An Economic Opportunity for Bulgaria, Czech Republic, Romania, and Serbia", Policy Note of the World Bank. Available online at https://openknowledge.worldbank.org/handle/10986/12905 (Accessed 4.1.2014)

Appendix A: Summary of regressions used for data imputation

In both tables, the independent variable is age, the asterisks represent the significance of each estimate so that:

*** mean p<0.001;** p<0.01; * p<0.05; no asterisk p<0.1

Data imputation for computation of hypothesis 1:

	Intercept	Slope	Quadratic term	Multiple R-squared	Number of imputed values	P-value
Unemployment						
A	1519.17	29.66		0.04401	11	0.2491
D	-1279.36	69.78 *		0.2016	14	0.02773
Value added tax (monthly)						
A	-715,59	87,3066**	-0,9363**	0.2456	5	0.004728
B	-265,16	82,5501***	-0,9216***	0.6095	2	2.649e-09
Consumption taxes (yearly)						
A	-8194,02	650,565*	-7,678**	0.1737	6	0.02933

	Intercept	Slope	Quadratic term	Multiple R-squared	Number of imputed values	P-value
B	5348,8876*	18,7078	-0,5558	0.07891	2	0.178

Data imputation for computation of hypothesis 2:

	Intercept	Slope	Quadratic term	Multiple R-squared	Number of imputed values	P-value
PIT						
4	13252.2*	187.2		0.04557	6	0.2113
5	-1286.9	615.5**		0.2636	13	0.008663
6	-71589.13	4728.72*	-50.32*	0.2473	4	0.0141
Employment contributions						
4	-56290.94	8141.51*	-91.89*	0.1844	6	0.02551
5	-209673.52	16667.90	-192.35	0.3146	13	0.008902
6	-386299.6***	24684.3***	-270.2***	0.662	5	9.803e-09
Value added tax						
1	-8584.67	1047.47**	-11.23**	0.2455	5	0.004733
2	-3182.534	990.601***	-11.059***	0.6094	2	2.673e-09
4	-5378.76	1670.96	-21.70	0.1337	16	0.2215

5	-13345.48	2060.21	-26.63*	0.3145	17	0.01571
6	-21044.40	2399.53*	-29.08*	0.1645	13	0.1057
Consumption taxes						
1	5108.99**	-17.64		0.006034	6	0.6338
2	6273.04***	-28.89		0.07499	2	0.06872
4	10859.42**	-101.39		0.08275	17	0.1832
5	4310.8431*	0.6877		1.357e-05	16	0.9861
6	3961.23*	51.18		0.04546	13	0.2668

Appendix B: Summary of the data sources and computation details of individual channels for hypothesis 1, 2 and the total loss

Channel	Source of the data	Computation details
Personal income tax	Probability of being employed (Census), Expected value of the tax (SILC, taxben)	Probability*Expected value
		Roma adjustment: Probability of being employed multiplied by 0.85; Expected tax multiplied by 0.85
Social security contributions and health insurance	Probability of being employed (Census), Expected value of contributions (SILC, taxben)	Probability*Expected value
		Roma adjustment: Probability of being employed multiplied by 0.85; Expected contribution multiplied by 0.85
Welfare benefits	For each welfare benefit: Probability of obtaining the benefit (SILC,	Probability*Expected value

	taxben); Expected value of the benefit (SILC, taxben)	
		No Roma adjustment
Unemployment benefit	Probability of obtaining the unemployment benefit (SILC, taxben); Expected value of the benefit (SILC, taxben)	Probability*Expected value
		The probability of being unemployed is multiplied by a constant adjustment factor 3.26
Value added tax and consumption taxes	Expected value of VAT and each consumption tax (The Household Budget Survey)	Expected value
		No Roma adjustment
Incarceration costs	Probability of being incarcerated (The Prison Service of the Czech Republic); Daily average costs per one incarcerated person (The Prison Service of the Czech Republic)	Probability*daily cost*365
		The probability of being incarcerated is multiplied by a constant adjustment factor 1.66

Costs of education	The yearly norms per student according to the educational level (The Ministry of Education Youth and Sports in the Czech Republic); The duration of individual levels of education (The Ministry of Education Youth and Sports in the Czech Republic)	The yearly norms assigned to the corresponding ages and educational categories.
		No Roma adjustment

Appendix C: Questionnaire for clients of Verda

Czech version:

1) Nejvyšší dokončené vzdělání, v jakém roce jste dokončil/a tuto školu? Jaký to byl obor? Nebo pokud stále studujete, tak jakou školu?

2) Kolikrát jste dostal/a od Verdy stipendium? Pomohlo Vám k tomu, abyste studoval/a? Co by se podle Vás stalo, kdybyste stipendium nedostal/a?

3) Jste zaměstnaný? Pokud ano, co je Vaše práce?

4) Pomáhá Vám Vaše vzdělání při hledání práce? Nebo při práci?

5) Co jste dělal/a předtím (od dokončení školy)? Byl/a jste zaměstnaný/á nebo ne? A jaké práce jste dělal/a?

6) Pobíráte právě teď nějaké sociální dávky? Pokud ano, jaké? (Podpora v nezaměstnanosti, příspěvek na bydlení, příspěvek na živobytí atd.)

7) Jaké vzdělání má Vaše matka? Je právě teď zaměstnaná? Pokud ano, co dělá?

8) Jaké vzdělání má Váš otec? Je právě teď zaměstnaný? Pokud ano, co dělá?

9) Kde bydlíte? (Vlastní byt, dům, podnájem, v bytě rodičů, partnera/ky) S kým bydlíte? (S rodiči, s partnerem/kou, dětmi atd.)

English translation:

1) What is your highest achieved education? When did you finish this school? What did you study? And if you still study, what school is it?

2) How many times did you obtain the scholarship? Did it help you to study? What would happen if you did not get it?

3) Are you employed? If yes, what is your job?

4) Does your education help you when searching for a job?

5) What did you do between your current job and the end of school? Were you employed? What jobs did you do?

6) Do you currently receive any welfare benefits? If yes, which ones? (unemployment benefits, housing benefit, subsistence benefit etc.)

7) What education does your mother have? Is she employed? If yes, what does she do?

8) What education does your father have? Is he employed? If yes, what does he do?

9) Where do you live? (Own house, flat, rental flat etc.) With whom? (Parents, partner, children etc.)

Appendix D: The transformation of educational categories

Categories for Hypothesis 1	SILC (TAXBEN)	The Household Budget Survey	Census 2011	Prison service of the Czech Republic
0: No education	1: First stage of elementary school	0: No education	1: No education	Special school
		1: First stage of elementary school	2: Not finished elementary school	No education
				Not finished elementary education
A: Elementary education	2: Second stage of elementary school	2: Second stage of elementary school	3: Elementary school	Elementary school
B: Secondary school without maturity exam	3: Secondary education without maturity exam	3: Secondary education without maturity exam	4: Secondary education without maturity exam	Vocational training without maturity exam
				Secondary education without maturity exam
C: Secondary school with maturity exam	4:Secondary education with maturity exam	4:Secondary education without	5: Secondary education without	Vocational training with maturity exam

		maturity exam	maturity exam	
	5: Post-secondary non-tertiary education	5: Post-secondary non-tertiary education	6: Post-secondary non-tertiary education	Secondary technical education with maturity exam
	6: Short-cycle tertiary education	6: Short-cycle tertiary education	7: Short-cycle tertiary education	Secondary general education with maturity exam
D: University education	7: Tertiary education: bachelor degree	7: Tertiary education: bachelor degree	8: Tertiary education: bachelor degree	Tertiary education: bachelor degree
	8: Tertiary education: master degree	8: Tertiary education: master degree	9: Tertiary education: master degree	Tertiary education: master degree
	9: Tertiary education: Ph.D.	9: Tertiary education: Ph.D.	10: Tertiary education: Ph.D.	Tertiary education: Ph.D.