10.14754/CEU.2016.11

Essays on the Evaluation of Educational Policies

by

Anna Adamecz

Submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

at Central European University

Advisor: Gábor Kézdi

Budapest, Hungary

10.14754/CEU.2016.11

CENTRAL EUROPEAN UNIVERSITY
DEPARTMENT OF ECONOMICS

The undersigned hereby certify that they have read and recommend to the Department of Economics for acceptance a thesis entitled **"Essays on the Evaluation of Educational Policies"** by **Anna Adamecz**.
Dated: October 28, 2016

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Chair of the Thesis Committee:

Botond Kőszegi

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Advisor:

Gábor Kézdi

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Internal Examiner:

Róbert Lieli

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
External Examiner:

Hessel Oosterbeek

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Internal Member:

Attila Rátfai

10.14754/CEU.2016.11

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

External Member:

András Simonovits

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

External Member:

Ágnes Szabó-Morvai

10.14754/CEU.2016.11

CENTRAL EUROPEAN UNIVERSITY

DEPARTMENT OF ECONOMICS

Author: Anna Adamecz
Title: Essays on the Evaluation of Educational Policies
Degree: Ph.D.
Dated: October 28, 2016.

Hereby I testify that this thesis contains no material accepted for any other degree in any other institution and that it contains no material previously written and/or published by any other person except where appropriate acknowledgment is made.

Signature of the author ....
Anna Adamecz

10.14754/CEU.2016.11

Disclosure of co-authors contribution

Title of paper (Chapter 2): The Effects of Increased Compulsory School Leaving Age on the Teenage Fertility of Roma Women, a Disadvantaged Ethnic Minority
Co-authors: Flóra Samu and Ágota Scharle

This project started as a research project of the Budapest Institute where I worked as a Senior Researcher that time. The original research idea of using the change in the CSL age legislation as a natural experiment was formulated by the three of us. We applied for funding to the National Scientific Research Program (OTKA) with this research idea under the leadership of Ágota Scharle. Our application was successful; thus, the first phase of this research was supported by the OTKA Grant No. 112792. The original focus of the research was on the teenage fertility of children of disadvantaged parents, using solely the Census data. After some exploratory data work done by Flóra Samu based on my instructions I explicitly asked my co-authors to let me work alone on this research. I realized that the original idea of the proposal is not feasible but we can measure the effect of increased CSL age on the teenage fertility of Roma women. After developing my empirical strategy for Chapter 1, I applied the same methodology to examine this research question in Chapter 2. Ágota Scharle wrote a brief summary of the literature explaining why CSL age matters with respect to teenage pregnancy to the original research report submitted to the National Scientific Research Program in December 2015, commented on the presentation of the paper at various stages and suggested the use of the Vital Statistics data in addition to the Census.

Title of paper (Chapter 3): Integrated Education of Disadvantaged Ethnic Minorities: The Effect of the OOIH Demonstration Program on Roma and Non-Roma Students in Hungary
Co-authors: Gábor Kézdi and Éva Surányi

This paper is a re-evaluation of the OOIH Demonstration Program. The program was evaluated at first by Kézdi and Surányi (2009) and the first results of the original evaluation were published as a research report. The matching strategy was worked out and all the data were collected by Kézdi and Surányi (2009). My contribution includes taking multiple testing seriously and applying appropriate testing procedures (aggregate index of outcome variables, p-value correction methods) on the original data. Furthermore, I applied small sample inference on the school-level data such as the Fisher's exact testing procedure and the simulation-based exact testing method of Bertrand, Duflo and Mullainathan (2002). In addition to this, I conducted a matching sensitivity analysis (Rosenbaum bounds) to see how sensitive the results are to a hypothetical unobserved variable affecting both program assignment and the outcome variables.

10.14754/CEU.2016.11

Abstract

In my thesis I evaluate the effects of educational policy interventions. Educational outcomes are important factors of economic and social success. According to the human capital theory, individuals invest time and effort in their education for immediate and future gains. In the theoretical model of schooling investment and consumption decisions, one decides about how much time and effort to invest in learning maximizing the difference of the expected present value of lifetime wages and non-monetary benefits from schooling, and the actual costs of going to school and taking efforts. Both theoretical and empirical evidence show that such decision-making process may lead to lower-than-optimal schooling investment decisions, especially in the case of children of low socio-economic background, because they discount future returns more heavily, and also, learning may require more efforts from them. Educational policies can influence the production of educational outcomes through two main channels. First, they may aim to induce individuals to invest more into learning. Second, they can increase the productivity of the learning process within schools to produce higher outcomes. I examine examples to these two types of policy measures in my dissertation. In the first two chapters I estimate the impacts of increasing the compulsory school leaving (CSL) age in Hungary. CSL age legislation introduces a constraint into the mechanism in which one decides about how much time to invest in going to school. I make use of a legislation change that increased CSL age from 16 to 18. In the first chapter, I estimate the effects of increased CSL age on secondary school track choice which occurs at age 14 and secondary school dropout rates. I find that the legislation change resulted in an increased probability that children would choose the academic high school track instead of vocational training schools. At the same time, those choosing vocational training schools are more likely to drop out under the higher CSL age scheme. Potential explanations of increased dropout rates include a decrease in the quality of teaching in vocational training schools due to supply constraints, and a shift in student composition to include more students from lower socioeconomic backgrounds. The second chapter is a joint work with Flóra Samu and Ágota Scharle. We are looking at the effects of increased CSL age on the teenage fertility of Roma women, a disadvantaged ethnic minority in Hungary. We provide evidence that the legislation change decreased the probability of teenage motherhood among Roma women by 6.8 percentage points. This effect is temporary as higher CSL age delayed first birth-giving by two years. We exploit a unique database that covers live births, miscarriages, abortions, and still births, and contains information on the time of conception by weekly precision. We propose that the impact of the legislation change can be explained exclusively by the incapacitation effect of education, which keeps women physically in school: the higher CSL age decreases the probability of getting pregnant during the school year but not during summer and Christmas breaks. The third chapter, which is a joint work with Gábor Kézdi and Éva Surányi, considers educational policy from a different angle. It estimates the effects of a change in the technology of educational production within schools by looking at the OOIH demonstration program in Hungary. The program supported teachers and the management of schools with disadvantaged Roma students, mixed with non-Roma students, and aimed at helping the development of all students in an integrated school environment. We find that the program had significant positive effects on academic development (especially for Roma students), socio-emotional skills (in both ethnic groups), and inter-ethnic attitudes of non-Roma students. In my thesis I document that increasing the CSL age affects forward looking decision making about secondary school track choice, and impacts the distribution of students in school. I find that these effects are the strongest among children of low-educated parents. Furthermore, I provide evidence that higher CSL age can reduce teenage pregnancy solely through the incapacitation effect of being in school, even in a case when no human capital effects of education can be detected. Lastly, I document that a sensitive approach to the integrated education of Roma and non-Roma students is beneficial for all parties involved.

Chapter 1: Increased Compulsory School Leaving Age Affects Secondary School Track Choice and Increases Dropout Rates in Vocational Training Schools

In the first chapter, I estimate the effects that an increase of the CSL age from age 16 to 18 in 1996 had on schooling outcomes in Hungary using a regression discontinuity design (RDD) strategy. The new CSL age came into force with students starting elementary school in September 1998. Identification is based on the age of elementary school start rule. Children compliant with the age rule started elementary school under the new CSL age scheme if they were born on June 1, 1991, or later. Those compliant with the age rule and born before this date had started elementary school in the previous year, under the old CSL age scheme. Thus, a natural cutoff point occurs at this date of birth which allows me to construct a fuzzy RDD strategy

to estimate the intention to treat (ITT) and local average treatment effects (LATE) of the increase using the 2011 and 2001 Hungarian Census data. The Hungarian reform is unique because the first treated cohort knew already at age 6 that they would have to stay in school for two years longer. This fact allows me to test whether this information affected forward-looking decision-making as measured by secondary school track choice at age 14. I find that as a result of the CSL age increase, children at age 14 were more likely to choose the more demanding and more beneficial 4-year academic high school track instead of the 4-year vocational training school track. The legislation change did not influence more individuals to start secondary school, but those who did decide to start were more likely to opt for an academic high school rather than a vocational training school. At the same time, those who did choose the vocational training school track were more likely to drop out under the new scheme. The data suggest two potential explanations for this adverse effect. First, the financial and human resources allocated to vocational training schools were not adequate for the sudden increase in the number of students. Second, as the students from higher socioeconomic standings, and with stronger abilities chose academic high schools instead, and lower standing, lower performing students stayed in vocational training school longer, the distribution of students in vocational training schools shifted towards lower socioeconomic status students. The last takeaway from this analysis is that increasing the CSL age may not always be a good instrumental variable (IV) for education. It harms the monotonicity assumption of the instrument if the quality of education for some students is negatively affected by the increase (Cygan-Rehm and Maeder, 2013). The monotonicity assumption requires students to be impacted by the instrument in the same way (Angrist and Pischke, 2009). In this context, it would assume that the legislation change induced some individuals to have more education and for no one to have less education, both in terms of length, tracks, quality, and earned degrees. In the case of the Hungarian reform, this concern is valid if one wants to use the increase of the CSL age as an IV to education, as the legislation change did increase dropout rates in vocational training schools.

Chapter 2: The Effects of Increased Compulsory School Leaving Age on the Teenage Fertility of Roma Women, a Disadvantaged Ethnic Minority (joint work with Flóra Samu and Ágota Scharle)

In the second chapter, which is a joint work with Flóra Samu and Ágota Scharle, we estimate how higher CSL age affects the teenage childbearing of Roma women. Roma make up the largest ethnic minority in Hungary, belonging to the Roma minority is highly correlated with poverty, social exclusion, and long term unemployment. Teenage fertility of non-Roma women is very low in Hungary, while among Roma women, it is comparable to levels measured in the Congo and Kenya. Teenage fertility is one of the most important sources of intergenerational poverty transmission. The literature presents evidence on the negative health, social, and economic consequences of teenage childbearing. Due to high opportunity costs, the prevalence of teenage motherhood has been declining in most developed countries. However, teenage motherhood is still very common in several communities of disadvantaged ethnic minorities living in developed countries. Examples include Mexican women in the US, women of Pakistani and Bangladeshi origin in the UK, Turkish women in Belgium and France, and Roma women throughout Europe. Despite the large selection of literature that looks at the effects of education on early fertility in general, we know very little about the effects of education on teenage motherhood of ethnic minority women in particular. Using the same identification strategy as in the previous chapter, we construct a fuzzy regression discontinuity design (RDD) identification strategy to estimate the intention-to-treat (ITT) effect of this legislation change on the probability of both early childbearing and pregnancy. We exploit a unique data set of all known pregnancies, including live births, abortions, fetal losses, and still births, linked to a large subsample of the 2011 Hungarian Census. We find that the higher CSL age decreased the probability of teenage motherhood among Roma women by 6.8 percentage points. This effect is temporary as it only creates a two-year delay in motherhood. We find no effect among non-Roma women, where teenage fertility is rare. We use our rich data sets to reconstruct the conception time of all pregnancies of Roma women. We show that the legislation change decreased the probability of getting pregnant during the school year but not during summer and Christmas breaks. Thus, we find no evidence of any impact through the human capital channel. This result is in line with the argument of the literature: teenage childbearing in disadvantaged ethnic minority communities might be more likely to be influenced if human capital development happens alongside increased economic opportunities. In addition to our contribution to the literature on the fertility of disadvantaged ethnic minorities, this is also the first known paper to document that being physically present in school contemporaneously lowers the probability of getting pregnant.

10.14754/CEU.2016.11

Chapter 3: Integrated Education of Disadvantaged Ethnic Minorities: The Effect of the OOIH Demonstration Program on Roma and Non-Roma Students in Hungary (joint work with Gábor Kézdi and Éva Surányi)

The third chapter is a joint work with Gábor Kézdi and Éva Surányi. In this paper we evaluate a program in Hungary that aimed at transforming the management and teaching practices of schools with a sizable disadvantaged Roma minority to provide high-quality education to all students in a mixed environment. The target schools were regular primary schools covering grades 1 through 8, with 20 to 40 percent of Roma students. The evaluation is non-experimental as the participating schools applied to an open call and were selected by the program administrators. We matched a control school to each of the 30 program schools in the sample, based on pre-program characteristics of student composition and aggregate test scores. Survey data were collected on school characteristics, student background, and academic outcomes including standardized cognitive tests, socio-emotional skills and inter-group tolerance. The treated and control samples are balanced both in terms of pre-program characteristic and student composition measured by subsequent surveys. Preliminary results of the evaluation were published by Kézdi and Surányi (2009). In this study we take a more systematic approach and address potential problems arising from clustered samples and multiple inference. We find that the effect of the program was positive overall. The effect is positive on academic development, especially for Roma students, and no negative effects are found on non-Roma students. The effects are positive on socio-emotional skills in both ethnic groups. Anti-Roma sentiments and social distance of non-Roma students are decreased due to the program.

10.14754/CEU.2016.11

Acknowledgments

I am grateful to my supervisor, Gábor Kézdi, for the tremendous advice on how to proceed with my research as well as for his continuous support. I owe him for not letting me quit when I did not see the light at the end of the tunnel. I am indebted to my examiners, Hessel Oosterbeek and Robert Lieli, for their thorough revisions, useful comments and great suggestions. I have learned a lot from their revisions. I also thank to Ágota Scharle and Balázs Váradi for their ideas and encouragement along the way.

I would like to express my gratitude to my professors and fellow students at the Central European University; being a student of CEU truly has been an eye-opening experience. I am grateful for the participants of the weekly therapy sessions of Gábor Kézdi at CEU, and the participants of the many seminars at the CEU, the Budapest Institute and the IE-CERSHAS in Szirák. I especially thank for all the help regarding data access to Gábor Kézdi and the IE-CERSHAS, in particular to Tibor Czeglédi and the research room team of the Hungarian Central Statistical Office. Some of this research was partly financed by the Hungarian National Scientific Research Program (OTKA), Grant no. 112792.

Last but not least, I owe my husband for his patience and encouragement in the crazy days of completing my dissertation. This dissertation would not have been completed without his love.

Contents

1 **Increased Compulsory School Leaving Age Affects Secondary School Track Choice and Increases Dropout Rates in Vocational Training Schools** 4
1.1 Introduction . . . 4
1.2 Data . . . 5
1.3 The Hungarian Education System and the Legislation Change . . . 6
1.3.1 The Context of Hungarian Education and the Public Education Act (1996) . . . 6
1.3.2 Compulsory Education in Hungary . . . 7
1.4 Identification and Empirical Methodology . . . 10
1.4.1 Identification Strategy . . . 10
1.4.2 Intention to Treat (ITT) Effects . . . 14
1.4.3 Local Average Treatment Effects (LATE) . . . 15
1.5 Results . . . 17
1.5.1 The Implementation of Increased CSL Age . . . 17
1.5.2 Effects on School Choice . . . 18
1.5.3 Effects on Dropout Rates . . . 19
1.5.4 Effects on School Completion . . . 20
1.5.5 Local Average Treatment Effects (LATE) of the Legislation Change . . . 22
1.5.6 The Heterogeneity of the Effects of Gender and Roma Ethnicity . . . 22
1.6 Heterogeneous Effects by Parental Education . . . 24
1.6.1 Data Sources Used to Estimate Parental Education Data . . . 24
1.6.2 Children of Less Educated Parents Are the Most Likely to Be Affected by the Increase 25
1.6.3 The First Stage by Maternal Education . . . 25
1.6.4 The Implementation of Increased CSL Age . . . 26
1.6.5 Effects on School Choice . . . 27
1.6.6 Effects on Dropout Rates . . . 29
1.6.7 Effects on School Completion . . . 31
1.6.8 Local Average Treatment Effects (LATE) of the Legislation Change . . . 33
1.7 Robustness Checks . . . 35
1.7.1 Parental Education Data from Other Sources Give Similar Results . . . 35
1.7.2 No Significant Effects at Cutoffs in the Before and After Years . . . 36
1.7.3 Alternative Optimal Bandwidth Choices Give Similar Results . . . 38
1.7.4 Global Polynomial Estimation Gives Similar Results . . . 39
1.8 Supportive Evidence from Other Data Sources and Potential Channels . . . 41
1.8.1 The Share of those Attending Academic High Schools in the NABC Data . . . 41
1.8.2 Potential Explanation to the Increase in Dropouts: Change in Student Composition . 41
1.8.3 Potential Explanation to an Increase in Dropouts: Supply Constraints in Vocational Training Schools . . . 42
1.9 Discussion . . . 43

2 The Effects of Increased Compulsory School Leaving Age on the Teenage Fertility of Roma Women, a Disadvantaged Ethnic Minority 45
2.1 Introduction . . . 45
2.2 Data and the Teenage Fertility of Roma Women . . . 46
2.2.1 Data . . . 46
2.2.2 Identifying Roma in the 2001 and 2011 Hungarian Census Data . . . 48
2.2.3 Teenage Fertility in Hungary . . . 49
2.3 Institutional Background and the Legislation Change . . . 49
2.3.1 Roma Students in the Hungarian Education System . . . 49
2.3.2 Compulsory Education and the Legislation Change . . . 50
2.4 Identification Strategy and Empirical Methods . . . 52
2.4.1 Identification Strategy . . . 52
2.4.2 Empirical Methods . . . 59
2.4.3 The Number of Completed School Years Around the Cutoff . . . 60
2.4.4 Teenage Fertility Around the Cutoff . . . 61
2.5 Estimation Results . . . 62
2.5.1 The Effects of the CSL Age Increase on Teenage Motherhood . . . 62
2.5.2 The Mechanics of the Incapacitation Effect . . . 64
2.6 Robustness Checks . . . 66
2.6.1 Different Bandwidth Choices . . . 66
2.6.2 Effect of Cutoffs in 1990-1992 . . . 67
2.6.3 Controlling for Monthly and Yearly Seasonality (Parametric Approach) . . . 68
2.7 The Potential Benefits of Delayed Motherhood . . . 69
2.8 Discussion . . . 71

3 Integrated Education of Disadvantaged Ethnic Minorities: The Effect of the OOIH Demonstration Program on Roma and Non-Roma Students in Hungary 72
3.1 Introduction . . . 72
3.2 Background . . . 73
3.2.1 The Roma Minority in Hungary . . . 73
3.2.2 The Hungarian School System . . . 73
3.2.3 The OOIH Demonstration Program . . . 73
3.3 Data . . . 74
3.3.1 Measuring Ethnicity . . . 75
3.3.2 Outcome Measures . . . 75
3.3.3 Controlling for Social Desirability . . . 77
3.4 Evaluation Framework . . . 78
3.4.1 Matching Control Schools . . . 78
3.4.2 Selection into Program and Control Groups . . . 79
3.4.3 Multiple Outcomes and Hypothesis Testing . . . 82
3.4.4 Estimation . . . 83
3.5 School-Level Estimates of Overall Effects . . . 84
3.5.1 Sensitivity Test for Unobserved Characteristics . . . 84
3.6 Individual Estimates: Effects on Roma and Non-Roma Students . . . 86
3.6.1 Academic Achievement, Socio-emotional Skills and Anti-Roma Sentiments . . . 86
3.6.2 Effects on Roma and Non-Roma Students . . . 88
3.7 Conclusions . . . 90

4 *References* 91
4.1 References for Chapter 1 . . . 91
4.2 References for Chapter 2 . . . 94
4.3 References for Chapter 3 . . . 98

A Appendix for Chapter 1 102
A.1 Appendix . . . 102
A.2 Appendix . . . 109
A.3 Appendix . . . 111
A.4 Appendix . . . 112

B Appendix for Chapter 2 113
B.1 Appendix . . . 113

C Appendix for Chapter 3 115
C.1 Appendix . . . 115
C.2 Appendix . . . 118
C.3 Appendix . . . 119

Chapter 1

Increased Compulsory School Leaving Age Affects Secondary School Track Choice and Increases Dropout Rates in Vocational Training Schools

1.1 Introduction

Compulsory school leaving (CSL) age policies introduce a constraint in making a decision about how much time and effort to invest in attending school. A wide range of literature estimates the effects that increasing the CSL age has on social and economic outcomes such as wages (Meghir and Palme, 2005; Grenet, 2013, etc.), mortality (Lleras-Muney, 2005), fertility (Black et al., 2008), crime (Lochner and Moretti, 2004), or voting behavior (Milligan et al., 2004). The evidence of this literature is mixed. In some cases, increasing the CSL age has had positive wage returns (Oreopoulos, 2007; Devereux and Hart, 2010), while in other cases there have been no returns (Oosterbeek and Webbink, 2007; Pischke and Wachter, 2008). Some increases in the CSL age have had impact on fertility decisions (Cygan-Rehm and Maeder, 2013; Black et al., 2008), while in other cases there have been no impacts (McCrary and Royer, 2011). Grenet (2013) tries to open the black box of CSL age legislation changes by comparing increases of similar sizes in Britain and France, which resulted in positive wage returns in only one country. He finds that after the reform, the share of students dropping out decreased sharply in Britain but not in France. He comes to the conclusion that a higher CSL age brings labor market advantages only if it leads to increased rates of school completion as well.

In spite of the enormous literature on the social and economic impacts of an increased CSL age, we know very little about its effects within schools: on school quality, student composition, secondary school tracking choice, or school completion rates. Learning more about its in-school effects may deepen our understanding of why CSL age reform is successful in one context but not in another. The emerging evidence on the in-school effects of increasing the CSL age is quite controversial. Higher CSL age has been shown to reduce the effort levels that teachers put into teaching (Green and Navarro Paniagua, 2012), and to increase criminal behavior of students within schools (Anderson et al., 2013). Cabus and De Witte (2011) show that increasing the CSL age decreased dropout rates in the Netherlands, while Landis and Reschly (2011) find that a higher CSL age has an effect on the timing of dropping out, but not on high school completion rates.

This paper estimates the effects that an increase of the CSL age from age 16 to 18 in 1996 had on schooling outcomes in Hungary using a regression discontinuity design (RDD) strategy. The new CSL age came into force with students starting elementary school in September 1998. Identification is based on the age of elementary school start rule. Children compliant with the age rule started elementary school under the new CSL age scheme if they were born on June 1, 1991, or later. Those compliant with the age rule and born before this date had started elementary school in the previous year, under the old CSL age scheme. Thus, a natural cutoff point occurs at this date of birth which allows me to construct a fuzzy RDD strategy to estimate the intention to treat (ITT) effects of the increase using the 2011 Hungarian Census data. Typically, children start school either according to the age rule, or they will start a year later; early school start is rare. I estimate compliance with the age rule by birth month bins from an earlier 2001 Hungarian Census, which captures the relevant cohorts during their first 3-4 grades of elementary school. On average, about one-fifth of all first graders start primary school one year later than what is determined by the age rule, and this ratio is 54% right before the cutoff (see Figure 1.2 in Section 1.2). Compliance with the age rule on the treated side of the cutoff is 87%. Using actual compliance rates as a first stage, I estimate the Local Average Treatment Effects (LATE) of increasing the CSL age in a two-sample two-stage least squares (TS 2SLS) setup (Angrist, 1990). The probability of starting elementary school under the higher CSL age scheme jumps about 0.33 at the cutoff causing LATE coefficients to be roughly 3 times as big as the ITT impacts. The almost 90%-compliance on the treated side of the cutoff is close to the situation of one-sided non-compliance. This setup gives the rare possibility to identify the Average Treatment Effect on the Non-treated (ATNT), which is close to the LATE in this special case.

To put the analysis in context, it is important to note that the Hungarian reform is unique for three reasons. First, the increased CSL age was introduced to students starting elementary school in September 1998. Thus, the first treated cohort knew already at age 6 that they would have to stay in school for two years longer. This fact allows me to test whether this information affected forward-looking decision-making as measured by secondary school choice at age 14. Second, the increase of the CSL age was mostly an administrative change with no supply-side expansion at the time when the first treated cohort reached age 16. While most papers find substantial positive returns on higher CSL age, they usually examine the effect of comprehensive education reforms, including sometimes massive school expansion (Harmon and Walker, 1995; Oreopoulos, 2007; Devereux and Hart, 2010). On the contrary, the Hungarian case allows for testing the pure marginal effects of a change in CSL age legislation. Third, the increase happens in an education system with early tracking and strong selection mechanisms (OECD, 2015). Higher CSL age in this environment may strengthen the already existing phenomenon that students from lower socioeconomic backgrounds get selected to low quality schools.

I find that as a result of the CSL age increase, children at age 14 were more likely to choose the more demanding and more beneficial 4-year academic high school track instead of the 4-year vocational training school track. The legislation change did not influence more individuals to start secondary school, but those who did decide to start were more likely to opt for an academic high school rather than a vocational training school. At the same time, those who did choose the vocational training school track were more likely to drop out under the new scheme. The data suggest two potential explanations for this adverse effect. First, the financial and human resources allocated to vocational training schools were not adequate for the sudden increase in the number of students. Second, as the students from higher socioeconomic standings, and with stronger abilities chose academic high schools instead, and lower standing, lower performing students stayed in vocational training school longer, the distribution of students in vocational training schools shifted towards lower socioeconomic status students.

The last takeaway from this analysis is that increasing the CSL age may not always be a good instrumental variable (IV) for education. It harms the monotonicity assumption of the instrument if the quality of education for some students is negatively affected by the increase (Cygan-Rehm and Maeder, 2013). The monotonicity assumption requires students to be impacted by the instrument in the same way (Angrist and Pischke, 2009). In this context, it would assume that the legislation change induced some individuals to have *more* education and for no one to have *less* education, both in terms of length, tracks, quality, and earned degrees. In the case of the Hungarian reform, this concern is valid if one wants to use the increase of the CSL age as an IV to education, as the legislation change did increase dropout rates in vocational training schools.

The remainder of the paper is organized as follows. Section 2 introduces the main data sources. The Hungarian education system and the legislation change is presented in Section 3. Section 4 presents the identification strategy and the empirical methods, and Section 5 shows the main results. Section 6 looks at the heterogeneous effects of the legislation change by parental education. Section 7 provides several robustness checks, and Section 8 shows additional evidence on the findings and the potential mechanisms behind them. Section 9 summarizes and discusses the results.

1.2 Data

As I have to refer to the data when presenting the Hungarian education context and my identification strategy, I start by briefly describing the main data sources used in this paper. Some additional data sources which are used to estimate heterogeneous effects by parental education will be presented in Subsection 1.6.1.

Four main data sources are used in this analysis: the 2001 Hungarian Census, the National Assessment of Basic Competencies (NABC) database, the Public Education Statistics (PES) of the Public Education Information System, and the 2011 Hungarian Census.

As will be detailed in Section 1.4.1, I estimate the effects of increased CSL age in a fuzzy RDD framework. Compliance with the age of elementary school entry rule creates a discontinuity in the probability of being exposed to the new CSL age scheme around a cutoff in date of birth. This jump in the probability of starting elementary school under the new CSL age scheme around the cutoff is going to be the first stage of my RDD strategy. I estimate the first stage using the 2001 Hungarian Census. The 2001 Hungarian Census data were collected in the spring of 2001 when the cohort of interest was 9-10 years old. It contains information on the birth year and month of the individuals, and, for those in school, it registers which grade of school they were attending at the time. Knowing their grade level in 2001 allows me to estimate the jump in probability of starting school under the new CSL age regime in birth year and month bins.

The NABC [1] administrative database is used to provide a robustness check to school start compliance rates estimated from the 2001 Hungarian Census. The NABC database registers the results of centrally organized low stake math and reading tests taken each year in Grades 6, 8, and 10. The NABC database provides information on the students' birth year and month. Thus, it allows estimating the probability of starting elementary school under the higher CSL age scheme in birth year and month bins, just as the 2001 Hungarian Census. The main difference between these two data sources is that while the 2001 Hungarian Census covers the whole cohort of interest, the NABC data cover a subsample of 10$^{\text{th}}$-graders only[2].

The PES of the Public Education Information System[3] is the official Hungarian school census database. It collects extensive information on schools, school programs, and students. For the cohorts and periods of interest for this paper, the PES provides aggregate level data across school cohorts and academic years. The PES data, along with the 2001 Hungarian Census and the NABC, are used to demonstrate that compliance with the age of elementary school entry rule did not change as a result of the CSL age increase (see Table 1.1).

The outcome measures used in this paper are constructed from the 2011 Hungarian Census data. While the 2001 Hungarian Census data were collected in the spring, the 2011 data were collected in October 2011, when the cohort of interest was 19-20 years old. Another important difference between the 2001 and the 2011 Censuses is that the latter contains information on the exact date of birth of individuals, including the day of birth, whereas the former does not. The 2011 Hungarian Census covers several educational outcome variables, i.e. the number of years completed successfully in the education system by school type, the highest earned degree, and whether one was in school at the time of the Census. Similarly to the 2001 Census, the 2011 data did not capture the important information on grade repetition. All educational outcome variables constructed from the 2011 Census data are defined in Table A.3 in Appendix A.

1.3 The Hungarian Education System and the Legislation Change

1.3.1 The Context of Hungarian Education and the Public Education Act (1996)

The Hungarian education system has long faced challenges in providing high quality education for students of differing backgrounds (OECD, 2015). Student achievements continue to stand below the OECD average, and the effects of socioeconomic background on test scores were noted as being among the largest in the 2012 Program for International Student Assessment (PISA) study (OECD, 2014). Free elementary school choice and early tracking have hindered equity and have caused a high variance of student achievements across schools. In an attempt to improve this situation and specifically to reduce the number of early school

[1] In English: `http://edecon.mtakti.hu/?q=node/15`, in Hungarian: `http://www.oktatas.hu/kozneveles/meresek/kompetenciameres/alt_leiras`

[2] The first NABC wave available to use is the one taken in the spring of 2006. Those starting elementary school in 1997, just before the CSL age increase, typically reached Grade 8 in the 2004/2005 academic year and Grade 10 in the 2006/2007 academic year. Those falling under the new CSL age legislation reach these grade levels a year later. Thus, the Grade 10 waves cover both those born right before, and right after, the cutoff and can be used to estimate the jump in the probability of starting elementary school under the higher CSL age scheme (see Tables A.1 in Appendix A for detailed information about the sample and estimation method).

[3] *KIR-STAT* in Hungarian.

leavers, and to close the education and employment gap between those of higher and lower socioeconomic backgrounds, the compulsory school leaving (CSL) age was increased from 16 to 18 in 1996.

Before the legislation change, students were obligated to attend school until the end of the academic year in which they turned 16. The Public Education Act (1996) increased compulsory school attendance from age 16 to age 18, requiring students to spend two more years in the education system. The new legislation was grandfathered in, as it first became binding with those students starting elementary school in September 1998. Thus, students knew already by age 6 that they had to stay in school two years longer. Although the Act introduced other measures as well, the increase in the CSL age was the only element causing sharp changes for those starting elementary school in the 1997/98 academic year versus those starting in the 1998/99 academic year. The Act also prescribed the gradual adaptation of the secondary school structure to meet the new CSL age by forcing all secondary school programs to have at least 4 grades (and thus not to end before age 18), a process that began during the 1998/1999 academic year. As a result, the first treated cohort was beginning secondary school at age 14 in 2006 at a time when the adjustments to secondary school program length had been adapted half a decade earlier.[4]

All actors within the education system supported the increase of the CSL age at the time of its enactment. However, it has been viewed controversially since the first treated cohort reached the age of 16. As it was grandfathered in, the Act pushed all implementation costs to the future government of 2008. This responsibility came about at the time that the first affected cohort reached age 16, the age at which students who would have dropped out in absence of the legislation change had to stay in school (National Institute of Public Education, 2010). Although the CSL age change and the number of potentially affected students had been predicted well in advance, the education policy at this time did not actively support the implementation of the increased CSL age. The schools and their leading bodies began to realize that they lacked the tools to handle the problems emerging in 2008 (National Institute of Public Education, 2010). The increased number of students put so much strain on unprepared schools that most school principals viewed the CSL age increase unfavorably, according to a 2009 survey (Mártonfy, 2011a). The most frequently expressed problems included that schools had no methods to engage unmotivated students in learning, that they were unable to offer a credible perspective on life to these mostly low socioeconomic standing and low-skilled students, and that they had no expertise in the development of students from troubled backgrounds (Mártonfy, 2011b). Due to the emerging problems and some other, mostly political considerations, the National Public Education Act (2011) reduced the CSL age from 18 back to 16, starting from September 2012. This paper exclusively evaluates the increase that occurred in 1996.

Despite the lapsed time since the CSL age increase was made and then retracted, there has not been any quantitative assessments on the casual effects of this legislation change to date. The only known attempt to evaluate the impact of the CSL age change was compiled by Dobos (2014) in an unpublished study. Dobos attempted to use a "discontinuity band" identification strategy but she was unable to capture any causal effects of the legislation change because her treated and control groups are too far apart in age. This paper aims to fill the gap by providing a quantitative evaluation to find a causal effect.

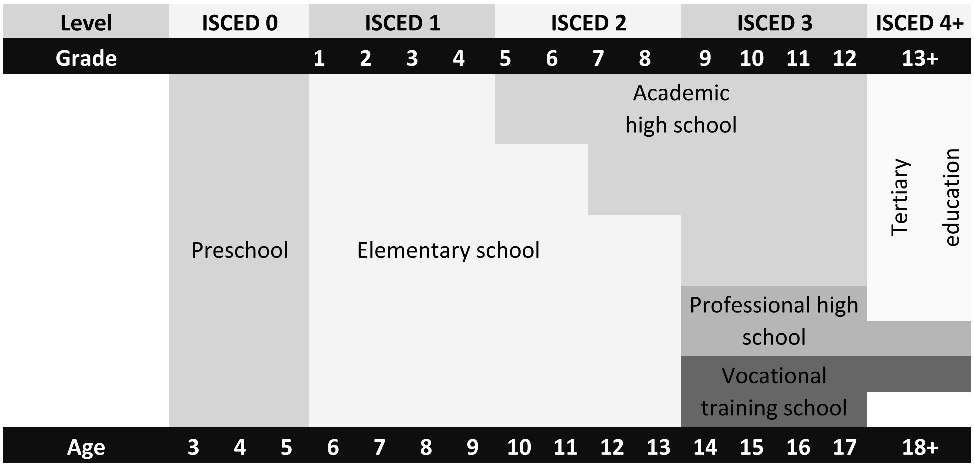
1.3.2 Compulsory Education in Hungary

Typically, Hungarian students reach CSL age while in secondary school. There are three types of secondary schools in Hungary: vocational training school, professional high school, and academic high school (see the structure of the Hungarian compulsory education system in Figure 2.1). The core programs in all types of secondary school last for 4 years during that time, both right before and right after the CSL age increase. There is a difference in the returns of vocational training school from those of the two types of high schools. Vocational education is considered as a dead end in the sense that, although theoretically it is possible, most students do not continue their studies after completing a vocational training school. The Mincer-type returns of vocational education are low but still positive in comparison to the returns of an elementary school

[4]Schooling outcomes may be affected by other schooling-related legislation changes. Such legislation may be the Child Benefit ("családi pótlék") which is a family-type cash allowance given to parents after children, either until CSL age, or, until at most age 20 if children continue their studies in the formal education system until then. In 2010, its availability conditions were tightened and it was supposed to be stopped if the child accumulated at least 50 teaching hours unjustified absence from school in one academic year. The new legislation came into force in the 2010/2011 academic year (on Aug 31, 2010), when the cohorts relevant to this paper were already 18-19 years old. Thus, the change in Child Benefit legislation could not have affected the results of this research.

degree (Kézdi, Köllő and Varga, 2009). A high school degree, on the other hand, brings substantial benefits.[5] Firstly, it is the prerequisite for starting tertiary education. Secondly, it gives high returns in the labor market, both in terms of employment probability, and wages. The average wage advantage of a high school degree is estimated to be about 25-30% higher when compared to a vocational degree, amounting to around 30,000 USD during a lifetime (Hajdú et al., 2015).

Figure 1.1: The Structure of the Hungarian Secondary Education System



Source: Horn, 2014. For the cohort of interest, preschool was compulsory from age 5. The CSL age was 16 in the case of those starting elementary school in September 1997 or earlier. The CSL age was 18 in the case of those starting elementary school in September 1998 but was then reduced to 16 again in September 2012 for students in any grades.

Prior to the CSL age increase in 2001, 13% of 15-year-olds were attending primary school, and 83% were attending secondary school right before reaching the actual CSL age of 16 that was in place at that time.[6] Out of those in secondary school, 16% were in vocational training schools, 45% in professional high schools, and 39% in academic high schools. [7]

Students decide on their secondary school tracking choice at an early age in Hungary (OECD, 2015). Some highly selective elite academic high schools already recruit top-talent students during Grade 4 and Grade 6. However, most students choose their secondary school at age 14, before entering Grade 9. The cohorts of interest to this study chose between the 4+ year vocational training school, the 4+ year professional high school and the 4+ year academic high school tracks.

Primary school starts at age 6 and has 8 grades. According to the age of elementary school start rule, compulsory schooling starts on September 1 of the same year in which one reached age 6 by May 31. Those born on June 1, or later, during the same year start elementary school one year later. Thus, those compliant with the age rule and born before June 1, 1991, entered elementary school in 1997 under the old CSL age

[5] Both high schools involve the completion of a "maturity exam" ("érettségi") at the end of Grade 12, which is similar to "Matura" or "Baccalaureat" examinations found in many European countries (Kézdi and Surányi, 2008).

[6] Data Source: own estimation from the 2001 Hungarian Census. For 4% of the sample, the information on the type of school is either missing (3%), or they are not in school due to living with disabilities (1%). See more on the data sources used in this paper in Section 1.2.

[7] Compulsory schooling obligation can be fulfilled in homeschooling as well. However, even those in homeschooling belong to a school and are supposed to be covered by education statistics as those in school. Homeschooling is rare; in the 2013/14 academic year, the share of those in homeschooling was 0.68% (Education Office of the Ministry of Human Resources, 2014).

10.14754/CEU.2016.11

scheme. Those compliant with the age rule and born on June 1, 1991, or later, entered school in 1998 under the new CSL age scheme. This discontinuity in date of birth at June 1, 1991, is the base of my identification strategy.

In addition to the age rule, the school starting year is a joint decision of parents, preschool teachers, and in some cases, pedagogical and psychological counselors employed by public pedagogical service centers[8]. The decision itself is made during preschool. At the time of the legislation change, preschool attendance was compulsory from age 5. The decision process about elementary school entry starts with an official opinion of preschool teachers about whether the child is ready to start school. In the case of any doubts, preschool teachers can ask for a "school readiness examination" from the local pedagogical service center.

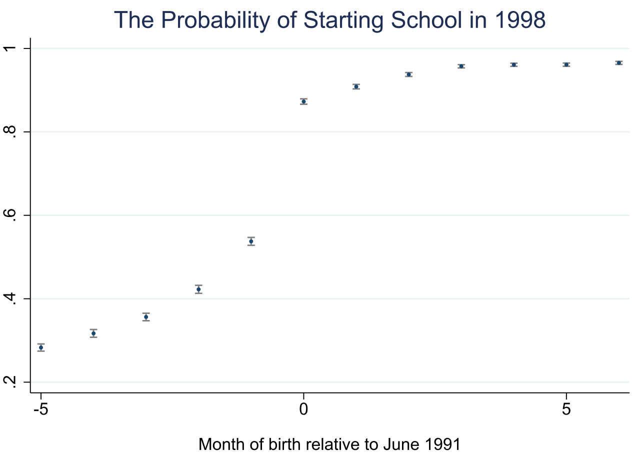
Table 1.1: Compliance with the Age of Elementary School Start Rule, PES Data

	Early starters (never-takers)	Compliers	Late starters (always-takers)	No. of students in Grade 1
Compliance by academic years, share of those starting elementary school at the given year				
1997/1998	0.02	0.80	0.18	127,214
1998/1999	0.02	0.78	0.20	125,875
1999/2000	0.01	0.78	0.21	121,424
Compliance by cohorts, share of cohort size				
Born between June 90-May 91	0.02	0.79	0.19	129,489
Born between June 91-May 92	0.02	0.78	0.20	126,294

Data Source: Public Education Statistics (PES) of the Public Education Information System (KIR-STAT). "Early starters" refers to those entering elementary school without reaching age 6 by May 31. "Compliers" refers to those entering elementary school according to the age of elementary school start rule at age 6. "Late starters" refers to those entering elementary school a year later than expected under the age rule. Individual level data are not available for this period in the PES.

On average, about 80% of a cohort starts elementary school according to the age rule, while the rest start a year later. Early school start is rare, at about 2%, according to the aggregate statistics of the PES (see Table 1.1). Around 54% of those born in May 1991, and around 87% of those born in June 1991 started school under the new CSL age scheme (see Figure 1.2). Thus, the probability of starting elementary school under the new CSL age scheme jumps 0.33 around the cutoff at June 1, 1991.

Figure 1.2: The Probability of Starting School under CSL Age 18



The average probability of starting school in 1998 among those born in 1991, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June 1991. No. of individual observations: 123,486.

[8] *Pedagógiai szakszolgálat* in Hungarian.

1.4 Identification and Empirical Methodology

1.4.1 Identification Strategy

Children compliant with the age of elementary school start rule entered elementary school under the higher CSL age scheme if they were born on June 1, 1991, or later. As mentioned in Section 1.2, the 2001 Hungarian Census registers which year of elementary school those born in 1991 were attending in 2001. Knowing their grade level in 2001 allows me to estimate the jump in probability of starting school under the new CSL age regime, assuming that grade repetition patterns did not change between 1997 and 2001. Those who started elementary school under the old CSL age scheme were in Grade 4, and those who started elementary school under the new CSL age scheme were in Grade 3 in 2001, as long as they had not repeated a grade by that time. Although the 2001 Hungarian Census does not register grade repetitions, this information is available on an aggregate level from the PES database. Table 1.2 summarizes the share of students repeating a year during Grades 1-3, and shows that the prevalence of grade repetition in these grades is low in general, and did not change in this period (National Institute of Public Education, 2006). Those who started elementary school under the old CSL age scheme in 1997 should have been in a higher grade in 2001 than those who started school a year later, and thus, they are 1.5 percentage points more likely to repeat a grade by the time of observation (see Table 1.2, 3rd row). Consequently, this method will result in estimating a 1.5 percentage point lower jump in the probability of starting school under the new scheme because I cannot distinguish between a late school start and a grade repetition.

Table 1.2: Grade Repetition in Grades 1-3, % of Students in Grade

Academic year	Grade 1	Grade 2	Grade 3	CSL age
1995/1996	4.0	1.9	1.6	16
1996/1997	3.9	2.0	1.5	16
1997/1998	3.9	1.9	1.5	16
1998/1999	4.0	1.8	1.5	18
1999/2000	3.9	1.9	1.4	18
2000/2001*	4.2	1.9	1.4	18

*The 2000/2001 is the academic year in which the 2001 Census is taken. Data Source: National Institute of Public Education, 2006. Table 4.28 in the Appendix, page 478.

The 2001 Hungarian Census captures an individual's year and month of birth, but not the day. Thus, compliance with the age rule can be estimated in birth month bins. According to this data, 54% of those born in May, and 87% of those born in June, started elementary school under the new CSL age regime (see Figure 1.2 in Section 1.2). This allows me to identify the causal effects of the CSL age increase in a regression discontinuity design (RDD) by setting a cutoff at the June 1, 1991, date of birth. As the probability of falling into the treated cohort jumps from about 0.54 to 0.87 at the cutoff, this is a fuzzy RDD setup where compliance with the age of elementary school start rule is endogenous. Being born right after rather than before the cutoff is used as an instrument for starting school under the new CSL age legislation. As a first step, I estimate the intention to treat (ITT) effects of the legislation change around the cutoff using the 2011 Hungarian Census data that captures the cohort of interest at ages 19-20. In Section 3.3, the analysis is extended to a 2-stage procedure taking compliance with the age rule into account in a first stage, estimated from the 2001 Hungarian Census as mentioned before.

I identify the ITT effects of the rise in a potential outcome framework. Let's consider the population of those born in a small neighborhood within the cutoff date and define Z as:

$$Z = 1(\text{being born on or after June 1, 1991}).$$

Z is a binary instrumental variable for the treatment, i.e. entering elementary school under the new CSL age scheme, for this population. The potential treatment indicators are then defined as

$$D(0) = 1(\text{entering elementary school under CSL age 18 if one had Z=0})$$

$$D(1) = 1(\text{entering elementary school under CSL age 18 if one had Z=1}).$$

The actual treatment indicator is $D = ZD(1) + (1 - Z)D(0)$. For those born before the cutoff, or $Z = 0$, the intended school starting date is 1997, which falls under the old CSL age of 16. However, some parents choose to withhold their child to start school one year later in 1998, which falls under CSL age 18. Therefore, before the cutoff,

$$D(0) = 0 \text{ for compliers and}$$

$$D(0) = 1 \text{ for always-takers (i.e., late starters).}$$

Both $D(0) = 0$ and $D(0) = 1$ are possible, and the choice between them is endogenous. Compliance with the age rule is 54% on this side of the cutoff (see Figure 1.2 in Section 1.2). For those born after the cutoff, or $Z = 1$, the intended school start date is 1998, under the new CSL age of 18. Therefore, after the cutoff:

$$D(1) = 1 \text{ for compliers, and}$$

$$D(1) = 0 \text{ for never-takers (i.e., early starters).}$$

As shown by the data, an early school start is rare and there is a 87% probability that D(1) = 1 occurs (see Figure 1.2 in Section 1.2).

My identification strategy is based on three assumptions. First, I assume that the instrument is exogenous: whether the student was born right before or after the cutoff is random. As the Act was introduced in 1996, five years after the relevant cohort was born, manipulation of birth because of the legislation change is not an issue. However, there is a dispute in the literature about whether children born in different months of the year are inherently different from each other with regard to their outcomes later in life (see Buckles and Hungerman, 2013 vs. Fan et al., 2014). The primary concern within the literature is related to those born during the winter as opposed to those born in spring. This literature argues that babies born in the winter are more likely to be born to less educated women, and that the winter months may not provide as favorable of an environment to a newborn as the spring. In this paper, I compare children born before and after June 1. To the best of my knowledge, the question of whether children are inherently different when born in May or in June has not been raised in the literature. In spite of this, the exogeneity assumption of the date of birth is going to be relaxed for a robustness check in Section 1.7.4.

The second assumption is the exclusion restriction: the legislation change is the only channel through which being born before or after the cutoff affects the outcomes. This is not a trivial assumption in this case. Specifically among those staying in school until the CSL age, starting elementary school at an age closer to 6 rather than closer to 7 may result in the student spending a longer or shorter time in school, independently from the legislation change (Angrist and Kruger, 1991; Hámori and Köllő, 2011). Without the legislation change, both those born in May and June 1991 would have to stay in school until the end of the 2006/2007 academic year. Those born in May, however, if compliant with the age rule, were supposed to start elementary school one year earlier. Although the CSL age increase might have created two extra years to the schooling career of those born in June, the net impact of the legislation change around the cutoff on the compliers is closer to one year. I ran two robustness checks to learn more about this issue. In Sections 5, 6 and 1.7.2, impacts estimated around the real cutoff in 1991 are compared to the same cutoffs in other years. Excluding the probability of starting a vocational training school and earning a vocational school degree, I find no effects around cutoffs in other years, the effects of starting school at a later age and spending less time in school balance each other out.[9] In Section 1.7.4, I directly control for any potential impacts around cutoffs in other years.

The third identification assumption is that the instrument is continuous and no defiers exist. On the one hand, I assume that being born after the cutoff and thus being subject to an increased CSL age did

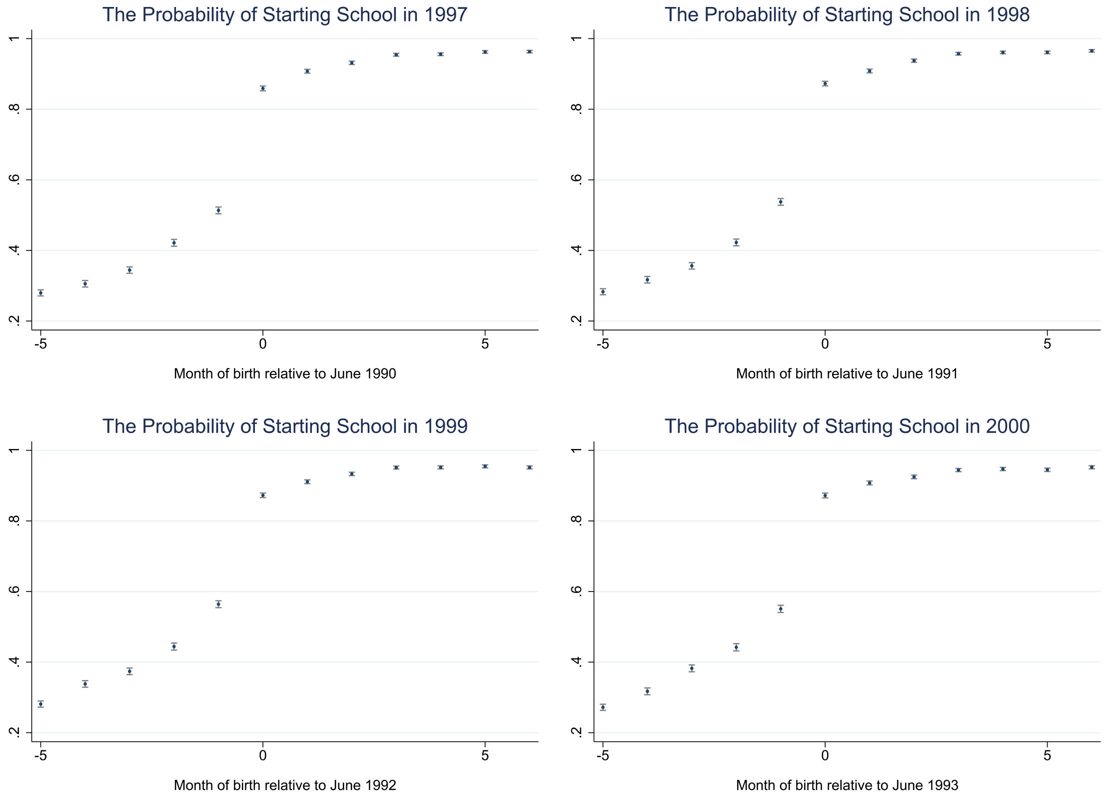
[9] I discuss this issue further in 1.7.2.

not induce anyone to start school one year earlier to avoid the extra two compulsory years in school. This assumption would theoretically be violated if parents who "dislike" schooling want to manipulate the system and therefore send their children born after the cutoff to school one year earlier, to "save" them from the extra two compulsory school years. Similarly, I also assume that none of those born before the cutoff are sent to school one year later for the purpose of being subjected to the new CSL age. I assume that always-takers on the left side of the cutoff, and never-takers on the right side of the cutoff, start school late or early because of their general preferences on what is the ideal time to start school, and not because they want to violate the age rule due to the legislation change.

Such violations are highly unlikely. First of all, the timing of the legislation change was not in favor of those who might want to manipulate school start. It was accepted in 1996, the first treated cohort started school in 1998, and they reached age 16 in 2008. To start school early in 1997, parents "disliking" school had to have been aware of the increase already in late 1996 or early 1997 to ask for an early school readiness examination in pre-school. Pre-school was compulsory from age 5 at that time, and some potential early-starters had just entered pre-school when the new legislation came out. Assuming that those parents who "dislike" school might be of a lower socioeconomic background, it is unlikely that they are informed of the details of the legislation change far enough in advance. The increase of the CSL age, as it became practically binding only 12 years later, did not receive much attention in the media in 1996. The Act made several prompt changes in the education system, and the media concentrated much more on those instead. It also seems quite unlikely that pre-school teachers suggested to parents in 1996-1997 to send their children to school early to avoid longer schooling. It sounds more reasonable that the information about longer schooling reached parents at the time that their children entered elementary school in 1998. By then it was too late to avoid longer compulsory school attendance. Furthermore, it was also unlikely that parents who "like" schooling manipulated the system by sending their children born before the cutoff to school a year later because CSL age is binding downwards. Parents who were fond of schooling could keep their children in school until whatever age they prefer, regardless of the actual CSL age legislation.

Secondly, I find no evidence in the data that such defiance occurred. As detailed in Section 1.2, I observe school starting patterns from three data sources. The first data source is the 2001 Hungarian Census, which allows me to estimate the school starting patterns of those born between 1990 and 1993, by birth month bins. Figure 1.3 shows that the share of never-takers and always-takers do not differ significantly across these four cohorts around the June 1 cutoff.

Figure 1.3: The Probability of Starting School in 1997-2000, Census Data



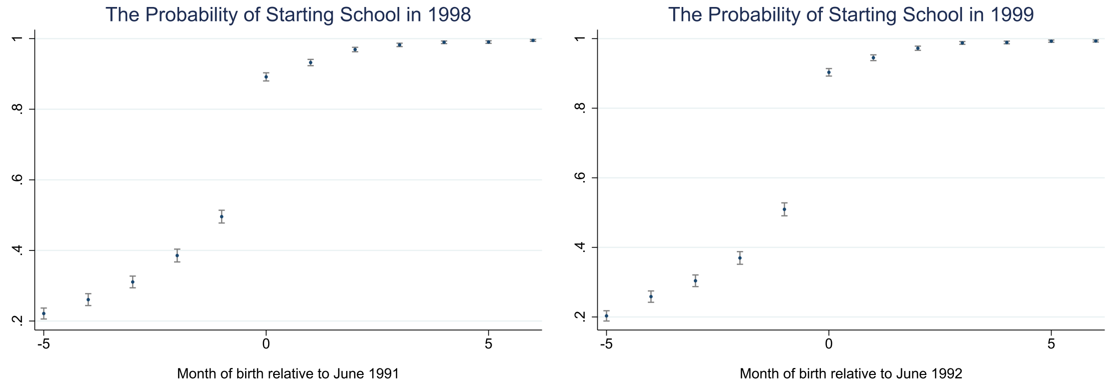
The average probability of starting school in 1997 for those born in 1990, in 1998 for those born in 1991, in 1999 for those born in 1992, and in 2000 for those born in 1993, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June. No. of individual observations: 123,719 in 1990, 123,486 in 1991, 118,500 in 1992, and 113,582 in 1993. The number of observations is decreasing in line with the number of live births.

The second data source is the individual level data of the Hungarian National Assessment of Basic Competencies (NABC) administrative database. This source allows for estimates of the probability of starting elementary school in 1998 among those born in 1991, and in 1999 among those born in 1992, for comparison.[10] Similarly to the 2001 Hungarian Census, the NABC data also suggest that the share of always-takers and never-takers are the same in both cohorts (see Figure 1.4).[11]

[10]The data do not cover the cohort born a year before, in 1990 (see in Section 1.2).

[11]Note that the size of the jump around the cutoff in the probability of starting school in 1998 is slightly larger in the NABC data than in the 2001 Hungarian Census. This difference comes from the fact that while the 2001 Hungarian Census captures the whole cohort, the NABC data cover only a subsample of those attending Grade 10 (see Section 1.2 for more details.). During this time, about 10% of students dropped out of elementary school before completing Grade 8, and the share of dropouts is even larger in Grades 9-10 in secondary school (National Institute of Public Education, 2006). Thus, those observed in the NABC data are positively selected, which results in a somewhat lower share of always-takers on the left-hand side of the cutoff.

Figure 1.4: The Probability of Starting School in 1998-1999, NABC Data



The average probability of starting school in 1998 for those born in 1991, and in 1999 for those born in 1992, plotted with the 95% confidence intervals of the means. Data source: own estimation from the 2006-2010 waves of the Hungarian Assessment of Basic Competencies Survey (NABC) taken in 10th grade. 0 on the x-axis refers to being born in June. No. of individual observations: 34,016 and 33,523, respectively.

The last data source on elementary school entry is the aggregate level administrative data of the Public Education Statistics (PES) of the Public Education Information System, which shows the share of early starters (never-takers) at 1-2%, compliers at 78-80%, and late starters (always-takers) at 18-20% in Grade 1 as stable between 1997 and 1999. This was true both across cohorts measured by date of birth, and measured by academic years (see Table 1.1 in Section 1.2). Consequently, all available data support the assumption that the legislation change did not alter school starting behavior.

1.4.2 Intention to Treat (ITT) Effects

The ITT effects of the CSL age legislation change are estimated from the 2011 Hungarian Census by using both a nonparametric and a parametric estimation strategy. **Nonparametric estimates** are generated by estimating weighted local linear regressions on both sides of the cutoff, within a certain bandwidth (Hahn et a., 2001; Imbens and Lemieux, 2008). For simplicity, weights are computed by applying a rectangular kernel function on the distance between each observation and the cutoff in terms of date of birth, measured in days. This is the standard method of RDD estimation as it has excellent properties in estimating the difference of two conditional expectations evaluated at the boundary points of the cut-off (Cheng, Fan, and Marron, 1997).

The following local linear models are estimated within a certain bandwidth:

$$y_i = \alpha_{ITT} + \beta_{ITT} * Z_i + \gamma_{ITT} * x_i + \delta_{ITT} * x_i * Z_i + \varepsilon_i,$$

where

y_i is the outcome variable;
Z_i is the instrument, which is 1 if individual i was born on June 1, 1991, or later, and 0 otherwise;
x_i is the running variable, number of days in date of birth before or after June 1, 1991, (and 0 if individual i was born on June 1, 1991); and
$x_i * Z_i$ is an interaction term of x_i and Z_i , allowing the local linear function to be different on the two sides of the cutoff.

Although recent literature agrees on using local linear regressions in this setup as the appropriate method (McCrary and Royer, 2011; Gulesci and Meyersson, 2013), there is an ongoing discussion about how to set the bandwidths. In other words, how close one has to go to the cutoff to believe that being born below or above is still random. In practice, such nonparametric estimates may be very sensitive to bandwidth choice. Traditional bandwidth selectors are usually balancing the trade-off between fit (minimizing squared errors of the local functions) and the variance of the RDD estimator, typically leading to too "large" bandwidths (Calonico, Cattaneo and Titiunik, 2014). Such optimization procedures include the family of rule-of-thumb bandwidths as in Lee and Lemieux (2010), the optimal bandwidth routine of Imbens and Kalyanaraman

(2012), and the various cross-validation procedures. If the bandwidth is too large, estimates are going to be biased as the estimated means are going to be too far from the cutoff. This paper follows a conservative strategy and uses the strictest procedure: the optimal bandwidth routine from Calonico, Cattaneo and Titiunik (2014), abbreviated as CCT in the rest of the paper, along with the 50-150% versions of the optimal bandwidths as robustness checks. The details of the procedure can be found in Appendix A.

The optimal bandwidths set by the method are 100-150 days wide below and above the cutoff, depending on the outcome variable and the sample. As the optimal bandwidth is case specific, all main results are re-estimated using a 100-day bandwidth, making the estimated coefficients comparable across subsamples (see Appendix A).

A **parametric approach** is used for one of the robustness checks on a 5-year sample of individuals born in 1988-1992 using 4th-order global polynomial models. The estimated ITT models are the following:

$$y_i = \alpha_P + \beta_{ITT,P} * t_i + f(x_i, Z_i) + \varepsilon_i$$

where

$f(x_i, Z_i)$ is a 4th-order polynomial function of the running variable, which is different on the two sides of the cutoff.

There are two reasons to complement the nonparametric analysis with parametric models. First, they can accommodate additional control variables. In particular, birth month fixed effects will capture the impacts of any potential monthly seasonality of child quality, and birth year fixed effects will capture the potential effects of business cycles. An interesting feature of the Census data (and the Hungarian health system) is that the day of the week matters with respect to the probability of being born. One is more likely to be born Tuesday through Friday rather than Saturday through Monday, and this probability difference is weakly related to the educational status of the mother.[12] This paper does not want to document this phenomenon. However, because June 1 in 1991 fell on a Saturday, day of the week fixed effects are also included in the parametric models to control for this pattern.

Secondly, although recent RDD papers use the nonparametric approach, the early literature mainly used global polynomial models (diff-in-diffs) in a similar fashion (see Table A.2 in Appendix A). Thus, I find it reassuring that both approaches lead to the same conclusion.

I estimate robust standard errors clustered by birth year and month throughout this paper. Furthermore, I test the impact of the legislation change on more outcomes and more subsamples at the same time. Testing several statistical hypotheses together increases the probability of finding significant effects by chance, known as the problem of multiple inference (Anderson, 2008). In considering a set of statistical inferences simultaneously, the probability of committing a type-I error increases. In other words, hypothesis tests that incorrectly reject the null are more likely to occur than initially intended by a single test at a time. Thus, I correct all hypotheses tests of ITT effects by estimating the number of tests done at once using the multiple testing procedure of Benjamini and Hochberg (1995). The procedure controls for the false discovery rate (FDR), which is defined as the expected share of type-I errors among all rejections. This is done by correcting upward the p-values of the tests according to two factors: the number of tests conducted together, and the relative magnitude of each p-value to the rest. A smaller original p-value of a hypothesis test results in a larger upward penalty imposed on the p-value by the procedure. I use the multiple testing p-value correction procedure in the parts of this analysis when I test whether an increase in the CSL age has an effect on a set of outcome variables. Original, uncorrected standard errors are reported along with corrected p-values for these cases. However, the procedure lowers the power of the tests by each additional hypothesis, and therefore increases the probability of type-II errors or false non-rejections. Thus, I do not use multiple testing correction when I estimate the magnitudes of the already established significant relationships (see LATE in the next subsection), nor for the robustness checks.

1.4.3 Local Average Treatment Effects (LATE)

I estimate the LATE of the legislation change around the cutoff using information about compliance with the age of elementary school entry rule. As the first stage is estimated from the 2001 Hungarian Census in birth month bins while the reduced from is estimated from the individual level data of the 2011 Hungarian

[12] Own estimation from the 2011 Census.

Census, this is a two sample procedure in the fashion of Angrist (1990). A two-sample 2SLS strategy is valid only if both samples are taken from the same population (Angrist, 1990), which is indeed the case in this analysis.

Based on the data, children typically start school either according to the age rule at age 6 or one year later, at age 7. Only a small fraction of children go to school earlier than that (1-2% of a whole year cohort). Thus, compliance with the age of elementary school start rule is over 90% on the treated side of the cutoff, and non-compliance is relevant on the non-treated side only (see Figure 1.2 in Section 1.2). This is close to the situation of one-sided non-compliance, in which case the Average Treatment Effect on the Non-treated (ATNT) is close to the LATE (Angrist and Pischke, 2009). This fact makes estimating LATE even more interesting as it is quite rare to identify ATNT in practice. ATNT is the effect measured on the compliers on the non-treated side of the cutoff, i.e. it is interpreted as the effect on those who born before June 1, 1991, and started school in 1997, under the old CSL age scheme.

The LATE are estimated as the size of the ITT effects over the jump in the probability of being treated; thus, they are going to be roughly 3-times as large as the ITT effects. The 2-stage procedure simply re-scales the magnitude of the ITT effects.

Formally, the 2-stage procedure in the nonparametric approach involves estimating the same ITT equation on the individual-level data of the 2011 Hungarian Census just as before:

$$y_i = \alpha_{ITT} + \beta_{ITT} * Z_i + \gamma_{ITT} * x_i + \delta_{ITT} * x_i * Z_i + \varepsilon_i,$$

and, estimating the first stage equation using the same bandwidth as in the case of the ITT equation:

$$p\hat{98}_m = \alpha_{FS} + \beta_{FS} * Z_i + \gamma_{FS} * x_i + \delta_{FS} * x_i * Z_i + u_i,$$

where:
$p\hat{98}_m$ stands for the probability of starting elementary school under the new legislation, estimated in birth year and month bins from the 2001 Hungarian Census.

The LATE coefficient is estimated as:

$$\hat{\beta}_{LATE} = \frac{\hat{\beta}_{ITT}}{\hat{\beta}_{FS}}$$

Following Bjorklund and Jantti (1997), a bootstrap method is used to calculate the 95% confidence intervals of estimated LATE coefficients. First, I drew a birth year and month stratified bootstrap sample from the 2001 Hungarian Census data, and estimate the average probability of starting school under the new CSL age scheme in birth year and month bins ($p\hat{98}_{m,1}$). Then, I drew a birth year and month stratified bootstrap sample from the 2011 Hungarian Census data, from which I estimate the ITT parameter ($\hat{\beta}_{ITT,1}$), and using ($p\hat{98}_{m,1}$), the first stage parameter, $\hat{\beta}_{FS,1}$. Then, I estimate a LATE parameter as $\hat{\beta}_{LATE,1} = \hat{\beta}_{ITT,1}/\hat{\beta}_{FS,1}$. Repeating these four steps $B = 1000$ times yields an empirical distribution of $1,000$ estimated LATE coefficients ($\hat{\beta}_{LATE,1} - \hat{\beta}_{LATE,1000}$). The 95% confidence interval of $\hat{\beta}_{LATE}$ is then set as the middle 95% of this empirical distribution.

Table 1.3: Educational Outcomes of Those Born in Spring vs. Summer 1991

	Born March-May 1991	Born June-Aug 1991	Difference (t-test p-values)
A. Probability of being in school at the time of the Census			
Any school	0.619	0.679	0.060 (0.000)
B. Number of years completed successfully in the school system			
No. of school years	12.2	12.0	-0.264 (0.000)
C. Probability of starting a secondary school			
Any secondary	0.916	0.918	0.002 (0.371)
Vocational training school	0.220	0.206	-0.014 (0.000)
Professional high school	0.343	0.350	0.008 (0.045)
Academic high school	0.454	0.466	0.012 (0.003)
No. of observations	33,350	33,919	
D. Probability of dropping out of secondary school conditional on starting one			
Any secondary school	0.045	0.049	0.004 (0.019)
	[30,375]	[30,939]	
Vocational training school	0.122	0.140	0.018 (0.002)
	[7,347]	[6,972]	
Professional high school	0.022	0.029	0.006 (0.003)
	[11,446]	[11,941]	
Academic high school	0.016	0.016	0.001 (0.690)
	[15,095]	[15,783]	
E. Probability of earning a secondary degree			
Any secondary	0.814	0.790	-0.024 (0.000)
Vocational training school	0.144	0.123	-0.021 (0.000)
Professional high school	0.258	0.244	-0.014 (0.000)
Academic high school	0.414	0.425	0.011 (0.004)
No. of observations	33,350	33,919	

Data Source. own estimation from the 2011 Hungarian Census. The table compares the average educational outcomes of those born 3 months before and after the cutoff at June 1, 1991. Two-tailed t-test p-values are in parenthesis. In block D, no. of observations are in brackets.

1.5 Results

Results are presented based on the following logic: general ITT effects are presented on the number of completed school years in Section 1.5.1, on school choice in Section 1.5.2, on the probability of dropping out of secondary school in Section 1.5.3, and on the probability of earning a secondary degree in Section 1.5.4. I estimate LATE in Section 1.5.5, and heterogeneous ITT effects with respect to gender and ethnic minority (Roma) status in Section 1.5.6.

1.5.1 The Implementation of Increased CSL Age

This subsection investigates whether those born after the cutoff stayed in school longer than those born before the cutoff. The number of completed years in school cannot be compared directly across these two groups because about two-thirds of the estimation sample is still in school at age 20, when their schooling outcomes are observed. The share of those still in school is higher among those born right after the cutoff (62% among those born before vs. 68% among those born after the cutoff, see Table 1.3, Block A). Some of this difference may be caused by the CSL age change, and some may come from the fact that they started school later. Until the time of the Census, those born after the cutoff could have spent fewer academic years in the school system by design than those born before the cutoff (13 vs. 14 years). To take into account the share of those still in school, the effects of the CSL age change on the number of successfully completed years is estimated in an exponential survival model with right-side censoring at the time of the Census. Table 1.4 shows that increasing the CSL age decreased the probability of exiting school at any point in time with 8%

on average, among those still in school at that time. In other words, the legislation change made students born after the cutoff more likely to stay in school longer.

Table 1.4: Effects on the Number of Successfully Completed School Years

	ITT effects	Robust clustered SE's	Multiple-testing corrected p-values	No. of obs.	Bandwidth (in days)
Probability of school exit	-0.080***	0.017	0.000	67,971	100

The probability of school exit is estimated in an exponential survival model, controlling for the same linear function of the running variable below and above the cutoff as in the case of local linear regressions, using a 100-day bandwidth. Negative coefficients indicate completing more years in school. Standard errors are robust and clustered by birth year and month. P-values are corrected by the number of hypothesis tests (13) done together in Tables 1.5 - 1.7 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after the multiple testing correction.

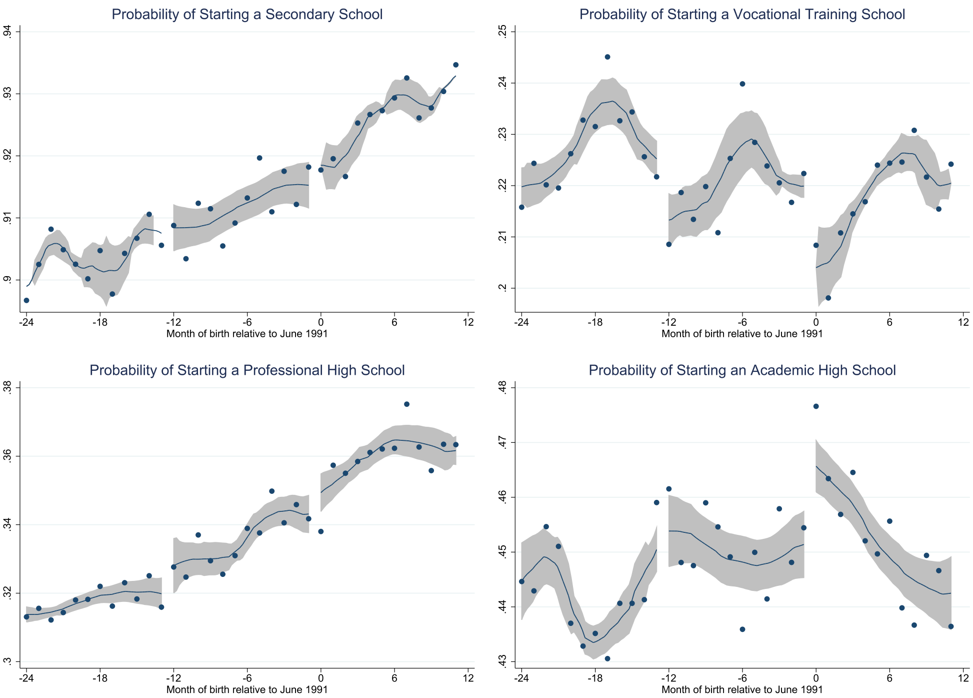
1.5.2 Effects on School Choice

The unique feature of the Hungarian reform is that it allows an estimation of the effects of the higher CSL age on secondary school tracking choice made at age 14. Figure 1.5 plots the average share of those starting each type of secondary school in birth year and month. On average, the majority of children (91%) started secondary school before the increase[13]. The share of those starting a secondary school is increasing in time, and the same is true for those starting a professional high school. The share of those starting a vocational training school or an academic high school is fluctuating between 20-24% and 43-47% during these years, respectively. To test whether the seasonality in the data is related to the cutoff, Figure 1.5 plots local linear regression functions below and above two cutoffs: the real cutoff in 1991, and the same cutoff in the previous year, in 1990. Indeed, the share of students starting a vocational training school decreased around the cutoff in 1990 as well as in 1991. For the rest three outcome variables I find so significant break around the cutoff in the previous year.

The increase of CSL age did not affect the probability of starting secondary school around the real cutoff in 1991 (see Table 1.5 and Figure 1.5). However, it did affect the choice of school tracks, even though all tracks were at least 4 years in length both before and after the increase. Students on average were 1.5 percentage points more likely to choose the academic high school track under the higher CSL age scheme.

[13]Data Source: own estimation from the 2011 Hungarian Census.

Figure 1.5: Effects on School Choice



Local linear regressions fit separately below and above the cutoff in 1990 and in 1991 on birth year and month averages, using the rule-of-thumb bandwidth of lpoly in Stata with 90% confidence intervals. Month of birth -12 represents June 1990, 0 represents June 1991. No. of individual observations: 349,925.

Table 1.5: Effects on School Choice

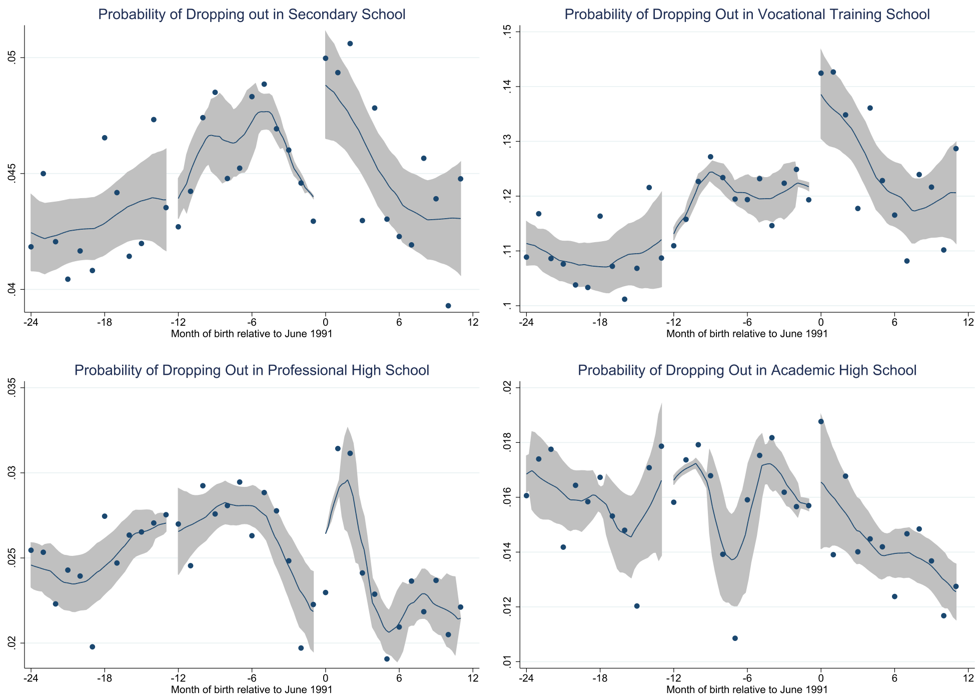
Effect on the probability of finishing at least one year in a secondary school					
	ITT effects	Robust clustered SE's	Multiple-testing corrected p-values	No. of obs.	CCT bandwidth (in days)
Any secondary school	-0.000	0.003	0.867	102,616	152.6
Vocational training school	-0.017*	0.006	0.066	80,505	118.2
Professional high school	-0.002	0.007	0.775	103,292	153.0
Academic high school	0.015*	0.006	0.074	86,292	127.6

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Standard errors are robust and clustered by birth year and month. P-values are corrected by the number of hypothesis tests (13) done together in Tables 1.5- 1.7 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after multiple testing correction. The same results estimated by a 100-day bandwidth are in Table A.5 in Appendix A.

1.5.3 Effects on Dropout Rates

The increase in CSL age causes an increase in the probability of dropping out of secondary school by 0.8 percentage points (see Table 1.6 and Figure 1.6). This effect is mostly driven by a 2.2 percentage point increase in dropout rates in vocational training schools.

Figure 1.6: Effects on Dropout Rates



Local linear regressions fit separately below and above the cutoff in 1990 and in 1991 on birth year and month averages, using the rule-of-thumb bandwidth of lpoly in Stata with 90% confidence intervals. Month of birth -12 represents June 1990, 0 represents June 1991. No. of individual observations: 314,933, 78,733, 114,076, and 156,416, respectively. Note that dropout rates are decreasing after the cutoff in all school types, parallel to the increasing probability of still being in school.

Table 1.6: Effects on Dropout Rates

	Effect on the probability of dropping out of ...				
	ITT effects	Robust clustered SE's	Multiple-testing corrected p-values	No. of obs.	CCT bandwidth (in days)
Any secondary school	0.008**	0.003	0.010	82,615	137.5
Vocational training school	0.022***	0.004	0.001	25,525	174.1
Professional high school	0.008*	0.004	0.085	32,570	138.7
Academic high school	0.003	0.002	0.177	48,077	156.9

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Standard errors are robust and clustered by birth year and month. P-values are corrected by the number of hypothesis tests (13) done together in Tables 1.5- 1.7 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after multiple testing correction. The same results estimated by 100-day bandwidths are in Table A.6 in Appendix A.

1.5.4 Effects on School Completion

Figure 1.7 plots the probability of earning a secondary degree in birth year and month bins. All of these functions are decreasing with date of birth: the younger is an individual in 2001 the less time s/he had to earn a secondary degree.

As a result of increasing the probability of choosing the academic high school track, higher CSL age increased the probability of earning an academic high school degree by 1.7 percentage points (see Table 1.7 and Figure 1.7).

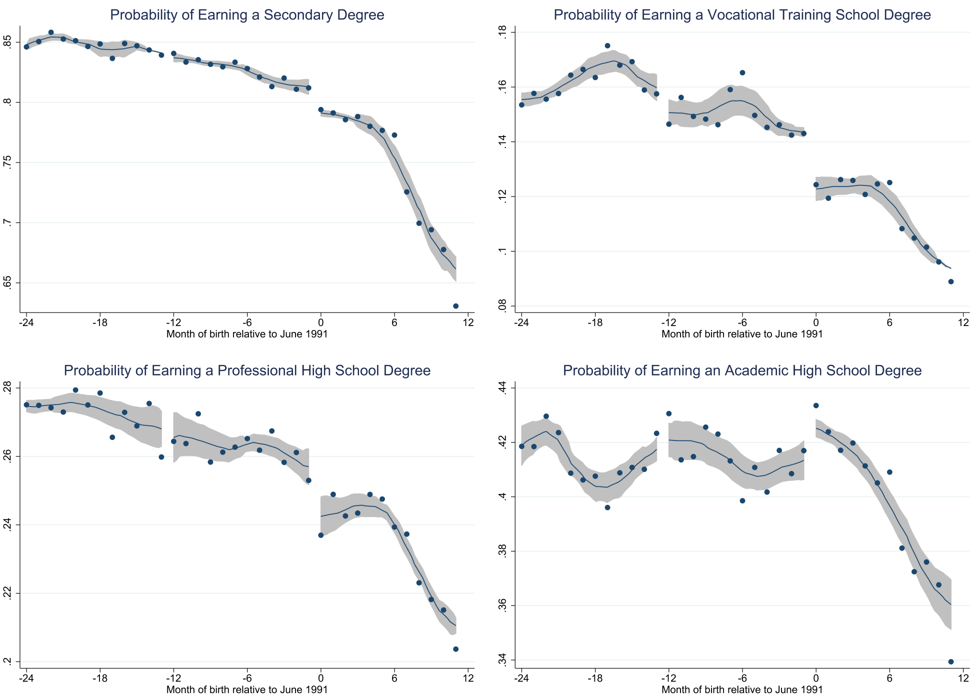
Table 1.7: Effects on School Completion

Effect on the probability of earning a secondary degree					
	ITT effects	Robust clustered SE's	Corrected p-values	No. of obs.	Bandwidth (in days)
Any secondary school degree	-0.015***	0.000	0.001	62,682	92.4
Vocational training school degree	-0.019***	0.002	0.005	70,804	104.4
Professional high school degree	-0.013*	0.006	0.085	83,827	123.5
Academic high school degree	0.017***	0.002	0.001	76,374	112.7

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Robust standard errors clustered by birth year and month are in parentheses. Observation numbers in the local regression are in brackets. Bandwidths in days are in braces. P-values are corrected by the number of hypothesis tests (13) done together in Tables 1.5- 1.7 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after multiple testing correction. The same results estimated by 100-day bandwidths are in Table A.7 in Appendix A.

At the same time, due to the fact that higher CSL age increases dropping out from this school type, the change decreased the probability of gaining a vocational degree as well. This effect is -1.9 percentage points large (see again Table 1.7 and Figure 1.7). Similarly, increased CSL age decreased the probability of gaining a professional high school degree as well.

Figure 1.7: Effects on Earning a Secondary Degree



Local linear regressions fit separately below and above the cutoff in 1990 and in 1991 on birth year and month averages, using the rule-of-thumb bandwidth of lpoly in Stata with 90% confidence intervals. Month of birth -12 represents June 1990, 0 represents June 1991. No. of individual observations: 349,925. Note that the decreasing trend comes from the fact that later a student was born the less likely s/he is to have earned a secondary degree by 2011.

The overall effect of the legislation change on the probability of gaining any secondary degree is significantly negative, -1.5 percentage points. My results suggest that the increased CSL age induced a selection mechanism. Some students chose and completed academic high school instead of vocational training school or professional high school. Meanwhile, those who were unable to upgrade their school choice and went to vocational training school ended up being less likely to earn a secondary school degree due to the increased dropout rates. I provide further evidence to support this hypothesis in Section 8.

1.5.5 Local Average Treatment Effects (LATE) of the Legislation Change

The three main results of this paper is that higher CSL age increased the probability of starting an academic high school, earning an academic high school degree, and dropping out of vocational training school. This subsection estimates the Local Average Treatment Effects (LATE) effect of the legislation change on these three outcomes to give an estimate on the magnitude of these effects around the cutoff. LATE estimates take into account that the jump in treatment probability at the cutoff is less than 1. Practically, they are estimated as the size of ITT effects over the first stage, which is the exact magnitude of the jump.[14] The LATE of the CSL age increase on the probability of starting an academic high school is 4.6 percentage points, while on the probability of earning an academic high school degree it is 5.5 percentage points (see Table 1.8). These local effects on the compliers are about three times as large as the ITT impacts. The average probability of starting an academic high school in the control group is 0.454 (see Table 1.3); thus, the LATE is 0.046/0.454=10%. Similarly, the average probability of earning an academic high school degree is 0.414; and, the higher CSL age increased the probability of earning an academic high school degree by 0.055/0.414=13% among the compliers.

Table 1.8: LATE on Starting an Academic High School, Earning an Academic High School Degree and Dropping Out of Vocational Training School

	Starting an academic high school	Earning an academic high school Degree	Dropping out of vocational training school
ITT effect	0.015**	0.017***	0.022***
	(0.006)	(0.002)	(0.004)
First stage (jump in the probability of starting school under the CSL age 18)	0.321***	0.311***	0.333***
	(0.025)	(0.022)	(0.025)
	[2,110.7]	[2,515.1]	[792.8]
Wald estimate of the LATE coefficient	0.046	0.055	0.065
95% bootstrapped confidence intervals of the LATE coefficient	0.004 - 0.086	0.008 - 0.100	0.015 - 0.114
No. of obs	86,292	76,374	25,525
Bandwidth (in days)	127.6	112.7	174.1

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Wald estimates are calculated as the ratios of ITT coefficients over the first stage coefficient (the size of the jump). Robust standard errors clustered by birth year and month are in parentheses, first stage F-statistics are in brackets. A */**/*** indicates significance on 10%/5%/1% level. The bootstrapped confidence intervals of the LATE coefficients are constructed following Bjorklund and Jantti (1997). First, a birth year and month stratified bootstrap sample is drawn from the 2001 Hungarian Census data, and birth year and month school entering probabilities ($\hat{p98}_m$) are estimated. Then, a birth year and month stratified bootstrap sample is drawn from the 2011 Hungarian Census data, from which the LATE parameters ($\hat{\beta}_{LATE,1}$) are estimated using the school entering rates estimated in the first step. Repeating these two steps $B = 1,000$ times yields an empirical distribution of $1,000$ estimated LATE coefficients ($\hat{\beta}_{LATE,1} - \hat{\beta}_{LATE,1000}$). The 95% confidence interval of $\hat{\beta}_{LATE}$ is set as the middle 95% of this empirical distribution.

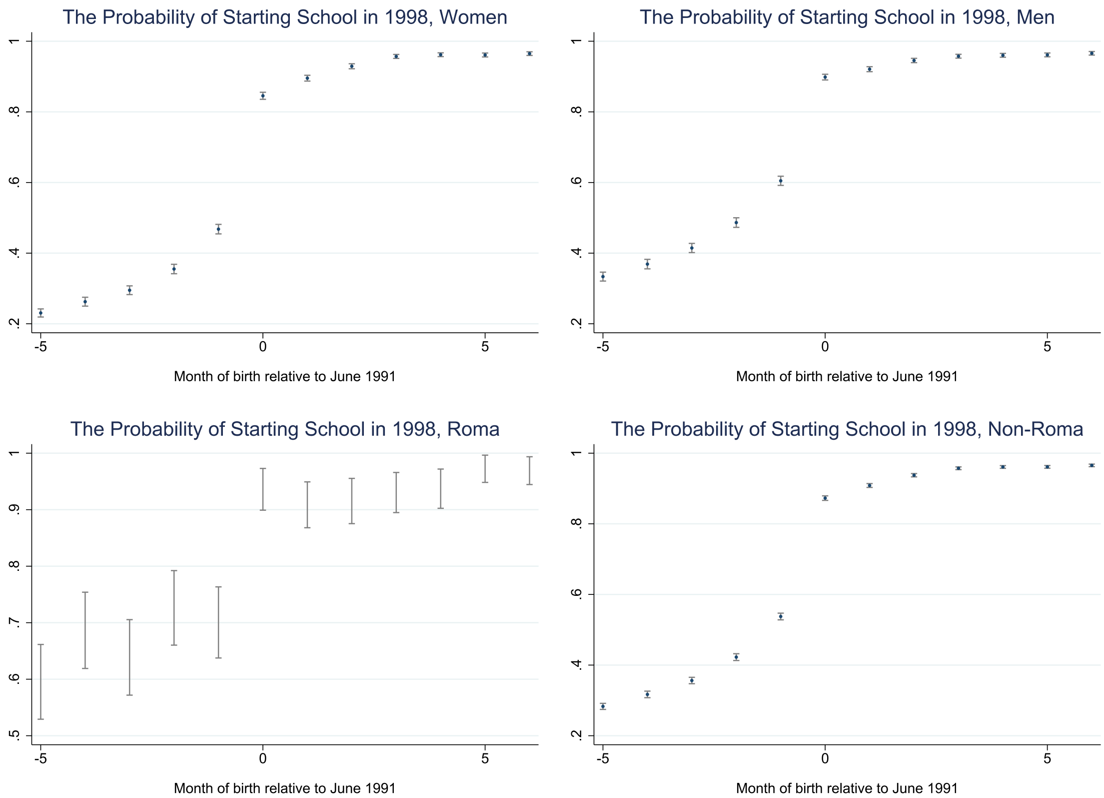
1.5.6 The Heterogeneity of the Effects of Gender and Roma Ethnicity

This subsection estimates the heterogeneous ITT effects of increased CSL age by gender and Roma ethnicity. The Roma are the largest ethnic minority in Hungary. The population of Roma in the 2011 Census is 315,583. However, based on earlier Roma studies, the actual number of the Roma population is estimated to be about two times larger than this (Hablicsek, 2007). Measurements on "being Roma" are not straightforward, which in part explains why the Census found half as many Roma people. Still, as being Roma is highly correlated

[14] First stage estimates are highly significant with high F-statistics. The reason for the high F-statistics, besides the strength of the first stage as demonstrated on Figures 1.2 and 1.3, is the following. F-statistics are estimated as $F = \frac{\frac{RSS_1 - RSS_2}{p_2 - p_1}}{\frac{RSS_2}{n - p_2}}$, where RSS_1 refers to the Residual Sum of Squares from a model of regressing the probability of starting school under the new CSL age scheme in month of birth bins, $\hat{p98}_m$, on a constant (model 1); RSS_2 refers to the Residual Sum of Squares from the first stage model, $\hat{p98}_m = \alpha_{FS} + \beta_{FS} * Z_i + \gamma_{FS} * x_i + \delta_{FS} * x_i * Z_i + u_i$, as detailed in Section 1.5.5 (model 2); p_1 and p_2 refer to the number of parameters estimated in the two models; and, n is the number of observations. Model 2 is a fully saturated model; thus, it provides high Explained Sum of Squares and low Residual Sum of Squares. Consequently, the value of F-statistics are going to be high. Due to this fact, looking at the raw data on Figure 1.2 provides a more credible support for the strength of the first stage.

with disadvantage, poverty, and poor access to public services, such as education, it makes sense to examine the heterogeneous effects of the CSL age increase with respect to Roma status.

Figure 1.8: The Probability of Starting School under CSL Age 18



The average probability of starting school in 1998 among those born in 1991, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June 1991. No. of individual observations: 60,302, 63,183, 4,626 and 113,102, respectively.

Figure 1.8 shows the first stage in these groups. The jump in the probability of starting school under the new CSL age scheme around the cutoff is higher for women (0.38) than for men (0.29), and it is highly significant for both. For the Roma, on the other hand, the data is much noisier. The size of the jump is 0.24, with large 95% confidence intervals of the means below and above the cutoff. The first stage of non-Roma individuals looks about the same as for the total sample.

The positive effect of the legislation change on starting an academic high school and earning an academic high school degree are of the similar magnitudes for women and men; however, they are significant in the case of women only (see Table 1.9). The negative effect of the increase, on the other hand, is present for men only. While men born around the cutoff are 4.2 percentage points more likely to drop out of vocational training school, the same impact on women is at 0.8 percentage points and not significantly different from zero.

Table 1.9: The Heterogeneity of the Effect by Gender and Roma Ethnicity (ITT effects)

	Starting an academic high school	Earning an academic high school degree	Dropping out of vocational training School
	The heterogeneity of the ITT effect by gender		
Women	0.024***	0.021***	0.008
	(0.004)	(0.004)	(0.013)
	{0.003}	{0.004}	{0.714}
	[32,744]	[32,744]	[5,244]
Men	0.020	0.015	0.042***
	(0.010)	(0.008)	(0.010)
	{0.177}	{0.175}	{0.009}
	[34,515]	[34,515]	[9,075]
	The heterogeneity of the ITT effect by Roma ethnicity		
Roma	-0.001	0.005	0.014
	(0.014)	(0.006)	(0.016)
	{0.949}	{0.612}	{0.630}
	[3,168]	[3,168]	[1,032]
Non-Roma	0.024***	0.020***	0.029***
	(0.004)	(0.003)	(0.006)
	{0.004}	{0.003}	{0.006}
	[61,574]	[61,574]	[12,701]
Information on ethnicity is missing	-0.004	-0.013	0.044
	(0.036)	(0.025)	(0.023)
	{0.918}	{0.702}	{0.179}
	[2,517]	[2,517]	[586]

Local linear kernel regressions using a 100-day bandwidth. Linear probability models. Robust standard errors clustered by birth year and month are in parentheses, p-values corrected by the fact that 15 hypotheses are tested together in this table using the FDR multiple testing procedure by Benjamini and Hochberg (1995) are in braces, number of observations are in brackets. A */**/*** indicates significance on 10%/5%/1% level after the multiple testing procedure.

Table 1.9 shows that the impact of the legislation change on the probability of starting an academic high school and earning an academic high school degree among the Roma is close to zero: -0.1 and 0.5 percentage points, respectively. The impact of higher CSL age on dropping out of vocational training school seems to be there for Roma individuals as well; however, the effect is not significant, probably due to the low sample size.

1.6 Heterogeneous Effects by Parental Education

1.6.1 Data Sources Used to Estimate Parental Education Data

This section will look at heterogeneous effects by parental education. For this, I use information on the education status of the parents from both the 2001 and the 2011 Censuses. Both Census data are collected at household and personal levels at the same time. They do not contain explicit information on parent-child relationships; however, they have information about the children of all individuals, including both mothers and fathers. The data register the year and month of birth of their first three children in 2001 and their first five children in 2011. This allows for linking the individuals of interest to their biological parents if they live in the same households. In the relevant cohort of those born in 1991, 99% of all individuals from the 2001 Census and 72% of all individuals from the 2011 Census live with at least one of their parents. For these subsamples, the education status of the parents is directly observed.

In order to gain more comprehensive coverage of parental education levels of the relevant cohort in the 2011 Census, I look at their Live Birth Records data. These records contain all live births in Hungary, and contain information about the exact date of birth, the residence of the parents, and several parental characteristics such as the exact birth date, and educational status, of the mother. I link the Live Birth Records data individually to the Census using two methods.

The first method (method 1) links the two databases by a two-level information comparison approach. In the case of those still living with their parents, the linking procedure uses all available information appearing in both datasets, including the date of birth of parents. In the case of those not living with their parents anymore, it uses maternal residence and exact date of birth only. With this method, 75% of those born in 1991 can be linked to the Census, out of which 80% live with their parents. The share of those living with their parents in the linked sample ratio is slightly higher than in the original Census sample of those born in 1991 (80% vs. 72%), because those living at home are easier to link.

For a potentially more comprehensive coverage of those living in smaller than 50,000-inhabitant settlements, the second linking method (method 2) links the two databases without using parental information and instead only uses residence at birth, date of birth, and gender. With this procedure, the data of an individual in the two databases can be linked together if on one given day, in one given settlement, no more than one boy and one girl were born. Under this method, 76% of those born in 1991 can be linked to the Census data.

Thus, three types of parental education data will be used in the paper. Main results are estimated using:

1. the educational status of the mother when the individual of interest was born, linked to the Census using method 1; while robustness checks are estimated using

2. the educational status of the mother when the individual of interest was born, linked to the Census using method 2; and

3. the educational status of the head of the household if the individual of interest lives with at least one of his/her biological parent.

1.6.2 Children of Less Educated Parents Are the Most Likely to Be Affected by the Increase

The CSL age increase has a direct effect on those for whom the new constraint is binding, i.e. who would otherwise exit school before reaching age 18. A potential subgroup of such students consists of the children whose mothers are less educated, i.e. finished primary school at most. The 2011 Hungarian Census data show that in the cohort born in 1990, right before the legislation change, children of less educated mothers completed on average 11.0 school years in formal primary and secondary education, while this average is close to, or above, 12.0 in all other parental groups (see Table A.4 in Appendix A).

Children of less educated mothers lag behind in all educational outcomes: they are more than 20 percentage points less likely to earn a secondary degree in general, and around 40 percentage points less likely to gain a high school degree in particular, than children of educated mothers. In fact, closing this gap was the explicit intention of the legislation change (Mártonfi, 2011a). Therefore, in the forthcoming analysis, heterogeneous effects are estimated with respect to maternal education, and special attention is paid to results estimated for children of less educated mothers.

The potential heterogeneity of the effects of an educational legislation change with respect to family background is well established both in theory and in practice. Oosterbeek and Van Ophem (2000) show that children from low socioeconomic backgrounds decide on their schooling investment and consumption along different parameters than others. They discount future returns more heavily, and have lower abilities in, and lower preferences for, schooling. Meghir and Palme (2005) find that a major educational reform in Sweden had insignificant average, but significant heterogeneous, effects. They estimate positive wage returns in the case of children with unskilled fathers, and negative wage returns in the case of children with skilled fathers. They argue that the reform might have decreased the quality of education for this group - hence the negative effect.

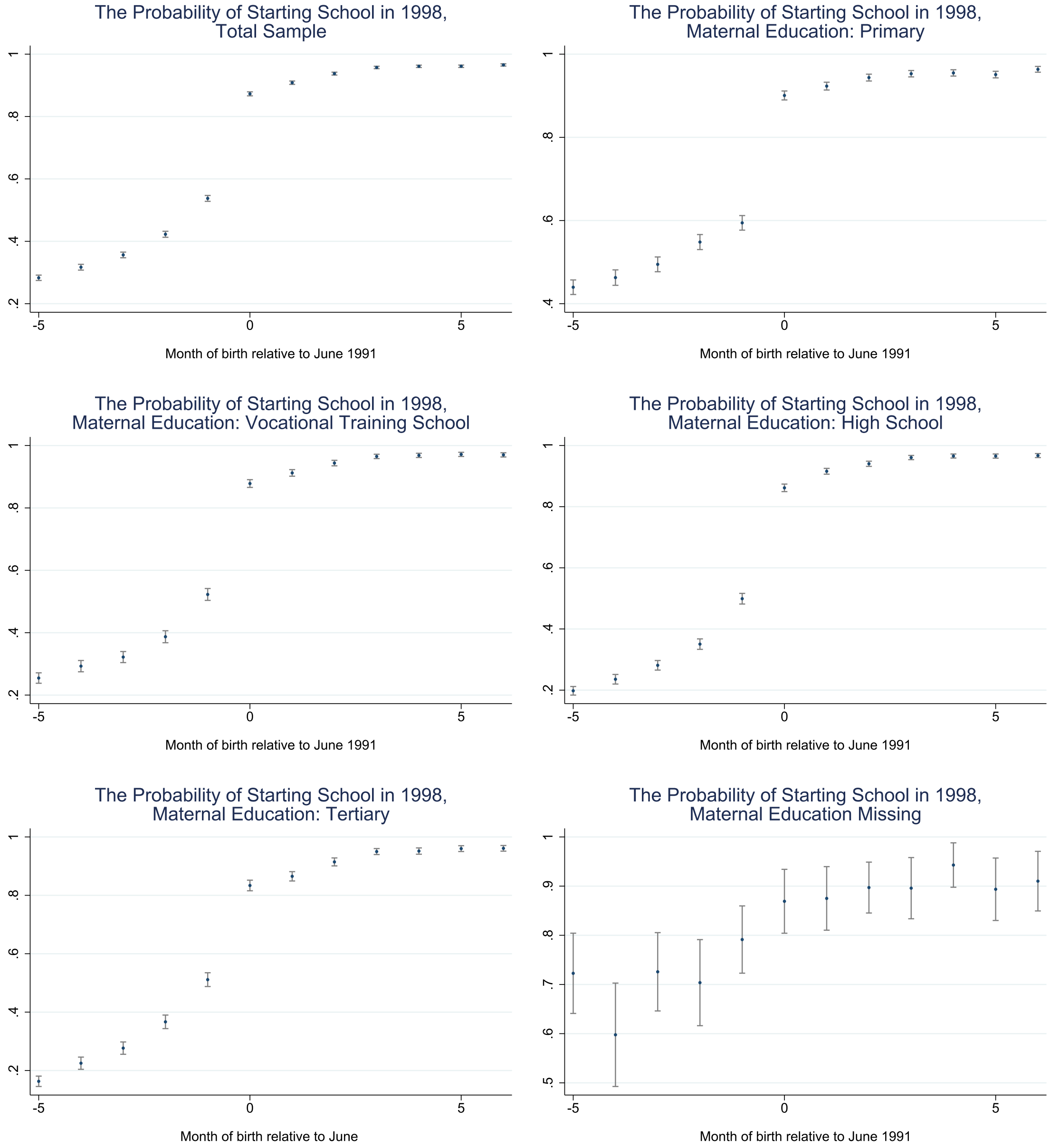
1.6.3 The First Stage by Maternal Education

The jump in the probability of starting school under the new CSL age scheme is significant in all groups of maternal education, and its size varies between 0.31-0.36 (see Figure 1.9). In the group of 1,308 observations out of the 123,486 individuals born in 1991, maternal education data is not available in the 2001 Hungarian Census. In their case, school starting information is extremely noisy and a significant first stage relationship

10.14754/CEU.2016.11

cannot be established (see Figure 1.9, second graph in the third row). The stability of the first stage across cohorts born between 1990 and 1993 are presented in Figures A.1 - A.5 in Appendix A.1.

Figure 1.9: The Probability of Starting School under CSL Age 18, by Highest Maternal Education



The average probability of starting school in 1998 among those born in 1991, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June 1991. No. of individual observations: 123,486; 35,979; 30,979; 36,197; 19,603; and 1,308; respectively.

1.6.4 The Implementation of Increased CSL Age

Heterogeneous effects by parental education are presented in the following four subsections in the same way as general effects are presented in Sections 1.5.1-1.5.5. The impact of increased CSL age on the number of completed school years is higher than average in the three lowest maternal education groups (see Table 1.10). Interestingly, the effect on schooling duration is the highest among children of mothers holding a high school degree.

Table 1.10: Effects on the Number of Successfully Completed School Years

	Total sample	Mother's highest education at giving birth: Primary	Vocational	High school	Tertiary	Missing
Probability of school exit	-0.080***	-0.095***	-0.130***	-0.177***	0.020	-0.001
	(0.017)	(0.023)	(0.019)	(0.047)	(0.065)	(0.019)
Corrected p-values	0.000	0.001	0.000	0.002	0.826	0.965
No. of obs.	67,971	15,588	12,097	16,697	5,950	17,127

The probability of school exit is estimated in exponential survival models, controlling for the same linear function of the running variable below and above the cutoff as in the case of local linear regressions, using a 100-day bandwidth. Negative coefficients indicate completing more years in school. Robust standard errors clustered by birth year and month are in parentheses. P-values are corrected by the number of hypothesis tests (78) done together in Tables 1.10 - 1.13 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after multiple testing correction.

1.6.5 Effects on School Choice

The unique feature of the Hungarian reform is that it allows for an estimation of the effects of the higher CSL age on secondary school tracking choices made at age 14. On average, the majority of children (91%), did start a secondary school track before the CSL age increase. However, this proportion was at only 80% among children of less educated parents (see Table A.4 in Appendix A). The increase of CSL age did not affect the probability of starting secondary school, neither in general, nor among children of less educated mothers. However, it did affect the choice of school tracks made at age 14, even though all tracks were at least 4 years in length both before and after the increase. Children of less educated mothers were 4.9 percentage points less likely to choose a vocational training school, and 4.1 percentage points more likely to choose an academic high school track under the higher CSL age scheme (Table 1.11). The effects on choosing academic high schools across maternal education groups are monotonic decreasing from 0.041 to 0.006 (see Table 1.11, last row.). The monotonicity of the coefficients adds credibility to the interpretation of the results.

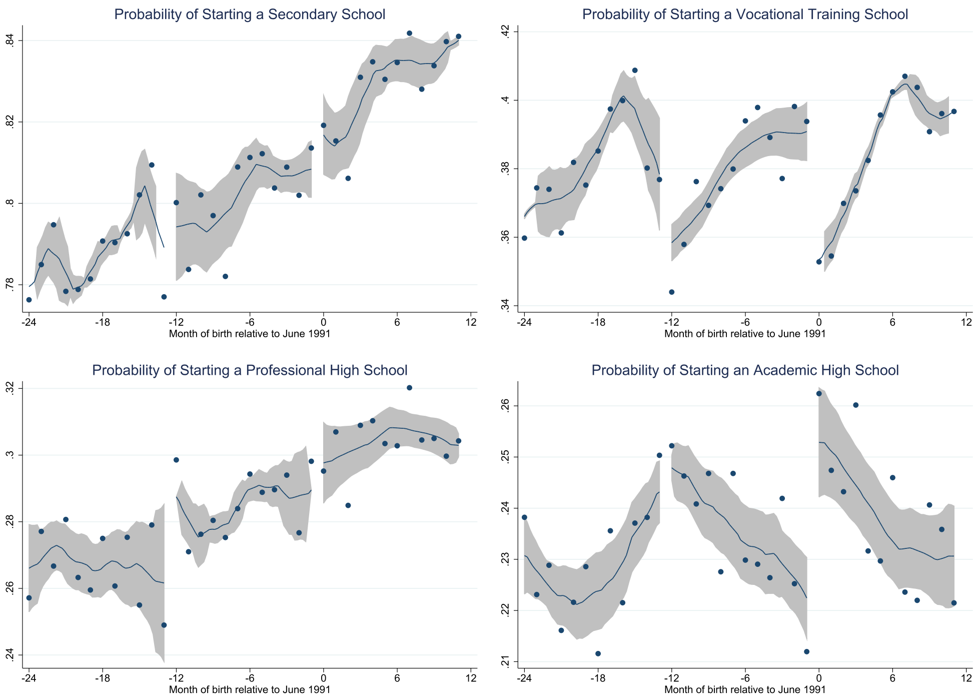
The positive effect on the probability of starting academic high schools instead of professional high schools or vocational training schools, although not robust across all specifications, is weakly present in the total sample (Figure 1.10). In fact, if this effect on the total sample was the only hypothesis to test, its coefficient (0.015) would be significant on 5% with a p-value of 0.046 (See Table 1.11, first item in the last row.). After the multiple testing procedure though, it is not considered as significant on 5% any more.

Table 1.11: Effects on School Choice

Tracks	Total sample	Mother's highest education at giving birth: Primary	Vocational	High school	Tertiary	Missing
Effect on the probability of finishing at least the first year in a secondary school						
Any secondary school	-0.000	0.003	-0.001	-0.004	-0.002	0.000
	(0.003)	(0.006)	(0.003)	(0.002)	(0.001)	(0.007)
	[102,616]	[26,242]	[19,474]	[25,567]	[9,061]	[16,928]
	{152.6}	{170.0}	{160.1}	{153.8}	{152.8}	{100}
Corrected p-values	0.913	0.743	0.901	0.238	0.232	0.970
Vocational training school	-0.017*	-0.049***	-0.045	-0.003	0.005	0.005
	(0.006)	(0.005)	(0.021)	(0.003)	(0.003)	(0.009)
	[80,505]	[19,323]	[14,234]	[29,280]	[7,829]	[16,928]
	{118.2}	{123.9}	{116.1}	{177.2}	{130.2}	{100}
Corrected p-values	0.084	0.000	0.144	0.376	0.241	0.742
Professional high school	-0.002	0.004	-0.010	-0.016	-0.033*	0.023**
	(0.007)	(0.010)	(0.014)	(0.010)	(0.012)	(0.006)
	[103.292]	[21,735]	[16,804]	[23,159]	[8,057]	[16,928]
	{153.0}	{139.6}	{138.1}	{138.2}	{134.9}	{100}
Corrected p-values	0.809	0.805	0.636	0.242	0.068	0.027
Academic high school	0.015	0.041***	0.032**	0.028**	0.006	-0.023
	(0.006)	(0.008)	(0.008)	(0.008)	(0.014)	(0.010)
	[86,292]	[24,082]	[21,467]	[25,220]	[9,247]	[16,928]
	{127.6}	{140.9}	{177.7}	{151.8}	{155.8}	{100}
Corrected p-values	0.099	0.003	0.010	0.021	0.814	0.132

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Robust standard errors clustered by birth year and month are in parentheses, observation numbers are in brackets, bandwidths in days are in braces. P-values are corrected by the number of hypothesis tests (78) done together in Tables 1.10 - 1.13 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after multiple testing correction. The same results estimated by a 100-day bandwidth are in Table A.5 in Appendix A.

Figure 1.10: Effects on School Choice, Children of Mothers with a Primary Degree



Local linear regressions fit separately below and above the cutoff in 1990 and in 1991 on birth year and month averages, using the rule-of-thumb bandwidth of lpoly in Stata with 90% confidence intervals. Month of birth -12 represents June 1990, 0 represents June 1991. No. of individual observations: 76,984.

1.6.6 Effects on Dropout Rates

The increase in CSL age causes an increase in the probability of dropping out of vocational training schools (Table 1.12). The average effect is an increase by 2.2 percentage points. Comparing this to the share of dropouts from vocational training schools in the previous cohort, which is 11.5% (see Table A.4), this is a 2.2/11.5=19% effect.

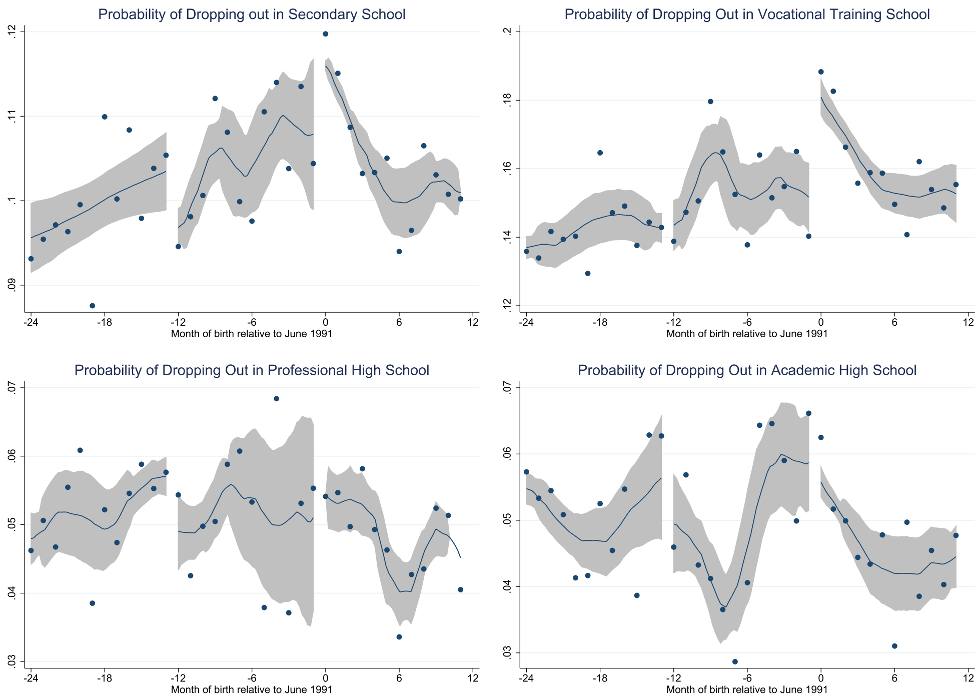
The ITT effect on dropping out of vocational training schools is not significant on 5% in the case of mothers with the lowest educational outcomes due to the multiple testing procedure (p-value=0.010).

Table 1.12: Effects on Dropout Rates

		Mother's highest education at giving birth				
Tracks	Total sample	Primary	Vocational	High school	Tertiary	Missing
	Effect on the probability of dropping out of ...					
Any secondary school	0.008**	0.011	0.010*	0.005	-0.002	0.006
	(0.003)	(0.009)	(0.005)	(0.003)	(0.003)	(0.006)
	[82,615]	[19,234]	[19,580]	[17,678]	[9,682]	[20,111]
	{137.5}	{160.5}	{178.5}	{117.3}	{164.1}	{156.2}
Corrected p-values	0.017	0.319	0.067	0.246	0.562	0.473
Vocational training school	0.022**	0.037***	0.017*	0.026	-0.028	0.016
	(0.004)	(0.012)	(0.007)	(0.010)	(0.044)	(0.013)
	[25,525]	[9,733]	[5,314]	[2,106]	[337]	[3,238]
	{174.1}	{165.2}	{161.4}	{112.0}	{167.3}	{100.0}
Corrected p-values	0.001	0.003	0.082	0.091	0.671	0.373
Professional high school	0.008	0.004	0.010	0.006	0.004	0.003
	(0.004)	(0.005)	(0.005)	(0.004)	(0.003)	(0.010)
	[32,570]	[8,122]	[9,913]	[9,600]	[1,534]	[4,985]
	{138.7}	{178.0}	{182.9}	{136.2}	{112.3}	{100.0}
Corrected p-values	0.134	0.617	0.173	0.245	0.335	0.797
Academic high school	0.003	-0.001	0.010***	0.000	-0.002	0.006
	(0.002)	(0.008)	(0.003)	(0.003)	(0.002)	(0.004)
	[48,077]	[6,049]	[6,423]	[13,067]	[9,472]	[8,289]
	{156.9}	{164.6}	{139.2}	{138.8}	{203.9}	{100.0}
Corrected p-values	0.249	0.970	0.001	0.965	0.425	0.351

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Robust standard errors clustered by birth year and month are in parentheses. Observation numbers are in brackets. Bandwidths in days are in braces. P-values are corrected by the number of hypothesis tests (78) done together in Tables 1.10 - 1.13 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after multiple testing correction. The same results estimated by 100-day bandwidths are in Table A.6 in Appendix A.

Figure 1.11: Effects on Dropout Rates, Children of Mothers with a Primary Degree



Local linear regressions fit separately below and above the cutoff in 1990 and in 1991 on birth year and month averages, using the rule-of-thumb bandwidth of lpoly in Stata with 90% confidence intervals. Month of birth -12 represents June 1990, 0 represents June 1991. No. of individual observations: 61,279, 29,254, 21,170 and 18,476, respectively.

1.6.7 Effects on School Completion

The negative effect of the CSL age increase on the higher probability of not earning a secondary degree, and in particular, a vocational training school degree, is the highest in the two lowest maternal education groups (see Table 1.13), mostly because they are the most likely to attend vocational schools (see Table A.4 in Appendix A.). In the same two groups, the higher CSL age has a quite large positive effect on the probability of gaining an academic high school degree, at 4.4 and 3.5 percentage points, respectively (see Table 1.13). However, it is worrying that this same effect is at a significant -5 percentage points for the group of children with missing maternal education data (see Table 1.13, last column). The group of children with missing maternal education data are the observations that could not be linked to the Vital Statistics records which contain information on maternal education. The overall positive effect of the legislation change on the probability of gaining an academic high school degree is 1.7 percentage points and highly significant. Thus, the findings of this paper holds without estimating any heterogeneous effects by parental education as well.

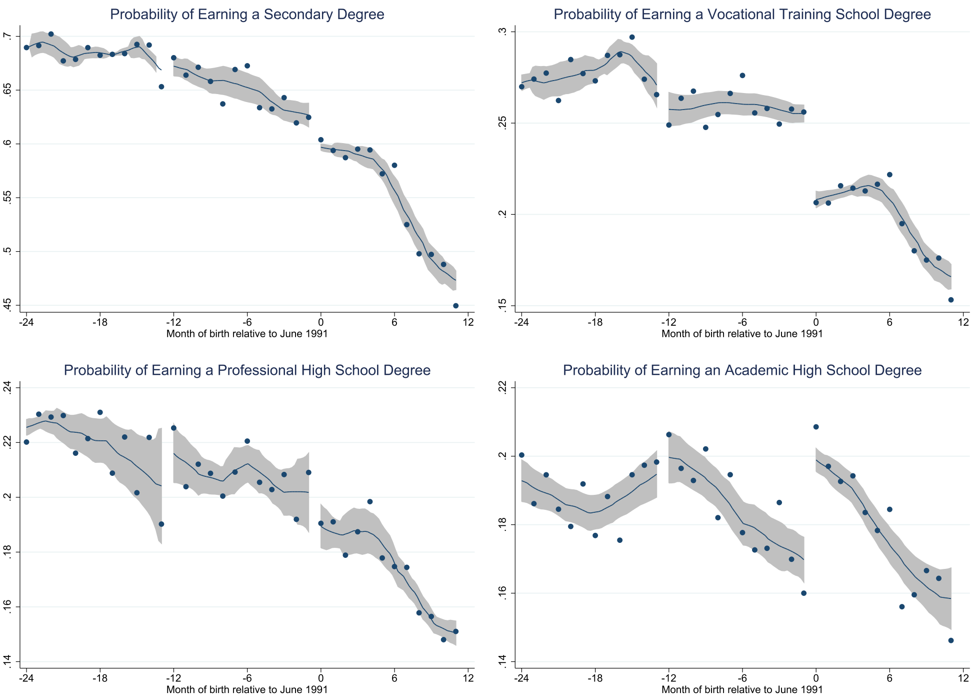
Table 1.13: Effects on School Completion

		Mother's highest education at giving birth				
	Total sample	Primary	Vocational	High school	Tertiary	Missing
Any secondary degree	-0.015***	-0.022**	-0.020**	-0.011***	-0.004	-0.017**
	(0.002)	(0.006)	(0.006)	(0.002)	(0.009)	(0.005)
	[62,682]	[20,359]	[12,615]	[14,319]	[4,819]	[16,928]
	{92.4}	{130.2}	{103.9}	{84.1}	{80.1}	{100}
Corrected p-values	0.002	0.020	0.040	0.005	0.762	0.029
Vocational training school degree	-0.019**	-0.054**	-0.045**	-0.011***	0.003	0.002
	(0.004)	(0.002)	(0.012)	(0.002)	(0.002)	(0.005)
	[70,804]	[20,359]	[16,343]	[21,901]	[8,109]	[16,928]
	{104.4}	{114.4}	{134.8}	{130.1}	{135.6}	{100}
Corrected p-values	0.009	0.000	0.014	0.003	0.348	0.800
Professional high school degree	-0.013	-0.010	-0.020*	-0.024*	-0.020	0.026*
	(0.006)	(0.008)	(0.007)	(0.008)	(0.012)	(0.004)
	[83,827]	[28,655]	[18,040]	[19,796]	[9,184]	[7,852]
	{123.5}	{185.7}	{148.8}	{117.3}	{154.5}	{46.9}
Corrected p-values	0.125	0.332	0.056	0.051	0.228	0.024
Academic high school degree	0.017***	0.044***	0.035***	0.028***	0.019	-0.050**
	(0.002)	(0.005)	(0.006)	(0.005)	(0.015)	(0.007)
	[76,374]	[23,200]	[16,695]	[19,111]	[7,526]	[9,693]
	{112.7}	{149.9}	{137.3}	{113.5}	{125.5}	{57.6}
Corrected p-values	0.001	0.000	0.002	0.005	0.350	0.020

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Robust standard errors clustered by birth year and month are in parentheses. Observation numbers in the local regression are in brackets. Bandwidths in days are in braces. P-values are corrected by the number of hypothesis tests (78) done together in Tables 1.10 - 1.13 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after multiple testing correction. The same results estimated by 100-day bandwidths are in Table A.7 in Appendix A.

10.14754/CEU.2016.11

Figure 1.12: Effects on School Completion, Children of Mothers with a Primary Degree



Local linear regressions fit separately below and above the cutoff in 1990 and in 1991 on birth year and month averages, using the rule-of-thumb bandwidth of lpoly in Stata with 90% confidence intervals. Month of birth -12 represents June 1990, 0 represents June 1991. No. of individual observations: 76,984.

1.6.8 Local Average Treatment Effects (LATE) of the Legislation Change

Similarly to the general results as presented in Section 1.5.5, LATE are about three times as large as the ITT impacts (see Table 1.14). The LATE of increased CSL age on the probability of starting an academic high school is the highest in the lowest maternal education group, at 14 percentage points. Comparing this to the share of children starting at an academic high school before the legislation change, 24% (see Table A.4 in Appendix A.2), this is a 58% LATE on the compliers. The LATE on the probability of earning a secondary degree is of the same magnitude in the group of children of the least educated mothers, 14.9 percentage points, or 0.149/0.192=78%.

Table 1.14: LATE on Starting an Academic High School and Earning an Academic High School Degree

	Total sample	Mother's highest education at giving birth: Primary	Vocational	High school	Tertiary	Missing
Effect on the probability of starting an academic high school						
ITT effect	0.015**	0.041***	0.032***	0.028***	0.006	-0.056***
	(0.006)	(0.008)	(0.008)	(0.008)	(0.014)	(0.006)
First stage (jump in the probability of starting school under CSL age 18)	0.321***	0.029**	0.361***	0.380***	0.310***	N/A
	(0.025)	(0.011)	(0.030)	(0.041)	(0.031)	
	[2,110.7]	[3,019.8]	[711.1]	[922.1]	[1263.7]	
Wald estimate of the LATE coefficient	0.046	0.140	0.089	0.074	0.018	-
95% bootstrapped confidence intervals of the LATE coefficient	0.004 - 0.086	0.069 - 0.215	0.017 - 0.159	0.010 - 0.141	-0.082 - 0.121	-
No. of obs.	86,292	24,082	21,467	25,220	9,247	6,990
Bandwidth (in days)	127.6	155.1	177.1	151.8	155.8	41.5
Effect on the probability of earning an academic high school degree						
ITT effect	0.017***	0.044***	0.035***	0.028***	0.019**	-0.050***
	(0.002)	(0.005)	(0.006)	(0.005)	(0.007)	(0.007)
First stage (jump in the probability of starting school under CSL age 18)	0.311***	0.297***	0.354***	0.346***	0.290***	N/A
	(0.022)	(0.010)	(0.034)	(0.034)	(0.028)	
	[2,515.1]	[2,2828.4]	[1,210.4]	[1,489.4]	[1,602.5]	
Wald estimate of the LATE coefficient	0.055	0.149	0.098	0.082	0.066	-
95% bootstrapped confidence intervals of the LATE coefficient	0.008 - 0.100	0.080 - 0.215	0.015 - 0.182	0.001 - 0.167	-0.065 - 0.206	-
No. of obs.	76,374	23,200	16,695	19,111	7,526	9,693
Bandwidth (in days)	112.7	149.9	137.3	113.5	125.5	57.6

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Wald estimates are calculated as the ratios of ITT coefficients over the first stage coefficient (the size of the jump). Robust standard errors clustered by birth year and month are in parentheses, first stage F-statistics are in brackets. A */**/*** indicates significance on 10%/5%/1% level. The bootstrapped confidence intervals of the LATE coefficients are constructed following Bjorklund and Jantti (1997). First, a birth year and month stratified bootstrap sample is drawn from the 2001 Hungarian Census data, and birth year and month school entering probabilities ($\hat{p98}_m$) are estimated. Then, a birth year and month stratified bootstrap sample is drawn from the 2011 Hungarian Census data, from which the LATE parameters ($\hat{\beta}_{LATE,1}$) are estimated using the school entering rates estimated in the first step. Repeating these two steps $B = 1000$ times yields an empirical distribution of $1,000$ estimated LATE coefficients ($\hat{\beta}_{LATE,1} - \hat{\beta}_{LATE,1000}$). The 95% confidence interval of $\hat{\beta}_{LATE}$ is set as the middle 95% of this empirical distribution.

Table 1.15: LATE on Dropping Out of Vocational Training School

	Total sample	Mother's highest education at giving birth: Primary	Vocational	High school	Tertiary	Missing
Effect on the probability of dropping out of vocational training school						
ITT effect	0.022***	0.037**	0.017**	0.026**	-0.028	-0.014
	(0.004)	(0.012)	(0.007)	(0.010)	(0.044)	(0.011)
First stage (jump in the probability of starting school under CSL age 18)	0.333***	0.284***	0.362***	0.344***	0.313***	N/A
	(0.025)	(0.016)	(0.033)	(0.033)	(0.030)	
	[792.8]	[3,257.6]	[672.6]	[1,707.0]	[1,524.3]	
Wald estimate of the LATE coefficient	0.065	0.130	0.047	0.076	-0.089	-
95% bootstrapped confidence intervals of the LATE coefficient	0.015 - 0.114	0.023 - 0.240	-0.029 - 0.121	-0.031 - 0.189	-0.365 - 0.179	-
No. of obs.	25,525	9,733	5,317	2,106	337	2,183
Bandwidth (in days)	174.1	165.2	161.4	112.0	167.3	67.1

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Wald estimates are calculated as the ratios of ITT coefficients over the first stage coefficient (the size of the jump). Robust standard errors clustered by birth year and month are in parentheses, first stage F-statistics are in brackets. A */**/*** indicates significance on 10%/5%/1% level. The bootstrapped confidence intervals of the LATE coefficients are constructed following Bjorklund and Jantti (1997). First, a birth year and month stratified bootstrap sample is drawn from the 2001 Hungarian Census data, and birth year and month school entering probabilities ($\hat{p98}_m$) are estimated. Then, a birth year and month stratified bootstrap sample is drawn from the 2011 Hungarian Census data, from which the LATE parameters ($\hat{\beta}_{LATE,1}$) are estimated using the school entering rates estimated in the first step. Repeating these two steps $B = 1000$ times yields an empirical distribution of $1,000$ estimated LATE coefficients ($\hat{\beta}_{LATE,1} - \hat{\beta}_{LATE,1000}$). The 95% confidence interval of $\hat{\beta}_{LATE}$ is set as the middle 95% of this empirical distribution.

Increasing the CSL age raised the probability of dropping out of vocational training school by 13 percentage points among children of mothers with primary education around the cutoff (see Table 1.15). In per cent terms, this is an 0.130/0.149=87% impact on the compliers.

1.7 Robustness Checks

This section supports the main findings of Section 5 and 6 using several robustness checks.

1.7.1 Parental Education Data from Other Sources Give Similar Results

Until this point, all results based on parental education in Section 7 were estimated using a subsample of individuals whose birth records data, containing information on maternal education, could be linked to the 2011 Hungarian Census using linking method 1. As detailed in Section 1.6.1, method 1 is based on a two-level information comparison approach. In the case of those still living with their parents, the linking procedure uses all available information appearing in both datasets, including the date of birth of parents. In the case of those not living with their parents anymore, it uses maternal residence and exact date of birth only. Thus, those living at home in 2011 are easier to link using this method. As it is also detailed in Section 1.6.1, there are two additional methods for gaining information on parental education: linking birth records without using any information on parents (linking method 2), and using the subsample of those still living at home at the time of the 2011 Hungarian Census. It would be worrying if the three methods would lead to different results. However, this is not the case. Table 1.16 shows that all three methods lead to similar results.

Table 1.16: Robustness Check 1 - Parental Education from Different Sources (ITT effects)

	Total sample	Highest parental education: Primary	Vocational	High school	Tertiary	Missing
Effect on the probability of starting an academic high school						
Maternal education at giving birth - linking with method 1	0.021***	0.050***	0.038***	0.036***	0.022	-0.023**
	(0.004)	(0.009)	(0.008)	(0.006)	(0.013)	(0.010)
(the same as in Table A.5 in Appendix A.3)	[67,259]	[15,588]	[12,097]	[16,696]	[5,950]	[16,928]
Maternal education at giving birth - linking with method 2		0.058***	0.010	0.046**	0.035	0.011**
		(0.008)	(0.010)	(0.015)	(0.022)	(0.005)
		[9,432]	[6,396]	[6,743]	[1,529]	[43,159]
Educational status of the household head		0.062***	0.041***	-0.001	0.013	0.011
		(0.005)	(0.007)	(0.009)	(0.008)	(0.007)
		[8,121]	[20,011]	[12,446]	[7,820]	[18,861]
Effect on the probability of earning an academic high school degree						
Maternal education at giving birth - linking with method 1	0.018***	0.055***	0.028**	0.030***	0.025	-0.030
	(0.003)	(0.005)	(0.010)	(0.007)	(0.018)	(0.010)
(the same as in Table A.7 in Appendix A.3)	[67,259]	[15,588]	[12,097]	[16,696]	[5,950]	[16,928]
Maternal education at giving birth - linking with method 2		0.065***	0.005	0.044**	0.048	0.004
		(0.009)	(0.012)	(0.016)	(0.027)	(0.004)
		[9,432]	[6,396]	[6,743]	[1,529]	[43,159]
Educational status of the household head		0.053***	0.032***	0.001	0.007	0.011
		(0.008)	(0.005)	(0.011)	(0.006)	(0.006)
		[8,121]	[20,011	[12,446]	[7,820]	[18,861]
Effect on the probability of dropping out of vocational training school						
Maternal education at giving birth - linking with method 1 (the same as in Table A.6 in Appendix A.3)	0.029**	0.052**	-0.002	0.028*	0.061***	0.016
	(0.005)	(0.012)	(0.007)	(0.012)	(0.012)	(0.013)
	[14,319]	[5,841]	[3,210]	[1,841]	[189]	[3,238]
Maternal education at giving birth - linking with method 2		0.035*	-0.023	0.012	-0.043	0.041***
		(0.012)	(0.013)	(0.021)	(0.070)	(0.007)
		[3,785]	[1,894]	[921]	[61]	[7,658]
Educational status of the household head		0.036*	0.033**	0.009	0.064**	0.014
		(0.015)	(0.006)	(0.009)	(0.007)	(0.012)
		[3,313]	[5,897]	[1,640]	[344]	[3,125]

Local linear kernel regressions using a 100-day bandwidth. Robust standard errors clustered by birth year and month are in parentheses, number of observations are in brackets. A */**/*** indicates significance on 10%/5%/1% level.

1.7.2 No Significant Effects at Cutoffs in the Before and After Years

One might be right to worry about whether the effects captured by this analysis are due to either starting elementary school at different ages (just over the age of 6 vs. almost age 7), or are just due to time (or academic-year) trends. If either of these were true, we should see the same effects for students born around the June 1 cutoff in other years. Two things happen around cutoffs in other years: (1) those born in June start school at an older age, and (2) those born in June spend one year less in school before reaching the CSL age. Theoretically, starting school at an older age can affect education outcomes positively or negatively. Black, Devereaux and Salvanes (2011) argue that starting school at an older age may be beneficial for learning because older children are at a more advanced stage of their developmental life. In addition, social development may depend on a child's age relative to the class. If being older than one's peers is beneficial, starting school at an older age would be beneficial as well. However, it is not clear whether this is really the case.

On the other hand, starting school at an older age may be harmful if children are able to learn more in school than in pre-school (or at home). Furthermore, parental investment in helping children with their school work may depend on school starting age as well – parents may provide less help to children if they start school when they are older. Black, Devereaux and Salvanes (2011) examine the effect of school starting age on education outcomes and they find very small positive effects of starting school when younger. In the Hungarian case, Hámori and Köllő (2011) examine the effect of school starting age on test results taken in Grades 4 and 8. They find a positive effect of starting school at an older age when in Grade 6 but the effect becomes much smaller by Grade 8. However, they cannot separate the effect of school starting age from the effect of age at time of the test, as these two are perfectly collinear.

Figures 1.5 - 1.7 in Section 5 plot all outcomes variables around the same cutoff in the previous year along with the cutoff of the legislation change in 1991. In the case of the probability of starting a vocational training school and earning a vocational school degree there is a break in the data around the cutoff in the previous year as well. Those born right after June 1 in both 1990 and 1991 are less likely to start a vocational training school and earn a vocational training school degree than those born right before that (see Figures 1.5 and 1.7). However, I do not find significant effects for the rest of the educational outcomes around the cutoff in 1990.

Table 1.17 shows the ITT effects of cutoffs in the year before and after 1991, alongside the real cutoff in 1991, for the three main outcome variables. On the probability of starting and completing academic high school, and, dropping out of vocational training school, the 1991 cutoff produces the largest, and the solely significant, coefficients. Thus, the effects of starting school when older and spending fewer years in school balance each other out in the comparison cohorts of those born in 1990 and 1992.

Table 1.17: Robustness Check 2 - Effects of Cutoffs in Other Years (ITT effects)

		Mother's highest education at giving birth				
Cutoff: June 1	Total sample	Primary	Vocational	High school	Tertiary	Missing
Effect on the probability of choosing an academic high school						
1990	-0.001	-0.004	-0.003	0.024	-0.015	-0.015
	(0.006)	(0.008)	(0.004)	(0.014)	(0.012)	(0.010)
	[65,408]	[14,969]	[11,490]	[15,872]	[5,824]	[17,253]
1991	0.021***	0.050***	0.038***	0.036***	0.022	-0.023**
	(0.004)	(0.009)	(0.008)	(0.006)	(0.013)	(0.010)
	[67,259]	[15,588]	[12,097]	[16,696]	[5,950]	[16,928]
1992	0.008	0.001	0.024*	-0.013	0.005	0.026
	(0.007)	(0.004)	(0.012)	(0.010)	(0.013)	(0.020)
	[64,903]	[15,250]	[12,091]	[16,223]	[5,971]	[15,3768]
Effect on the probability of earning an academic high school degree						
1990	0.003	-0.003	0.013	0.027	-0.015	-0.003
	(0.006)	(0.009)	(0.007)	(0.017)	(0.013)	(0.009)
	[65,408]	[14,969]	[11,490]	[15,872]	[5,824]	[17,253]
1991	0.018***	0.055***	0.028**	0.030***	0.025	-0.030
	(0.003)	(0.005)	(0.010)	(0.007)	(0.018)	(0.010)
	[67,259]	[15,588]	[12,097]	[16,696]	[5,950]	[16,928]
1992	-0.019	-0.013	-0.019	-0.039**	-0.029**	-0.010
	(0.010)	(0.008)	(0.010)	(0.014)	(0.011)	(0.013)
	[64,903]	[15,250]	[12,091]	[16,223]	[5,971]	[15,368]
Effect on the probability of dropping out of vocational training school						
1990	-0.005	-0.015	0.015	0.014	0.060	-0.022
	(0.006)	(0.010)	(0.017)	(0.016)	(0.035)	(0.019)
	[14,452]	[5,604]	[3,223]	[1,918]	[203]	[3,504]
1991	0.029**	0.052**	-0.002	0.028*	0.061***	0.016
	(0.005)	(0.012)	(0.007)	(0.012)	(0.012)	(0.013)
	[14,319]	[5,841]	[3,210]	[1,841]	[189]	[3,238]
1992	-0.002	-0.001	0.002	0.012	-0.008	-0.016
	(0.007)	(0.005)	(0.016)	(0.014)	(0.049)	(0.011)
	[13,882]	[5,801]	[3,053]	[1,762]	[190]	[3,076]

Local linear kernel regressions using a 100-day bandwidth. Robust standard errors clustered by birth year and month are in parentheses, number of observations are in brackets. A */**/*** indicates significance on 10%/5%/1% level.

1.7.3 Alternative Optimal Bandwidth Choices Give Similar Results

Another concern may be whether these results are sensitive to bandwidth choice. Tables 1.18 and 1.19 suggest that using the 50-150% versions of the CCT bandwidth gives very similar results, both in the terms of the magnitude of the coefficients and their p-values.

Table 1.18: Robustness Check 3/A - Effects Using 50-150% Versions of the CCT bandwidth (ITT effects)

Version of CCT (2014) optimal bandwidth	Total sample	Mother's highest education at giving birth: Primary	Vocational	High school	Tertiary	Missing
Effect on the probability of starting an academic high school						
50%	0.018***	0.055***	0.034**	0.034***	0.030**	-0.02
	(0.002)	(0.006)	(0.009)	(0.004)	(0.010)	(0.011)
75%	0.023***	0.045***	0.031***	0.034***	0.021*	-0.019
	(0.004)	(0.009)	(0.008)	(0.005)	(0.010)	(0.022)
100%	0.015**	0.041***	0.032***	0.028***	0.006	-0.023
	(0.006)	(0.008)	(0.008)	(0.008)	(0.014)	(0.010)
	[86,292]	[24,082]	[21,467]	[25,220]	[9,247]	[16,928]
	{127.6}	{140.9}	{177.7}	{151.8}	{155.8}	{100}
125%	0.018***	0.043***	0.022**	0.017	0.004	-0.016
	(0.005)	(0.008)	(0.009)	(0.010)	(0.013)	(0.015)
150%	0.016***	0.036***	0.030***	0.023***	0.015	-0.010
	(0.005)	(0.008)	(0.007)	(0.008)	(0.012)	(0.007)
Effect on the probability of earning an academic high school degree						
50%	0.016**	0.051***	0.042***	0.035***	0.042**	-0.048**
	(0.003)	(0.003)	(0.010)	(0.004)	(0.010)	(0.013)
75%	0.021***	0.052***	0.029**	0.030***	0.031	-0.046***
	(0.004)	(0.005)	(0.009)	(0.004)	(0.016)	(0.008)
100%	0.017***	0.044***	0.035***	0.028***	0.019	-0.050**
	(0.002)	(0.005)	(0.006)	(0.005)	(0.015)	(0.007)
	[76,374]	[23,200]	[16,695]	[19,111]	[7,526]	[9,693]
	{112.7}	{149.9}	{137.3}	{113.5}	{125.5}	{57.6}
125%	0.016**	0.045***	0.032***	0.019**	0.010	-0.033**
	(0.004)	(0.006)	(0.006)	(0.009)	(0.017)	(0.010)
150%	0.015***	0.043***	0.025***	0.017	0.004	-0.015
	(0.003)	(0.006)	(0.006)	(0.010)	(0.018)	(0.016)

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Robust standard errors clustered by birth year and month are in parentheses. Number of observations are in brackets, 100% CCT bandwidths in days in braces. A */**/*** indicates significance on 10%/5%/1% level.

Table 1.19: Robustness Check 3/B - Effects Using 50-150% Versions of the CCT bandwidth (ITT effects)

Version of CCT (2014) optimal bandwidth	Total sample	Mother's highest education at giving birth				
		Primary	Vocational	High school	Tertiary	Missing
Effect on the probability of dropping out of vocational training schools						
50%	0.029***	0.059***	0.002	0.012	0.051**	0.008
	(0.004)	(0.010)	(0.006)	(0.010)	(0.013)	(0.004)
75%	0.027***	0.050***	0.010	0.016	-0.020	0.010
	(0.005)	(0.011)	(0.007)	(0.007)	(0.043)	(0.004)
100%	0.022***	0.037**	0.017**	0.026**	-0.028	0.016
	(0.004)	(0.012)	(0.007)	(0.010)	(0.044)	(0.013)
	[25,525]	[9,733]	[5,314]	[2,106]	[337]	[3,238]
	{174.1}	{165.2}	{161.4}	{112.0}	{167.3}	{100.0}
125%	0.022***	0.033***	0.020**	0.019	-0.035	0.006
	(0.004)	(0.010)	(0.007)	(0.012)	(0.046)	(0.003)
150%	0.023***	0.036***	0.015**	0.027*	-0.039	0.006
	(0.004)	(0.009)	(0.006)	(0.013)	(0.037)	(0.003)

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014). Robust standard errors clustered by birth year and month are in parentheses. Number of observations are in brackets, 100% CCT bandwidths in days in braces. A */**/*** indicates significance on 10%/5%/1% level.

1.7.4 Global Polynomial Estimation Gives Similar Results

Even though cutoffs in others years do not yield statistically significant impacts, their coefficients are not significant zeros. As a result, one may still be worried about seasonality in the data or may question the assumption that date of birth is exogenous in general (see Section 1.4.1). To control for such seasonality effects, global polynomial models are estimated so that in addition to the 4th-order polynomial functions of the assignment variable below and above the cutoff, birth year and month fixed effects (FE) are explicitly controlled. Day of the week of birth, county, and settlement type FE's are also explicitly controlled. Table 1.20 shows that the estimated coefficients are very similar in their signs, magnitude, and significance to the earlier results.

Table 1.20: Robustness Check 4 - Global Polynomial Estimation (ITT effects)

		Mother's highest education at giving birth				
	Total sample	Primary	Vocational	High school	Tertiary	Missing
Effect on the probability of starting an academic high school						
Local linear effect	0.015**	0.041***	0.032***	0.028***	0.006	-0.023
	(0.006)	(0.008)	(0.008)	(0.008)	(0.014)	(0.010)
	[86,292]	[24,082]	[21,467]	[25,220]	[9,247]	[16,928]
	{127.6}	{140.9}	{177.7}	{151.8}	{155.8}	{100}
Global polynomial effect	0.011***	0.034***	0.022	0.018**	0.008	-0.023
	(0.004)	(0.009)	(0.010)	(0.008)	(0.013)	(0.012)
	586,283	136,513	103,329	143,259	51,010	[152,172]
Effect on the probability of earning an academic high school degree						
Local linear effect	0.017***	0.044***	0.035***	0.028***	0.019	-0.050**
	(0.002)	(0.005)	(0.006)	(0.005)	(0.015)	(0.007)
	[76,374]	[23,200]	[16,695]	[19,111]	[7,526]	[9,693]
	{112.7}	{149.9}	{137.3}	{113.5}	{125.5}	{57.6}
Global polynomial effect	0.013**	0.043***	0.024**	0.015	0.014	-0.025*
	(0.006)	(0.006)	(0.011)	(0.011)	(0.018)	(0.014)
	[586,283]	[136,513]	[103,329]	[143,259]	[51,010]	[152,172]
Effect on the probability of dropping out of vocational training schools						
Local linear effect	0.022***	0.037**	0.017**	0.026**	-0.028	0.016
	(0.004)	(0.012)	(0.007)	(0.010)	(0.044)	(0.013)
	[25,525]	[9,733]	[5,314]	[2,106]	[337]	[3,238]
	{174.1}	{165.2}	{161.4}	{112.0}	{167.3}	{100.0}
Global polynomial effect	0.028***	0.063***	0.004	0.004	-0.089	0.012
	(0.007)	(0.009)	(0.010)	(0.014)	(0.064)	(0.008)
	[132,209]	[52,303]	[28,937]	[17,194]	[1,797]	[31,978]

Local linear kernel regressions using the bandwidth optimization routine of CCT (2014) and global polynomial estimations using data from 1988-1992, controlling for a 4th order polynomial function of the running variable separately below and above the cutoff, and the following FE's: year of birth, month of birth, day of the week of birth, county, settlement type. Robust standard errors clustered by birth year and month are in parentheses. Number of observations are in brackets, CCT 2014 bandwidths in days are in braces. A */**/*** indicates significance on 10%/5%/1% level.

The fact that the results of my nonparametric and parametric empirical strategy coincide is important. Results in the CSL age literature tend to be sensitive to using an econometric method. For example, several papers use geographical and timing variations of CSL age legislation changes to assume a common time trend in these geographical units. They control separately for state and cohort fixed effects (Acemoglu and Angrist, 2001; Lochner and Moretti, 2004; Lleras-Muney, 2005; Oreopoulos, 2006). However, Stephens and Yang (2014) show that once state-specific time trends are introduced into the original models using the same US Census data that these papers used before, the estimated positive effects of longer education on wages, occupation, unemployment, and divorce will lose significance, and that many would even turn sign. Devereux and Hart (2010) reevaluate a CSL age increase in 1947 in England, evaluated by Harmon and Walker (1995) and Oreopoulos (2007). They use an RDD setup and get 3 to 5 times lower wage returns than earlier papers using other methods.

This global polynomial method is similar to the diff-in-diffs strategy applied by a wide strand of the CSL age literature (e.g. Meghir and Palme, 2005; Oreopoulos, 2007; Oosterbeek and Webbink, 2007; and Pischke and Wachter, 2008). It is relieving that in this particular case, the results of the diff-in-diffs and the RDD techniques coincide.

1.8 Supportive Evidence from Other Data Sources and Potential Channels

1.8.1 The Share of those Attending Academic High Schools in the NABC Data

One of the main findings of this paper is that the probability of starting at an academic high school is positively affected by the CSL age increase. Table 1.21 shows the share of students attending the three types of secondary school in the Grade 10 waves of the National Assessment of Basic Competencies (NABC) data. This may be a biased sample in the sense that the probability of making it to Grade 10 and being observed by the survey may be related to the CSL age increase by design. Nonetheless, not finding the same result in this data would be worrying. As Table 1.21 shows, the ratio of those attending academic high school in Grade 10 is significantly higher among those born within three months after the cutoff. The difference is that of about 4 percentage points, a little higher but quite similar in magnitude to the effects estimated in the Census data.

Table 1.21: The Probability of Attending Secondary Schools by Type, NABC data

	Born in March-May 1991	Born in June-Aug 1991	Difference	t-test p-value
Vocational training sch.	0.484	0.415	-0.069***	0.000
Professional high school	0.387	0.414	0.027	0.048
Academic high school	0.129	0.171	0.042***	0.000

Own estimation from the Grade 10 waves of the 2006-2010 NABC data. No. of observations: 5,669. A */**/*** indicates significance on 10%/5%/1% level.

1.8.2 Potential Explanation to the Increase in Dropouts: Change in Student Composition

One potential channel through which a higher CSL age may increase dropout rates could be its effect on student composition. This works in two ways. Firstly, if the change induces more students to start secondary school at the lower end of the distribution, lower ability students would mainly attend vocational training schools, which may mechanically lead to higher dropout rates. This hypothesis is not likely in our case. As is seen in Table A.5 and Table 1.11, higher CSL age did not increase the probability of starting a secondary school in general, neither on average nor among children of less educated mothers. However, the probability of starting at a vocational training school decreased by 4.9 percentage points on the lower end of the distribution, while there is an increase of similar size in the probability of starting at an academic high school. Thus, it is quite likely that the students who switched from the vocational to the high school track were higher ability students, which could have an effect on student composition in vocational training schools. Secondly, keeping students in school for longer should have a composition effect with regard to those who would have otherwise dropped out.

Table 1.22: Heterogeneity of the Effect with Respect to the Share of the Less Educated (ITT effects)

Effect on the probability of dropping out of vocational training school			
	ITT effect	Robust SE	No. of obs.
1st quintile	0.028	0.025	1,736
2nd quintile	0.000	0.001	2,144
3rd quintile	0.005	0.005	2,675
4th quintile	0.048**	0.017	3,525
5th quintile	0.043**	0.020	4,239

Outcome variable: probability of dropping out of vocational training schools. Local linear kernel regressions using a 100-day bandwidth. The first column indicates which quintile the settlement where the individual was born belongs to with respect to its share of less educated adults. All rows represent a separate regression. Robust standard errors are clustered by birth year and month. A */**/*** indicates significance on 10%/5%/1% level.

There are signs that the higher CSL age changed student composition in vocational training schools. Assuming that there is a composition effect, dropout rates will be higher in settlements where the share of students from lower socioeconomic background is larger. Table 1.22 shows the effect of the legislation change by quintiles with respect to the share of less educated adults (hence, signaling the share of children of less educated mothers) in the settlement where the individual was born. The effect of dropout rates is indeed the highest and most significant in the top two quintiles where the share of less educated mothers is also the highest.

Table 1.23: Student Composition in Vocational Training Schools

	Avg. share of Roma students	Avg. share of students receiving child protection subsidy	Avg. share of students living in financially deprived family	Avg. share of students with at least one unemployed parent
	In vocational training schools			
2006	0.232	0.336	0.484	0.318
2007	0.178	0.267	0.410	0.285
2008	0.265	0.351	0.499	0.371
2009	0.354	0.407	0.462	0.408
2010	0.373	0.413	0.506	0.444
2011	0.408	0.436	0.522	0.435

Data source: own calculation from the school (*telephely*) level data of the Hungarian Assessment of Basic Competencies database. Schools offering primary education or academic high school track are excluded.

Also, the school (*telephely*) level data of the NABC data suggest that the share of disadvantaged students increased quite dramatically in vocational training schools after the reform (Table 1.23). In fact, according to school principals, the share of Roma students almost doubled between 2006 and 2011, and the share of students with at least one unemployed parent grew by about a quarter.

1.8.3 Potential Explanation to an Increase in Dropouts: Supply Constraints in Vocational Training Schools

Another reason behind the increase of vocational training school dropout rates is that no sufficient resources were allocated to schools to compensate for their extra workload in terms of the number and the composition of students. In line with some earlier evidence from a survey research (Mártonfi 2011a, 2011b), the available administrative data support this hypothesis as well. Table 1.24 shows the number of students and the average expenditures of schools during this period.

Table 1.24: Number of Students and Financing of Vocational Training Schools

		Vocational training programs only in school			Both vocational training and professional high school programs in school		
			Expenditures			Expenditures	
	No. of students in vocational education+	Avg. no. of students in school	In the previous fiscal year	Per student in current year	Avg. no. of students in school	In the previous fiscal year	Per student in current year
			index, 2006=100			index, 2006=100	
2006	126,211	246	100.0	100.0	704	100.0	100.0
2007	124,466	272	101.6	73.4	721	111.9	99.5
2008	129,066	284	82.5	99.4	726	114.1	103.4
2009	128,848	244	116.6	75.6	775	119.3	90.9
2010	135,268	248	76.2	60.5	783	112.1	93.8
2011	138,489	257	69.4	-	674	116.8	-

A + indicates data from the Statistical Yearbook of Education (Ministry of Human Resources, 2013). All other data are based on own calculation from the school (*telephely*) level data of the NABC database. Schools offering primary education or academic high school tracks are excluded. Expenditures are calculated from current prices, 2006=100. Students in all vocational school programs.

The structure of the available data and the school system is such that pinpointing exact numbers is not straightforward. Data on the number of students are available from two sources: the administrative data of the yearly Yearbooks of Education, and the school level data of the NABC database. However, as one school may offer programs in more than one track (i.e. offer both vocational training and professional high school tracks within the same institution), it is impossible to follow the expenditures of vocational schools only. Table 1.24 presents the expenditures of schools by fiscal (not academic) years, as well as their average expenditures per student. The data shows that schools did receive extra funding in 2008, the year when the first treated cohorts reached age 16 and were forced to continue their studies. In spite of this, on average, per capita expenditures of schools decreased between 2006 and 2011, even by current terms. However, there are differences between vocational schools as well: the decrease of expenditures in small schools offering vocational training programs only is much larger than in big schools offering both vocational and professional high school tracks.

Table 1.25: Teachers in Vocational Training Schools

	Number of students in vocational education+	Number of teachers in vocational education+	Student per teacher ratio in vocational education	Share of schools where at least one person teaches with no qualifications ł	Avg. number of teachers with no qualification ł in school
2006	126,211	8,938	14.1	0.565	2.32
2007	124,466	8,947	13.9	0.532	2.40
2008	129,066	8,942	14.4	0.520	2.35
2009	128,848	8,706	14.8	0.426	1.76
2010	135,268	8,824	15.3	0.605	4.15
2011	138,489	9,314	14.8	0.592	4.96

A + indicates data from the Statistical Yearbook of Education (Ministry of Human Resources, 2013); A ł indicates data from my own calculation using the school (*telephely*) level data of the NABC database. Schools offering exclusively vocational tracks are included. Students in all vocational programs.

The same is true with respect to student-per-teacher ratios. Table 1.25 shows the number and distribution of teachers working in vocational training schools. The number of students per teachers increased between 2006 and 2011, as well as the share and number of teachers teaching without proper qualifications.

1.9 Discussion

Increasing the CSL age from 16 to 18 in Hungary has a robust effect on the probability of choosing an academic high school at age 14, and earning an academic high school degree. This effect is realized at an intensive margin as it induced students to choose academic high schools over vocational schools but it did not affect their choice on whether to enter any secondary school or not. Most of the positive impacts on the probability of choosing and completing an academic high school was realized among children of less educated parents, who completed primary school at most. Considering the benefits that a high school degree has in the labor market, and also the fact that in Hungary 19% of children (468,000 children) at or below age 16 live in a family where the head of the household is less educated[15], this is a favorable outcome. Educational attainment is shown to be highly correlated across generations (Hertz et al., 2007) and it may be one of the strongest channels through which disadvantaged status and poverty is inherited (Haveman and Wolfe, 1995, and d'Addio, 2007). Closing the gap between the educational outcomes of the rich and the poor may bring substantial short- and long-term social returns.

This paper, however, documents an unexpected negative effect of the higher CSL age that caused an increase in the probability of students dropping out from vocational training schools. The fact that the new legislation put a strain on vocational training schools became obvious when the treated cohort reached age 16, which resulted in a growth in the number of students in these schools (Mártonfi, 2011a). The reason why vocational (and not high) schools were negatively affected derives from the highly selective and segregative

[15] Own calculation from the 2011 Hungarian Census.

nature of the Hungarian education system, which through free school choice and early tracking pushes children of low socioeconomic backgrounds into vocational schools.

There are two potential channels of this adverse effect on dropout rates. Firstly, although the demand for education services increased in vocational schools, they could not handle the increased workload due to supply side constraints. This finding is in line with the experience of development programs using demand-side interventions only. The literature on conditional cash transfers, such as cash benefits given to the poor on the condition of school attendance or participation in medical check-ups, concludes that one of the main elements of success is finding the right balance between demand and supply side components (Adato and Hoddinott, 2010). Secondly, composition of students shifted quite heavily towards children of very low socioeconomic background, causing a mechanical increase in dropout rates.

Chapter 2

The Effects of Increased Compulsory School Leaving Age on the Teenage Fertility of Roma Women, a Disadvantaged Ethnic Minority

(co-authors: Flóra Samu and Ágota Scharle)

2.1 Introduction

Teenage fertility is one of the most important sources of intergenerational poverty transmission (e.g. Bonell, 2004). The literature presents evidence on the negative health, social, and economic consequences of teenage childbearing, such as lower academic attainment and labor market attachment rates, higher rates of benefit reception, and infant mortality (Chevalier and Viitanen, 2003; Fletcher and Wolfe, 2008; Wilson 2012). Negative impacts have been found to carry on to the next generation as well: Navarro Paniagua and Walker (2012) show that having a teenage mother decreases the probability of post compulsory education, and daughters of adolescent mothers are more likely to become teenage mothers themselves. Due to high opportunity costs, the prevalence of teenage motherhood has been declining in most developed countries.

However, teenage motherhood is still very common in several communities of disadvantaged ethnic minorities living in developed countries. Examples include Mexican women in the US, women of Pakistani and Bangladeshi origin in the UK, Turkish women in Belgium and France, and Roma women throughout Europe. Bean and Swicegood (1985) argue that fertility differences between minority and majority women come from social and economic exclusion. The opportunity costs of early childbearing tend to be lower for minority women because their perceived and actual future economic opportunities are constrained, independently from their fertility patterns. Kearney and Levine (2012) propose that teenage motherhood provides an opportunity for women to exit the workforce. Young women choose motherhood instead of investing in the development of their own human capital because "they feel they have little chance of advancement". If this is true, policy tools that reduce high teenage fertility through human capital development among the majority ethnic group of a society may work differently in the case of women of disadvantaged ethnic minorities.

A wide range of literature has already examined the effects of education on teenage childbearing in general, without making a distinction between women belonging to a minority or majority ethnic group of a country. Their identification strategy is usually based on changes in education policy (school entry rules: McCrary and Royer, 2011; length of school day: Berthelon and Kruger, 2011; length of schooling: Black et al., 2008; Silles, 2011; Wilson, 2012, Cygan-Rehm and Maeder, 2013; Clark, Geruso and Royer, 2014). Most papers find that having more education does reduce the probability of early childbearing. The literature considers two possible channels to explain this effect: incapacitation and human capital channels. The incapacitation (or incarceration) effect of education means that teenagers do not have the desire, time, or opportunity to have a child while they are in school. Additionally, according to the human capital effect, an increase in education increases the expected wage rate, which in turn increases the opportunity cost of having a teen birth. Black et al. (2008) find weak evidence of the incapacitation effect and somewhat larger effects through human capital development, while Wilson (2012) finds strong effects through both channels. Berthelon and Kruger (2011) concentrate on the incapacitation effect of education by showing that longer school days reduce the probability of motherhood and criminal behavior among teenage girls aged 15-19.

Despite the large selection of literature that looks at the effects of education on early fertility in general, we know very little about the effects of education on teenage motherhood of ethnic minority women in particular. The only evaluation we have found so far is by Bifulco, Lopoo and Oh (2015). They examine the effects of school desegregation in the US and find that it has not reduced the fertility of black teenagers. They conclude that although school desegregation has generated benefits for black students on several domains, early childbearing is not one of them.

Therefore, this paper examines how longer compulsory education affects teenage childbearing of Roma women. Roma make up the largest ethnic minority in Hungary, accounting for about 5-6% of the total population, or 500,000-600,000 people. Belonging to the Roma minority is highly correlated with poverty, social exclusion, long term unemployment, and access to low quality public services, such as health care and education (Kemény and Janky, 2003; Ladányi and Szelényi, 2002; Kertesi, 2005). Teenage fertility of non-Roma women is very low in Hungary, while among Roma women, it is comparable to levels measured in the Congo and Kenya.

For our analysis, we estimate the effects of longer schooling on the teenage fertility of Roma women by using a legislation change in Hungary in 1996. This new legislation increased the compulsory school leaving (CSL) age from 16 to 18 for those starting elementary school in September 1998. Our identification strategy is based on the age of elementary school entry rule. Children compliant with the age rule started elementary school under the new CSL age scheme if they were born on June 1, 1991, or later. Those compliant with the age rule and born before this date had started elementary school in the previous year, under the old CSL age. Thus, compliance to the age rule creates a discontinuity in the probability of starting elementary school according to the date of birth. Using June 1, 1991, as a cutoff, we construct a fuzzy regression discontinuity design (RDD) identification strategy to estimate the intention-to-treat (ITT) effect of this legislation change on the probability of both early childbearing and pregnancy.

We exploit a unique data set of all known pregnancies, including live births, abortions, fetal losses, and still births, linked to a large subsample of the 2011 Hungarian Census. We find that the higher CSL age decreased the probability of teenage motherhood among Roma women by 6.8 percentage points. This effect is temporary as it only creates a two-year delay in motherhood. We find no effect among non-Roma women, where teenage fertility is rare. We use our rich data sets to reconstruct the conception time of all pregnancies of Roma women. We show that the legislation change decreased the probability of getting pregnant during the school year but not during summer and Christmas breaks. Thus, we find no evidence of any impact through the human capital channel. This result is in line with the argument of the literature: teenage childbearing in disadvantaged ethnic minority communities is more likely to be influenced if human capital development happens alongside increased economic opportunities. In addition to our contribution to the literature on the fertility of disadvantaged ethnic minorities, this is also the first known paper to document that being physically present in school contemporaneously lowers the probability of getting pregnant. Our findings on the incapacitation effect of education are analogous to that of Jacob and Lefgren (2003), who show that teenagers are less likely to engage in criminal behavior on schooldays.

The rest of the paper is structured as follows. Section 2 describes the data and the fertility pattern of Roma adolescents. Section 3 summarizes the legislation change and the institutional background. Our identification strategy and empirical methods are detailed in Section 4. Section 5 presents our results, and Section 6 supplies several robustness checks. Section 7 discusses the potential relevance of delayed motherhood, while our conclusions can be found in Section 8.

2.2 Data and the Teenage Fertility of Roma Women

2.2.1 Data

We begin with an overview of our data sources, the data of which we will refer to in later sections. This paper uses three data sources: the 2001 Hungarian Census, the 2011 Hungarian Census, and the Vital Statistics database. As will be detailed in Section 2.4, we estimate the effects of increased CSL age in a fuzzy regression discontinuity design (RDD) framework, where compliance with the age of elementary school entry rule creates a discontinuity in the probability of being exposed to the new CSL age scheme. We use the 2001 Hungarian Census to demonstrate the first stage of our fuzzy RDD strategy, i.e. the jump in the probability of starting elementary school under the new CSL age scheme around the cutoff, which is June 1, 1991. The 2001 Hungarian Census data were collected in the spring of 2001 when the cohort of interest was 9-10 years old. It contains information on the birth year and month of the individuals, and, for those in school, it registers which grade of school they were attending at the time that the Census data was collected. Knowing their grade level in 2001 allows us to estimate the jump in the probability of starting school under the new CSL age regime in birth year and month bins.

Similarly to the earlier Census, the 2011 Hungarian Census also contains information about the entire Hungarian population. It captures the cohort of interest in October 2011, at ages 19-20. In addition to the birth year and months of individuals, it includes their birth date as well, along with information on gender, ethnicity, and the number of successfully completed years (grades) in school. It registers the year and months of birth of the first five children in a family, independent from whether the children live with their parents or not. We use this 2011 Hungarian Census data to estimate whether Roma women became mothers during their teenage years.

The Vital Statistics databases cover all pregnancy-ending medical events in Hungary. They record live births, still births, abortions, and miscarriages. The data come from surveys completed by women in the hospital at the time of the event. The Vital Statistics registers the mother's day of birth, the date of the event, how far along the pregnancy was at the time of the event (measured by week), the gender of the child in the case of giving birth, and earlier pregnancy history.

We are interested in understanding how school attendance affects Roma girls' fertility decisions. As teenagers usually drop out of school after having a child, we concentrate on their first live births, and pregnancy-ending events before having their first child. The Vital Statistics database does not have ethnic markers, while the 2011 Hungarian Census does. Therefore, we link all pregnancy-ending events of childless women, and all first live births by age 18 to the 2011 Hungarian Census, which will herein be referred to as "linked data".

Our linking procedure is based on the common variables in the two data sets: exact date of birth of the mother, place of residence, and, in the case of live births, the year and month of the event. We can link two observations from the two data sets together if they are not duplicates, i.e. they belong to exactly one woman born on one particular day who lives in one particular settlement. As a consequence, we can only use data of those living in settlements with less than 50,000 inhabitants for the linking procedure. From this subsample, we are able to link 40% of pregnancy-ending events to the 2011 Hungarian Census.

Our identification strategy is based on a discontinuity in date of birth at June 1, 1991 (see Section 2.4 for more details). Considering this fact and the purpose of this analysis, our linking procedure has to fulfill two requirements:

- Linking a certain pregnancy event to the 2011 Hungarian Census has to be random, i.e. it should not be correlated with individual characteristics related to fertility or education; and
- Linking a certain pregnancy event to the 2011 Hungarian Census cannot be correlated with being born right before or after June 1, 1991.

Tables B.1-B.3 in Appendix B show how these two requirements were fulfilled. The records of live births can be linked the most reliably to the 2011 Hungarian Census because the year and month of birth-giving are included in both databases. However, the linking of other pregnancy events seems to be related to both personal characteristics and CSL age legislation in that:

- The events of younger and more educated women are more likely to be linked; and
- The events of those born after June 1, 1991. are less likely to be linked.

Although our linking procedure is not random, the magnitude of the potential bias is small. We estimate the effect of the CSL age increase on outcome variables available both in the Census and the linked data to consider the size of the potential bias, and we find very similar results (see Table 2.8). Furthermore, controlling for year and month of birth, settlement size, and county fixed effects reduces the systematic relationship between the treatment status of women and the link-ability of their Live Statistics events substantially (see Table B.3

in Appendix B). Thus, we reproduce our main estimations using these additional control variables as well (see Table 2.11).

We construct the following outcome variables from the data. From the 2011 Hungarian Census data, we calculate directly at what age women gave birth for the first time between the ages of 16 and 20. We define age at first birth-giving as a continuous variable. Age categories are defined as (0,age]. For example, "having the first child by age 16" means giving birth for the first time at age (0;16]; either before, or exactly on, the day of the mother's sixteenth birthday. Giving birth one day later, therefore, is captured as "having the first child by age 17". We also calculate the approximate conception time of live births by age 18 assuming the pregnancies were 9 months long. In Hungary, the school year starts on September 1 and lasts until the middle of June. To test the incapacitation effect of education, we construct binary variables to capture whether the calculated conception time falls on a date during the school year (September through May) or during the summer break (June through August).

From our linked data, we also calculate whether the woman had her first child by age 18. We use this variable to compare the effects estimated on the 2011 Hungarian Census data to those estimated on the linked data in a robustness check. In addition to that, we construct variables on whether a woman became pregnant by age 18, whether she had an abortion, and whether or not she chose abortion to end her pregnancy, conditional on getting pregnant. The biggest value added of the linked data is from the information on the week of pregnancy at the time of the end-of-pregnancy event. Using this information, we calculate whether the week of conception of pregnancies occurred during the school year. The weekly precision of the data allows us to pin down the school year more precisely than the 2011 Hungarian Census: we are able to deduct the 2-week-long Christmas holiday from the school year and add the first two weeks of June to the school year. Furthermore, we are able to test the effect of the higher CSL age on conceptions during Christmas breaks from our linked data, which is not possible from the 2011 Hungarian Census.

2.2.2 Identifying Roma in the 2001 and 2011 Hungarian Census Data

This analysis is interested in the effects that the increased CSL age has on the teenage fertility of Roma women. Both the 2001 and the 2011 Hungarian Census rely on individuals' self-identification of Roma ethnicity, and both suffer from an underreporting problem to different extents. The 2011 Hungarian Census registers almost 50% more Roma people than the 2001 Census for three reasons. First, it allows for the respondents to have dual-nationality identities, whereas the 2001 Hungarian Census only allows three choices of answers to the question about national identity. The 2011 Hungarian Census first has a question about national identity, set up similarly to that of the 2001 Hungarian Census. Then, it explicitly asks the respondents whether they feel that they belong to any other nationality or ethnicity group, in addition to the one which was identified in the first question. Compared to the previous census, this new method allows for substantially improved identification of the Roma population (Messing, 2011). Second, civil organizations campaigned in 2011 to increase the number of Roma who revealed their ethnic identity. In some 2011 Hungarian Census tracts, census takers who themselves identified as Roma were employed (Budapest Institute, 2013). Unfortunately, the details of these initiatives are not documented publicly. According to statistics on the "We belong here!" (Ide tartozunk!) campaign, 1,046 Roma survey takers were employed. However, there is no data on which census tract they worked, or on how many questionnaires each that they registered (Data source: Open Society Foundation)[1]. Lastly, there must have been demographic changes in the Roma population between 2001 and 2011, including population growth and geographical migration. However, we have no information on the magnitude of these phenomena.

As a result of these three factors, the 2001 Hungarian Census reported a total of 217,097 Roma in Hungary, while the 2011 Hungarian Census reported 315,525 Roma[2]. In spite of this improvement, demographers estimate that the 2011 Hungarian Census still identifies only half of the total number of the Roma population (Hablicsek, 2007). We know from a regular household survey using both self- and public-assessments of Roma identity that individuals identifying themselves as Roma are more likely to live in smaller settlements, to be less educated, to be unemployed, and to live on lower and less stable incomes, than those identified as Roma by others but not by themselves (Tárki, 2013). Those who identify themselves as Roma in the Census thus belong to the lower half of the income distribution of the Roma in Hungary.

[1] The total number of Census tract units is over 47,000 (Data source: Hungarian Statistical Office).
[2] Data Source: own estimation from the 2001 and 2011 Hungarian Censuses.

We can estimate the consistent effects of the CSL age increase on Roma women if a declaration of being Roma is not related to the legislation change. This issue is discussed in Section 2.4.

2.2.3 Teenage Fertility in Hungary

Table 2.1 summarizes the timing pattern of first birth-giving of teenagers born before the CSL age increase. The prevalence of having the first child by age 18 is low on average (3.2%), and lower among non-Roma women (2.0%). However, it is very high among Roma women (26.0%). Almost half (47.2%) of Roma women become a mother by age 20. The share of teenage mothers among Roma women is similar to rates found in developing countries like Kenya, Eritrea, or the Congo (UNFPA, 2013).

Table 2.1: The Prevalence of Teenage Childbearing before the Legislation Change (cohorts born 1988-1990)

Probability of giving birth by reaching age ...	All women	Roma women	Non-Roma women	Ethnicity unknown
16	0.007	0.065	0.004	0.007
	(0.000)	(0.003)	(0.000)	(0.085)
17	0.017	0.150	0.010	0.019
	(0.000)	(0.004)	(0.000)	(0.136)
18	0.032	0.260	0.020	0.038
	(0.000)	(0.005)	(0.001)	(0.192)
19	0.052	0.379	0.034	0.063
	(0.001)	(0.006)	(0.001)	(0.243)
20	0.073	0.472	0.052	0.090
	(0.001)	(0.006)	(0.001)	(0.286)
No. of obs.	169,375	8,020	154,055	7,300

Data source: estimation based on the 2011 Hungarian Census data. Standard errors of the group means are in parenthesis. Age is a continuous variable; "by reaching age 16", for example, means giving birth either before, or exactly on, the mother's 16th birthday.

There could be several explanations for the prevalence of early childbearing among Roma women but little evidence exists on this. Teenage fertility has been historically high among Roma women; a 14-year-old girl or boy is treated as an adult in Roma communities. Most children of adolescent Roma mothers are born in stable relationships. Out of Roma mothers born in 1990 who gave birth by age 18, 81% were either married or lived with a long term partner in 2011[3].

Although teenage fertility has been decreasing in Hungary on average, even among Roma women, this is not observed within some groups of poor (Szikra, 2010). In particular, in some marginalized Roma communities, fertility is either on the rise or it has stabilized at a high rate (Durst, 2007). It should be noted, however, that Roma communities themselves are heterogeneous with respect to their fertility patterns (Janky, 2005).

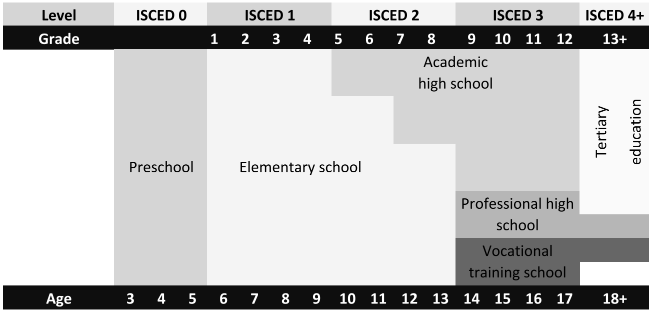
2.3 Institutional Background and the Legislation Change

2.3.1 Roma Students in the Hungarian Education System

Typically, Hungarian students reach CSL age while in secondary school. There are three types of secondary schools in Hungary: vocational training school, professional high school, and academic high school (see the structure of the Hungarian compulsory education system in Figure 2.1). There is a difference to the returns of vocational training school from the other two types of high schools. Vocational education is considered to be a dead end in the sense that, although theoretically it is possible, most students do not continue their studies after completing vocational training school (REF). The Mincer-type returns of vocational education are low but still positive in comparison to the returns of an elementary school degree (Kézdi, Köllő and Varga, 2009).

[3] Own calculation from the 2011 Census.

Figure 2.1: The Structure of the Hungarian Secondary Education System



Source: Horn, 2014. For the cohort of interest, preschool was compulsory from age 5. Compulsory school leaving (CSL) age is 16 in the case of those starting elementary school in September 1997 or earlier, age 18 in the case of those starting elementary school in September 1998, and again age 16 for those in any grades starting from September 2012.

A high school degree, on the other hand, brings substantial benefits[4]. First, it is a prerequisite for starting tertiary education. Secondly, it gives high returns in the labor market, both in terms of employment probability, and wages. The average wage advantage of a high school degree is estimated to be about 25-30% higher when compared to a vocational degree, amounting to around 30,000 USD during a lifetime (Hajdú et al., 2015).

Prior to the CSL age increase in 2001, 9% of 15-year-old non-Roma women were attending primary school, and 88% were attending secondary school right before reaching the actual CSL age of 16 that was in place at that time. The rates for Roma women were 46% and 29%, respectively[5]. Out of those in secondary school at age 15 in 2001, 12% of non-Roma women were in vocational training schools, and 46% of Roma women.

Besides being more likely to repeat grades and end up in vocational training schools, Roma students, both male and female, lag behind in several aspects. The achievement gap in standardized reading and math test scores between Roma and non-Roma students is comparable to the size of the black to white test score gaps of the United States in the 1980s (Kertesi and Kézdi, 2011; Kertesi and Kézdi, 2014). About 13-50% of this gap comes from the fact that Roma students do not have access to high-quality education, and the remainder is accounted for by differences in social backgrounds (Kertesi and Kézdi, 2014). The Hungarian education system is rated as one of the worsts among the OECD countries when it comes to dealing with social disadvantages. According to the 2012 Program for International Student Assessment (PISA) study, family background explains one of the largest shares of the variance in mathematics test results in Hungary in this country group (OECD, 2014). With free elementary school choice and early tracking, the Hungarian education system is highly segregative with respect to disadvantaged students in general, and with respect to Roma students in particular (Kertesi and Kézdi, 2009).

2.3.2 Compulsory Education and the Legislation Change

Before the legislation change, students were obligated to attend school until the end of the academic year in which they turned 16. The Public Education Act (1996) increased compulsory school attendance from age 16 to age 18, requiring students to spend two more years in the education system. The new legislation was

[4]Both high schools involve the completion of a "maturity exam" ("érettségi") at the end of Grade 12, which is similar to "Matura" or "Baccalaureat" examinations found in many European countries (Kézdi and Surányi, 2008).

[5]Own estimation from the 2001 Census. For the rest of the sample (3% and 15%, the information on the type of school is either missing, they live in institutions, or they are not in school due to living with disabilities.

Compulsory schooling obligation can be fulfilled in homeschooling as well. However, even those in homeschooling belong to a school and are supposed to be covered by education statistics as those in school. Homeschooling is rare; in the 2013/14 academic year, the share of those in homeschooling was 0.68%. (Education Office of the Ministry of Human Resources, 2014) We know very little about homeschooling in general.

grandfathered in, so that it first became binding with those students starting elementary school in September 1998. Thus, students already knew by age 6 that they would have to stay in school two years longer. Although the Act introduced other measures are well, the increase in the CSL age was the only element causing sharp changes for those starting elementary school in the 1997/98 academic year versus those starting in the 1998/99 academic year. The Act also prescribed the gradual adaptation of the secondary school structure to meet the new CSL age by forcing all secondary school programs to have at least 4 grades (and thus not to end before age 18), a process that began during the 1998/1999 academic year. As a result, the first treated cohort was beginning secondary school at a time when the adjustments to secondary school program length had been adapted half a decade earlier.

Primary school in Hungary starts at age 6 and has 8 grade levels. According to the age of elementary school entrance rule, compulsory schooling starts on September 1 of the same year in which a student reaches age 6 by May 31. Those born on June 1, or later, during the same year start elementary school one year later. Thus, those compliant with the age rule, and born before June 1, 1991, entered elementary school in 1997 under the old CSL age scheme. Those compliant with the age rule, and born on June 1, 1991, or later, entered school in 1998 under the new CSL age scheme. This discontinuity in date of birth at June 1, 1991 is the base of our identification strategy.

In addition to the age rule, the school starting year is a joint decision of parents, preschool teachers, and in some cases, pedagogical and psychological counselors employed by public pedagogical service centers [6]. The decision itself is made during preschool. At the time of the legislation change, preschool attendance was compulsory from age 5. The decision process about elementary school entry time begins with an official opinion from the preschool teachers about whether the child is ready to start school. In the case of any doubts, preschool teachers can request a "school readiness examination" from the local pedagogical service center.

Table 2.2: Compliance to the Age Rule of Compulsory Schooling

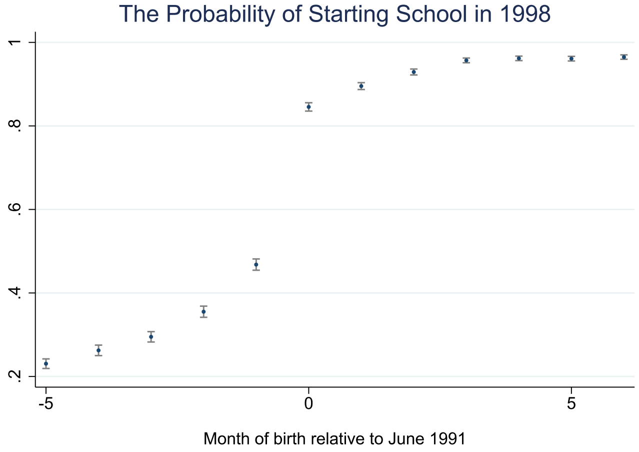
	Share of early starters	Share of compliars	Share of late starters	No. of school starters
Compliance by Academic Years, share of those starting elementary school at the given year (Total=100%)				
1997/1998	0.02	0.80	0.18	127,214
1998/1999	0.02	0.78	0.20	125,875
1999/2000	0.01	0.78	0.21	121,424
Compliance by Cohorts, share of cohort size (Total=100%)				
June 90-May 91	0.02	0.79	0.19	129,489
June 91-May 92	0.02	0.78	0.20	126,294

Data source: aggregate data from the Public Education Information System - Public Education Statistics (*KIR-STAT* in Hungarian). Individual-level data is not available for this period. "Early starters" refers to those entering elementary school before reaching age 6 by May 31. "Compliers" refers to those entering elementary school according to the age rule of compulsory schooling. "Late starters" refers to those entering elementary school a year later than what is determined by the age rule.

On average, compliance with the age rule is 78-80%. About 18-20% of a cohort start elementary school a year later then they are supposed to according to what is expected based on their date of birth. Early school start is rare, occuring at less than 2% (see Table 2.2). The share of late starters is the highest among those born right before the cutoff. Among women born in May 1991, 47% of them started elementary school under the higher CSL age scheme, while 85% of the women born in June 1991 began school under the higher CSL age scheme (see Figure 2.2).

[6] Pedagógiai Szakszolgálat in Hungarian.

Figure 2.2: The Probability of Starting School Under CSL Age 18, Women Born in 1991



The average probability of starting school in 1998, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June 1991. No. of individual observations: 60,302.

2.4 Identification Strategy and Empirical Methods

2.4.1 Identification Strategy

As detailed in Section 3.2, our identification strategy is based on compliance with the age of elementary school entry rule, which creates a discontinuity in the probability of starting elementary school under the increased CSL age regime at the date of birth of June 1, 1991. This discontinuity allows us to identify the ITT effects of the increase using an RDD, with being born on June 1, 1991, as the cutoff.

Table 2.3: Grade Repetition in Grades 1-3, % of Students in Grade

Academic year	Grade 1	Grade 2	Grade 3	CSL age
1995/1996	4.0	1.9	1.6	16
1996/1997	3.9	2.0	1.5	16
1997/1998	3.9	1.9	1.5	16
1998/1999	4.0	1.8	1.5	18
1999/2000	3.9	1.9	1.4	18
2000/2001*	4.2	1.9	1.4	18

*The 2000/2001 is the academic year in which the 2001 Census is taken. Data Source: National Institute of Public Education, 2006. Table 4.28 in the Appendix, page 478.

Establishing the first stage relationship, i.e. estimating the size of the jump in the probability of starting school under the new CSL age scheme around the cutoff, is straightforward for all women in general, but challenging for Roma women in particular. As mentioned in Section 2.2, we observe compliance with the age rule in birth year and month bins from the 2001 Hungarian Census. The 2001 Hungarian Census registers which grade of school the cohort of students born in 1991 was attending at the time the data was collected. Knowing their grade level in 2001 allows us to estimate the size of the jump, assuming that grade repetition patterns did not change between 1997 and 2001. Those who started elementary school in 1997, under the old CSL age scheme, were in Grade 4, and those who started elementary school in 1998, under the new CSL age scheme, were in Grade 3 in 2001. If we see an individual in a lower grade than where s/he is supposed to be based on the age rule, it means one of two things: s/he started school later, or s/he repeated grades. Although the 2001 Hungarian Census does not register grade repetitions, this information is available on an aggregate level from the Public Education Statistics of the Public Education Information System database. Table 2.3 summarizes the share of students repeating a year during Grades 1-3, and shows that the average prevalence of grade repetition is quite low, 1.4-4.0%, and its pattern does not change during this period

(National Institute of Public Education, 2006). The prevalence of grade repetition is the highest in Grade 1, and it is lower, at around 1.5%, in Grades 2 and 3. Those who started elementary school in 1997 should have been in a higher grade in 2001 and thus, they are 1.5 percentage points more likely to have repeated a grade by the time of observation. Consequently, the only problem with estimating the size of the jump in general is that this method will result in estimating a 1.5 percentage point lower jump in the probability of starting school under the new scheme because we cannot distinguish between a late school start and a grade repetition.

Table 2.4: Grade Repetition in Grades 1-4, % of Students in Grade

No. grade repetitions	Non-Roma students	Roma students	Roma girls	Roma boys
None	94.51	80.04	80.86	79.24
One	4.89	17.12	16.36	17.87
Two	0.40	2.20	1.86	2.53
More than two	0.02	0.37	0.37	0.36
No data	0.15	0.27	0.56	0.00
Total no. of obs.	8,930	1,092	538	554

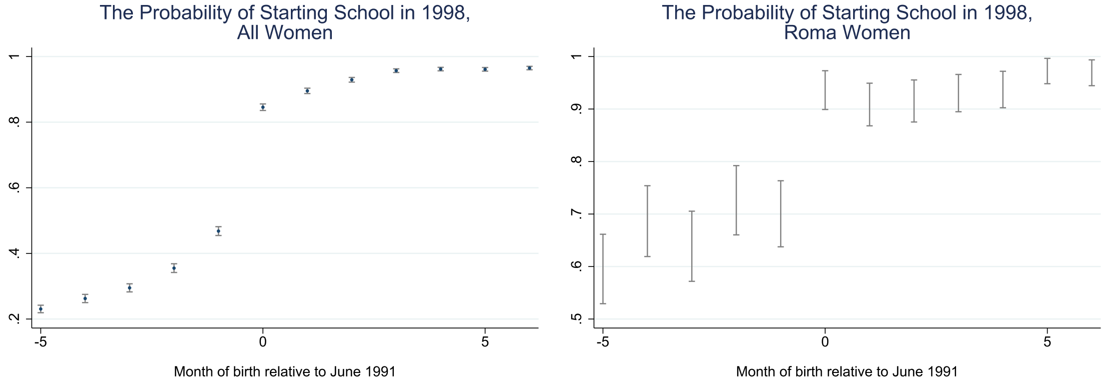
Data Source: own estimation from the Hungarian Life Course Survey.

We face several challenges when establishing the first stage relationship for Roma women in particular. First, although the 2001 Hungarian Census has ethnic markers, the underreporting problem is much more serious there than it is in the 2011 Hungarian Census (see Section 2.2). Second, the probability of repeating a grade at least once during Grades 1-4 is five times higher among Roma than among non-Roma students. The only data on grade repetition of Roma students in Hungary come from the Hungarian Life Course Survey[7]. The Hungarian Life Course Survey is a cohort study that follows 10,000 youth, who were in Grade 8 in the spring of 2006 (Hajdú, 2015). As Table 2.4 demonstrates, only 5% of non-Roma students repeated a grade at least once during grades 1-4, whereas 20% of Roma students repeated at least once. Thus, our method underestimates the size of the jump around the cutoff in the case of Roma students. Figure 2.3 compares the jump in the probability of starting school in 1998 between all women and Roma women specifically. The data show three main differences:

- The data of Roma women have a higher variance, and the 95% confidence intervals are much larger. This is partly due to the smaller sample size of Roma women (2,313 vs. 60,302).
- The main difference is on the left hand-side of the cutoff: Roma women born before the cutoff are more likely to be found in a lower grade than they are supposed to be. This phenomenon is due to the high number of grade repetitions.
- Due to grade repetitions, the estimated jump in the probability of starting school under the new CSL age regime is lower for Roma women than for all women in general (36 vs. 24 percentage points).

[7] http://edecon.mtakti.hu/?q=node/16

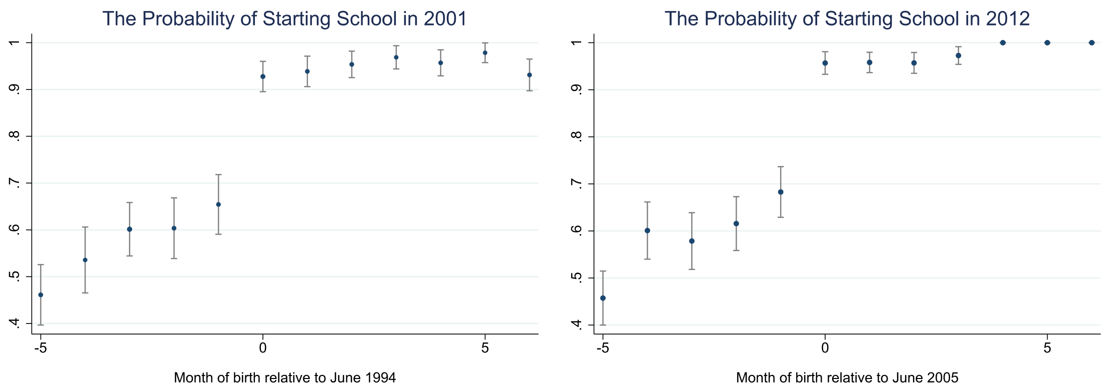
Figure 2.3: The Probability of Starting School in 1998, All Women vs. Roma Women



The average probability of starting school in 1998, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June 1991. No of individual observations: 60,302 and 2,313.

Comparing the share of Roma girls who have just started elementary school, and thus, have not been able to repeat a grade, to those who have not started school yet gives a better estimation about the size of the jump. The relevant cohorts for this comparison consist of those born in 1994 in the case of the 2001 Hungarian Census, and those born in 2005 in the case of the 2011 Hungarian Census[8]. For these cohorts, the probability of not starting school yet is equivalent to the treated status defined above, i.e. starting school the next year. Figure 2.4 proves that once grade repetition is out of the question, the size of the jump is indeed higher, at 28 percentage points in both Censuses.

Figure 2.4: The Probability of Starting School in 2001 and 2012, Roma Women



The average probability of starting school in 2001 and 2012, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 and 2011 Hungarian Census. 0 on the x-axis refers to being born in June 1991. No. of individual observations: 2,635 and 3,407.

As our data does not allow us to estimate the exact size of the jump for Roma women, this analysis will estimate ITT effects only. This is a fuzzy RDD setup where compliance with the age of elementary school entry rule is endogenous. Being born right after, rather than before, the cutoff is used as an instrument for starting school under the new CSL age legislation. We identify the ITT effects of the increase in a potential outcome framework. Consider the population of Roma women born in a small neighborhood within the cutoff date and define Z as:

$$Z = 1(\text{being born on or after June 1, 1991}).$$

[8]The difference between 2001-1994=7 and 2011-2005=6 comes from the fact that the 2001 Census was taken in the spring while the 2011 Census was taken in the autumn.

Z is a binary instrumental variable for the treatment, i.e. entering elementary school under the new CSL age scheme, for this population. The potential treatment indicators are then defined as:

$$D(0) = 1(\text{entering elementary school under CSL age 18 if one had Z=0})$$

$$D(1) = 1(\text{entering elementary school under CSL age 18 if one had Z=1}).$$

The actual treatment indicator is $\mathrm{D} = \mathrm{ZD}(1) + (1 - \mathrm{Z})\mathrm{D}(0)$. For those born before the cutoff, or $\mathrm{Z} = 0$, the intended school starting date is 1997, which falls under the old CSL age of 16. However, some parents choose to withhold their child to start school one year later, in 1998, which falls under the new CSL age of 18. Therefore, before the cutoff:

$$D(0) = 0 \text{ for compliers and}$$

$$D(0) = 1 \text{ for always-takers (i.e., late starters).}$$

Both $D(0) = 0$ and $D(0) = 1$ are possible, and the choice between them is endogenous. For those born after the cutoff, or $Z = 1$, the intended school start date is 1998 under the new CSL age of 18. Therefore, after the cutoff:

$$D(1) = 1 \text{ for compliers, and}$$

$$D(1) = 0 \text{ for never-takers (i.e., early starters).}$$

The size of the jump in the probability of starting school under the new CSL age scheme is $P(D(1) = 1) - P(D(0) = 1)$, which is estimated to be around 0.3 (see Figure 2.3 and 2.4).

Our identification strategy is based on five assumptions. First, we assume that the instrument is exogenous: whether the student was born right before or after the cutoff is random. As the Act was introduced in 1996, five years after the relevant cohort was born, manipulation of birth because of the legislation change is not an issue. However, there is a dispute in the literature about whether children born in different months of the year are inherently different from each other with regard to their outcomes later in life (see Buckles and Hungerman, 2013 vs. Fan et al., 2014). The primary concern within the literature is related to those born during the winter as opposed to those born in spring. This literature argues that babies born in the winter are more likely to be born to less educated women, and that the winter months may not provide as favorable of an environment to a newborn as the spring. In this paper, we compare children born before and after June 1. To the best of our knowledge, the question of whether or not children born in May or in June are inherently different has not been raised in the literature. In spite of this, the exogeneity assumption of the date of birth is going to be relaxed for a robustness check in Section 2.6.3.

The second is the exclusion restriction that whether one was born right before or after the cutoff does not affect fertility decisions if not through the legislation change. This is not a trivial assumption in this case. Specifically among those staying in school until the new CSL age, starting elementary school at an age closer to 6 rather than closer to 7 may result in the student spending a longer or shorter time in school, independent from the legislation change (Angrist and Kruger, 1991; Hámori and Köllő, 2011). Without the legislation change, both those born in May and June 1991 would have to stay in school until the end of the 2006/2007 academic year. However, those born in May, if compliant with the age rule, were supposed to start elementary school one year earlier. Although the CSL age increase might have created two extra years to the schooling career of those born in June, the net impact of the legislation change around the cutoff on the compliers is closer to one year. We ran two robustness checks to learn more about this issue. In Section 2.6.2, impacts estimated around the true cutoff in 1991 are compared to cutoffs in other years. We find no effects around cutoffs in other years; the effects of starting school at an older age but spending less time in school balance each other out[9]. In Section 2.6.3, we directly control for any potential impacts around the same cutoff in other years.

The third assumption is that the instrument is continuous and no defiers exist. On the one hand, we assume that being born after the cutoff and thus being subject to an increased CSL age did not induce anyone to start school one year earlier to avoid the extra two compulsory years in school. This assumption would theoretically be violated if parents who "dislike" schooling want to manipulate the system and therefore send their children born after the cutoff to school one year earlier to "save" them from the extra two compulsory school years. Similarly, we also assume that none of those born before the cutoff are sent to school one year later for the purpose of being subjected to the new CSL age. We assume that late starters (always-takers) on the left side of the cutoff, and early starters (never-takers) on the right side of the cutoff, start school late or early because of their general preferences on what is the ideal time to start school, and not because they want to violate the age rule due to the legislation change.

Such violations are highly unlikely because the timing of the legislation change is not in favor of those who might want to manipulate school start. The CSL increase was accepted in 1996, the first treated cohort started school in 1998, and they reached age 16 in 2008. To start school early in 1997, parents "disliking" school had to have been aware of the increase already in late 1996 or early 1997 to ask for an early school readiness examination in pre-school. Pre-school was compulsory from age 5 at that time, and some potential early-starters had just entered pre-school when the new legislation came out. Assuming that those parents who "dislike" school might be of a lower socioeconomic background, it is unlikely that they are informed of the details of the legislation change far enough in advance. The increase of the CSL age, as it became practically binding only 12 years later, did not receive much attention in the media in 1996. The Act made several prompt changes in the education system, and the media concentrated much more on those instead . It also seems quite unlikely that pre-school teachers suggested to parents in 1996-1997 to send their children to school early to avoid longer schooling. It sounds more reasonable that the information about longer schooling reached parents at the time that their children entered elementary school in 1998. By then it was too late to avoid longer compulsory school attendance. Furthermore, it was also unlikely that parents who "like" schooling manipulated the system by sending their children born before the cutoff to school a year later because CSL age is binding downwards. Parents who were fond of schooling could keep their children in school until whatever age they prefer, regardless of the actual CSL age legislation.

Additionally, we find no evidence in the data that such defiance occurred. On an aggregate level, the administrative data of the Public Education Statistics of the Public Education Information System[10], which is the Hungarian school census, show the share of early starters (1-2%), compliers (78-80%), and late starters (18-20%) in Grade 1 as stable between 1997 and 1999. This was true both across cohorts measured by date of birth, and measured by academic years (see Table 2.5).

Table 2.5: Compliance with the Age Rule of Compulsory Schooling, PES Data

	Early starters (never-takers)	Compliers	Late starters (always-takers)	No. of students in Grade 1
Compliance by academic years, share of those starting elementary school at the given year				
1997/1998	0.02	0.80	0.18	127,214
1998/1999	0.02	0.78	0.20	125,875
1999/2000	0.01	0.78	0.21	121,424
Compliance by cohorts, share of cohort size				
Born between June 90-May 91	0.02	0.79	0.19	129,489
Born between June 91-May 92	0.02	0.78	0.20	126,294

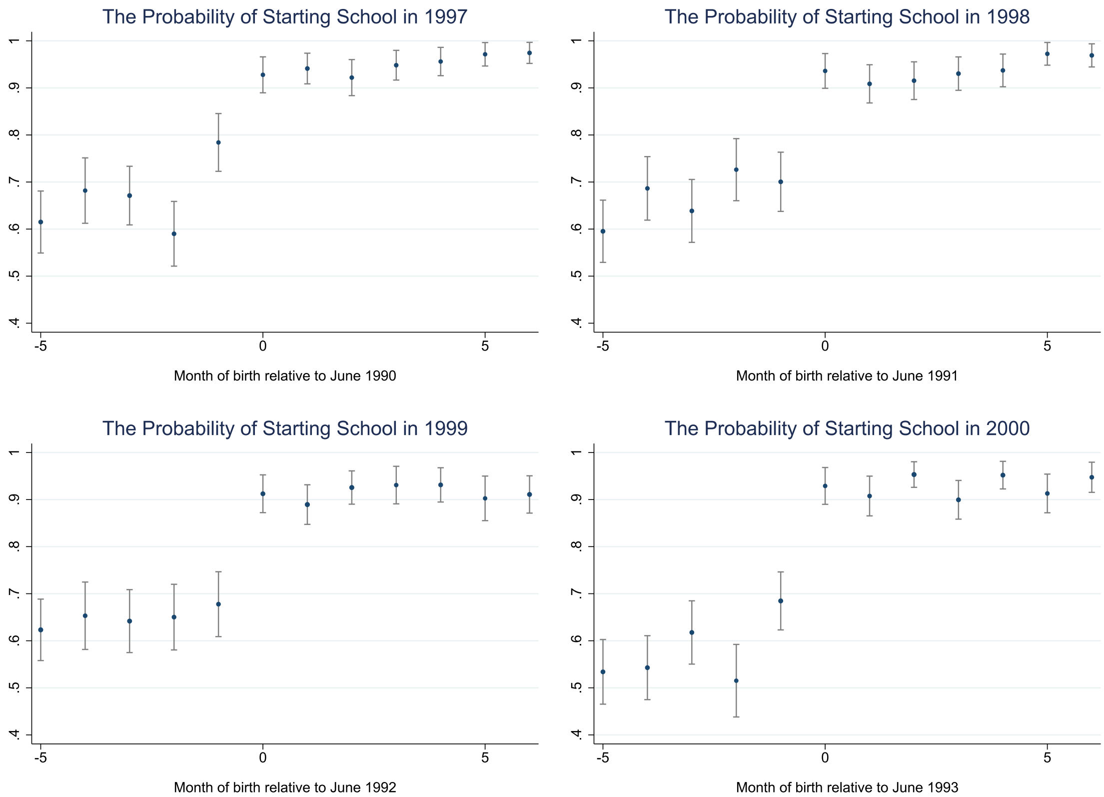
Data Source: Public Education Statistics (PES) of the Public Education Information System (KIR-STAT). "Early starters" refers to those entering elementary school without reaching age 6 by May 31. "Compliers" refers to those entering elementary school according to the age rule of compulsory schooling at age 6. "Late starters" refers to those entering elementary school a year later than expected under the age rule.

[9]This issue is discussed further in 2.6.2.

[10]KIR-STAT in Hungarian

In the case of Roma women in particular, although we systematically underestimate the size of the jump in the probability of starting school under the CSL age scheme from the 2001 Hungarian Census, and although the data is noisy, it does allow us to estimate the school starting patterns of those born between 1990 and 1993, organized by birth month bins. Figure 2.5 shows that the share of late starters and early starters do not differ significantly across these four years around the June 1 cutoff.

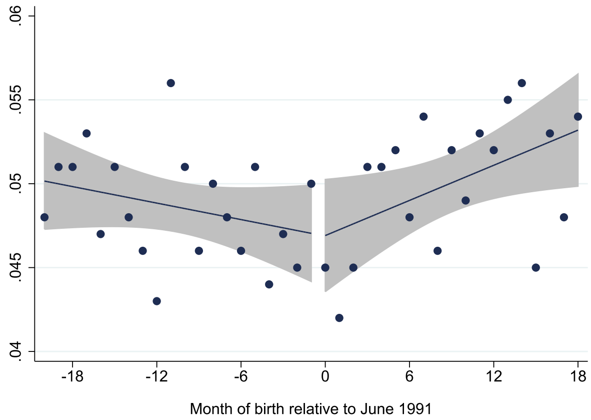
Figure 2.5: The Probability of Starting School in 1997-2000, Roma Women



The average probability of starting school in 1997-2000, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June 1991. No of individual observations: 2,308, 2,313, 2,282, and 2,386.

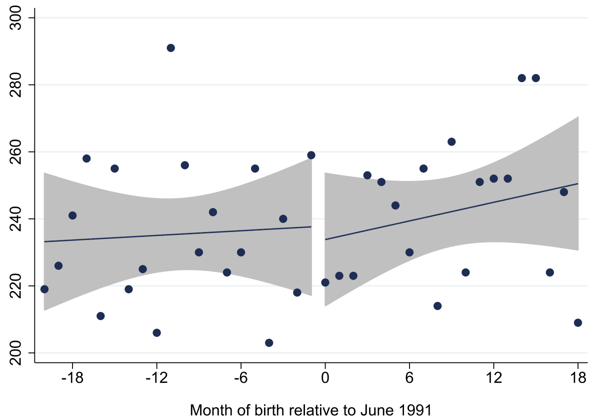
The fourth identification assumption is that the fact whether one reports herself as Roma or not is independent to the CSL age change. As detailed in Section 2.2, ethnicity is self-reported, and roughly every second Roma person does not report himself or herself as Roma in the 2011 Hungarian Census. In general, the effects of longer schooling on human capital accumulation could impact self-assessment in two ways. If human capital development induces individuals to be more conscious about their identity, those under the new scheme of the higher CSL age could be more likely to reveal their Roma ethnicity, meaning we overestimate the effect of the change. However, this works in an opposite fashion if human capital development induces people to hide their minority status. There is some evidence for such phenomenon in relation to the black population in the US (Fryer and Torelli, 2010). If this happens, our results would have a downward bias. To test the validity of this assumption, we compare the share of women identifying themselves as Roma below and above the cutoff. Figure 2.6 shows that we find no significant discontinuity in the share of Roma women at the cutoff. The same is true for the number of Roma women around the cutoff (see Figure 2.7).

Figure 2.6: The Share of Roma Women by Month of Birth, 1991



Data source: own estimation from the 2011 Hungarian Census. 0 on the x axis refers to being born in June 1991. Linear regression lines estimated separately below and above the cutoff, plotted with 95% confidence intervals. No. of individual observations: 263,298.

Figure 2.7: The Number of Roma Women by Month of Birth, 1991



Data source: own estimation from the 2011 Hungarian Census. 0 on the x axis refers to being born in June 1991. Linear regression lines estimated separately below and above the cutoff, plotted with 95% confidence intervals. No. of individual observations: 16,667.

Fifth, in addition to Roma self-identification being independent from the increased CSL age, we also assume its independence from teenage motherhood. In addition to a legislation change, ethnic identification may be affected by individual circumstances as well. Kézdi and Simonovits (2016) examine the causal effects of economic hardship on Roma identification of adolescents in Hungary. They build on a theory that "individuals are more likely to categorize themselves as members of a group if they perceive themselves to be more similar to the other members of that group". They show that adolescents of Roma descent are more likely to identify themselves as Roma if their family experienced economic hardship since being Roma is highly associated with poverty. It would be problematic if a similar mechanism occurred with respect to teenage childbearing, and that teenage mothers were more likely to identify as being Roma. Kézdi and Simonovits (2016) do not examine the causal effect of early childbearing on Roma identification. However, in their regressions of Roma identification on economic hardship, they do control for having a child. They find a negative and insignificant correlation between having a child in teenage years and identifying as Roma in the case of those of Roma descent. On the other hand, there is a positive correlation that is not significant in all specifications among those who are not of Roma descent.

We cannot test the causal relationship between teenage motherhood and Roma identification using our data. However, this paper concludes that increasing the CSL age decreased the probability of teenage pregnancy and motherhood among Roma women. If becoming a teenage mother made women more likely to identify as Roma, and increased CSL age decreases the probability of teenage motherhood among *some* women, it is possible that we underestimate the effect that the legislation change had on Roma women.

2.4.2 Empirical Methods

We estimate the ITT effects of the CSL age legislation change around the June 1, 1991 cutoff using both a nonparametric and a parametric estimation strategy. **Nonparametric estimates** are generated by estimating weighted local linear regressions on both sides of the cutoff, within a certain bandwidth (Hahn et a., 2001; Imbens and Lemieux, 2008). For simplicity, weights are computed by applying a rectangular kernel function to the distance from each observation to the cutoff in terms of day of birth. This is the standard method of RDD estimation as it has excellent properties in estimating the difference of two conditional expectations evaluated at the boundary points of the cut-off (Cheng, Fan and Marron, 1997).

The following local linear models are estimated within a certain bandwidth:

$$y_i = \alpha_{NP} + \beta_{NP} * itt_i + \gamma_{NP} * x_i + \delta_{NP} * x_i * itt_i + \varepsilon_i$$

where

y_i is the outcome variable;
itt_i is the intention-to-treat variable, which is 1 if individual i was born on June 1, 1991, or later, and 0 otherwise;
x_i is the running variable, number of days in date of birth before or after June 1, 1991 (and 0 if individual i was born on June 1, 1991);
$x_i * itt_i$ is an interaction term of x_i and itt_i, allowing for the local linear function to be different on the two sides of the cutoff; and
β_{NP} captures the ITT effect of the legislation change estimated using our nonparametric approach.

We follow a conservative strategy and use the strictest procedure to set the bandwidth of the local linear regressions: the optimal bandwidth routine from Calonico, Cattaneo and Titiunik (2014), abbreviated as CCT for the remainder of the paper, along with the 50-150% versions of the optimal bandwidths as robustness checks. The optimal bandwidths set by the method are 150-200 days wide below and above the cutoff, depending on the outcome variable and the sample.

A **parametric approach** is used as one of the robustness checks on the sample of individuals born in 1980-1993 using 4th order global polynomial models. The estimated parametric models are the followings:

$$y_i = \alpha_p + \beta_P * itt_i + f(x_i, itt_i) + u_i$$

where

$f(x_i, itt_i)$ is a 4th-order polynomial function of the running variable, which is different on the two sides of the cutoff; and
β_P captures the ITT effect of the legislation change estimated using our parametric approach.

There are two reasons to complement the nonparametric analysis with parametric models. First, they can accommodate additional control variables. In particular, birth month fixed effects are used to capture the impacts that any potential monthly seasonality has on the characteristics of the child, and thus relaxes the assumption of exogenous variation in month of birth, and birth year fixed effects are used to capture the potential effects of business cycles. An interesting feature of the Census data (and the Hungarian health system) is that the day of the week matters with respect to the probability of being born. That is, a child is more likely to be born Tuesday through Friday than Saturday through Monday, and this probability difference is weakly related to the educational status of the mother[11]. This paper will not document this phenomenon. However, because June 1 in 1991 fell on a Saturday, the day of the week fixed effects are also included in the parametric models to control for this pattern.

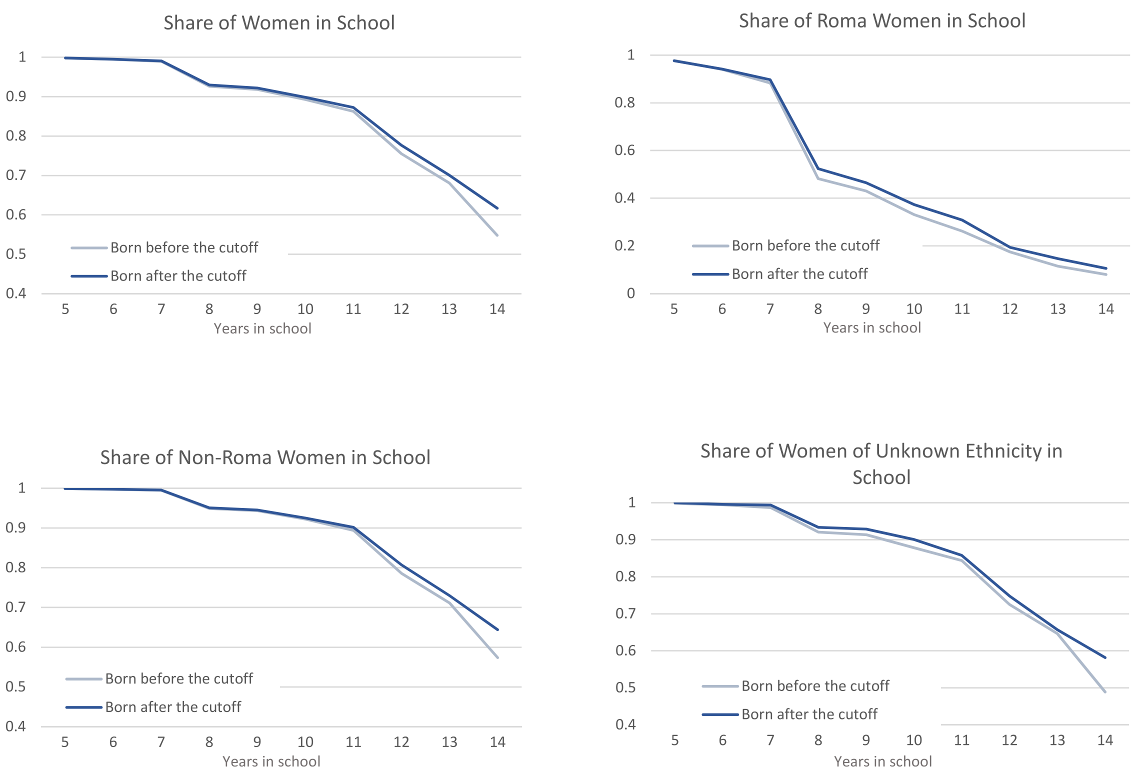
[11]Own estimation from the 2011 Census.

Second, we use two databases that are linked together. Linking observations related to one woman between the two data sets is not random. However, the systematic relationship between the linking procedure and being born after the cutoff disappears after controlling for additional individual characteristics (see Section 2.2.1). We do not control for these characteristics in our nonparametric RDD strategy, only in our parametric approach. Finding the same results in both approaches supports that any potential bias coming from non-random linking must be small.

2.4.3 The Number of Completed School Years Around the Cutoff

As detailed in Section 2.2, our data do not allow us to directly observe how long students stayed in school. We observe the number of successfully completed school years in the 2011 Hungarian Census at ages 19-20 but the data do not capture unfinished, incomplete years in school. In spite of this fact, the impact of increased CSL age is visible in the data among non-Roma women. Figure 2.8 shows the share of women that completed 1,2,...,13 years (i.e., grades) in school. Non-Roma women born right after the cutoff have successfully completed more school years by 2011 than non-Roma women born before the cutoff. However, the difference in the number of successfully completed school years is negligible among Roma women born around the cutoff. In line with the findings of Chapter 1, the legislation change has no affect on the education outcomes of Roma women.

Figure 2.8: The No. of Completed School Years Among Those Born Right Before and Right After the Cutoff (Kaplan-Meier survival functions with respect to the no. of successfully completed years in school)



Share of women who successfully completed a given number of school years. Data source: own estimation from the 2011 Hungarian Census. Born before the cutoff (control group): born at most 180 days before June 1, 1991. Born after the cutoff (intention-to-treat group): born at most 180 days after June 1, 1991. Number of observations: 56,628, 2,775, 53,634 and 2,219, respectively.

This result does not imply that Roma women born after the cutoff did not physically stay in school longer

10.14754/CEU.2016.11

than before the CSL age increase. We know from the qualitative study of Mártonfi (2011a and 2011b) that most Roma (and non-Roma) students followed the new CSL age legislation and showed up in school at older ages, as required. There are two possible reasons why we do not see the effects of the higher CSL age on the length of time that Roma women stayed in school. First, as mentioned in the previous paragraph, we do not have proper data to measure their length of schooling. Second, the Hungarian education system has long faced problems with providing quality education for students of various backgrounds (OECD, 2015), and especially for the Roma (Kertesi and Kézdi, 2009). There is a huge variation in the quality of education in Hungary, and Roma students are likely to go to low quality schools (Kertesi and Kézdi, 2009). Thus, even though Roma students did stay in school for longer, it is likely that they did not successfully complete more academic years on average.

2.4.4 Teenage Fertility Around the Cutoff

Table 2.6 presents the probability of having that a woman has her first child between ages 16 and 20 for both Roma and non-Roma women. According to this raw data, there is no difference in the prevalence of teenage motherhood between non-Roma women born right before or after the cutoff. However, among Roma women, the probability of beginning motherhood by age 18 and 19 is lower in the group for those women born right after the cutoff.

Table 2.6: Descriptive Statistics

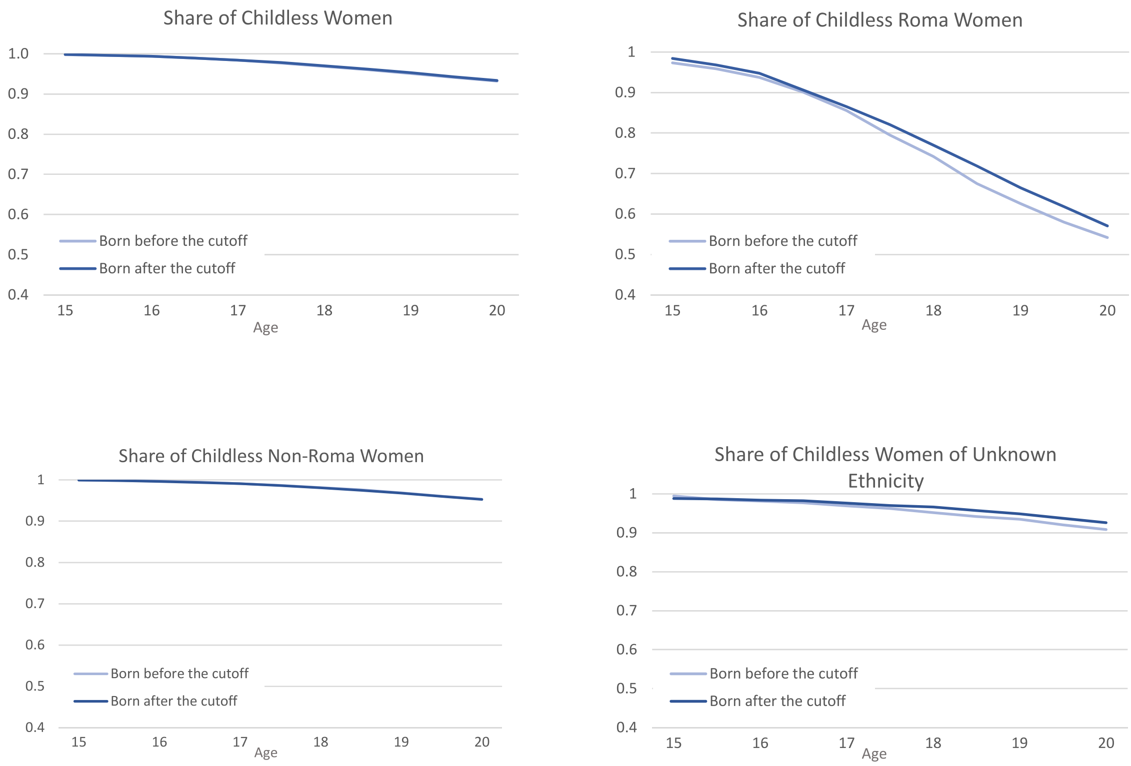
Having the	All women		Roma women		Non-Roma women		Ethnicity unknown	
first child by age ...	Control	Treated	Control	Treated	Control	Treated	Control	Treated
16	0.006	0.006	0.062	0.052	0.003	0.003	0.019	0.016
17	0.016	0.015	0.144	0.135	0.009	0.009	0.030	0.024
18	0.031	0.030	0.259	0.230	0.019	0.019	0.048	0.034
19	0.049	0.047	0.375	0.336	0.031	0.032	0.065	0.051
20	0.068	0.066	0.458	0.428	0.047	0.047	0.092	0.074
No. of obs.	29,275	29,353	1,381	1,394	26,770	26,864	1,124	1,095

Data source: own estimation from the 2011 Hungarian Census. Born before the cutoff (control group): born at most 180 days before June 1, 1991. Born after the cutoff (intention-to-treat group): born at most 180 days after June 1, 1991.

Figure 2.9 shows a similar picture by comparing the share of childless women by age among those born right before and after the cutoff. There is no difference between the share of childless women around the cutoff among non-Roma women. However, such differences are observed among the group of Roma women. The gap in their entry into motherhood begins after reaching age 17, it peaks between ages 18-19, and starts closing by age 20.

10.14754/CEU.2016.11

Figure 2.9: The Share of Childless Women Among Those Born Right Before and Right After the Cutoff (Kaplan-Meier survival functions with respect to the age at first birth-giving)



Share of childless women by age. Age is a continuous variable; giving birth by age 16, for example, means giving birth either before, or exactly on, the mother's 16th birthday. Data source: own estimation from the 2011 Hungarian Census. Born before the cutoff (control group): born at most 180 days before June 1, 1991. Born after the cutoff (intention-to-treat group): born at most 180 days after June 1, 1991. Number of observations: 56,628, 2,775, 53,634 and 2,219, respectively.

2.5 Estimation Results

2.5.1 The Effects of the CSL Age Increase on Teenage Motherhood

Table 2.7 (and Figure 2.10) presents the intention-to-treat (ITT) effects of the higher CSL age on the probability of having the first child by ages 16-20. The CSL age increase has a significant negative effect on the probability of first birth-giving by age 18 among Roma women. Considering that the share of Roma women who give birth by age 18 is at 26.0% in the control group (see Table 2.1), the estimated 6.8 percentage point effect is quite large, 6.8/26=26%. However, this effect is temporary and vanishes by the age 20 observation group.

Table 2.7: Effects on the Probability of First Birth-giving

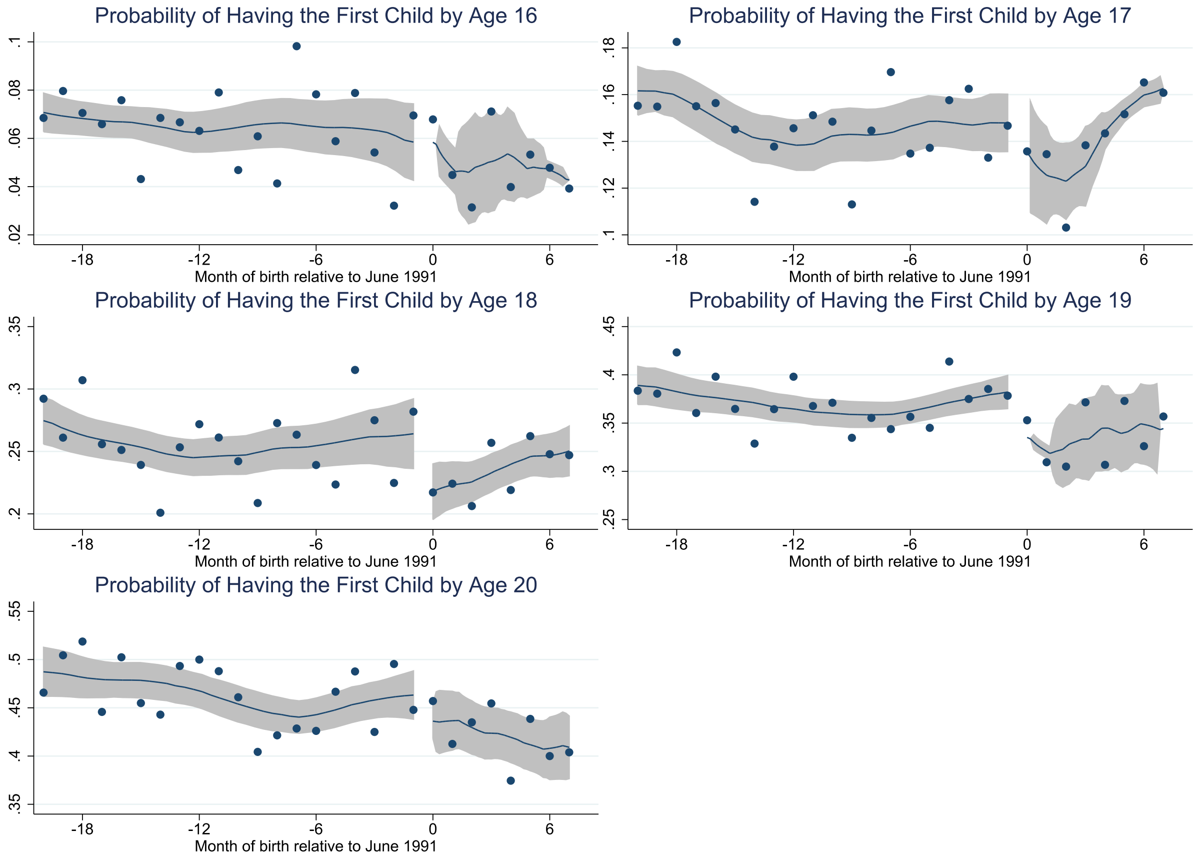
Probability of having the first child by age ...	All women	Roma women	Non-Roma women	Ethnicity Unknown
16	0.001	0.009	0.001	-0.003
	(0.002)	(0.020)	(0.001)	(0.011)
	[50,005]	[2,298]	[51,884]	[2,486]
	{152.2}	{148.2}	{173.6}	{201.3}
17	-0.002	-0.027	0.000	-0.007
	(0.001)	(0.020)	(0.002)	(0.011)
	[60,137]	[3,000]	[53,321]	[2,510]
	{184.3}	{194.8}	{178.4}	{203.0}
18	-0.004	-0.068***	0.000	-0.02**
	(0.002)	(0.022)	(0.003)	(0.008)
	[65,502]	[2,906]	[56,700]	[2,545]
	{201.2}	{188.3}	{190.8}	{206.7}
19	-0.008	-0.067**	-0.001	-0.01
	(0.004)	(0.023)	(0.004)	(0.013)
	[61,070]	[3,015]	[61,000]	[2,100]
	{187.9}	{195.7}	{205.3}	{168.6}
20	-0.003	-0.018	-0.001	-0.044**
	(0.004)	(0.030)	(0.004)	(0.017)
	[51,619]	[2,633]	[44,611]	[2,438]
	{157.2}	{169.8}	{148.0}	{197.7}

Linear probability models estimated by local linear regressions with rectangular kernel bandwidths set using the bandwidth optimization routine of Calonico, Cattaneo and Titiunik (2014). Age is a continuous variable; giving birth by age 16, for example, means giving birth either before, or exactly on, the mother's 16th birthday. Each reported coefficient comes from a separate regression. The coefficients on the treatment dummy variable are reported in the table. Control variables: linear functions of the running variable separately below and above the cutoff. Robust standard errors clustered by birth year and month are in parentheses, observation numbers in brackets, bandwidth measured in days in braces. *** $p<0.01$, **$p<0.05$,*$p<0.1$.

Our results coincide with the phenomenon also seen in Figure 2.9. The gap in motherhood in the treated and control groups of Roma women starts to open after the CSL age of 16, and fades out after age 19, when the CSL age is non-binding. This pattern suggests that the legislation change delayed motherhood by 2 years among Roma adolescents.

Figure 2.10 shows the same results in a visual manner that Table 2.7 showed in numbers. There is one important difference between these two representation of the results: Figure 2.10 plots birth year and month averages, and fits local linear regressions on these averages, while Table 2.7 uses individual level data and standard errors clustered by birth year and month. The two methods produce slightly different standard errors. We work with binary variables on the left hand side, resulting in intra-cluster correlation smaller than 1. Thus, some results are significant on a lower lever in Table 2.7 due to the larger "practical" number of observations.

Figure 2.10: The Effects on the Probability of Motherhood, Roma Women

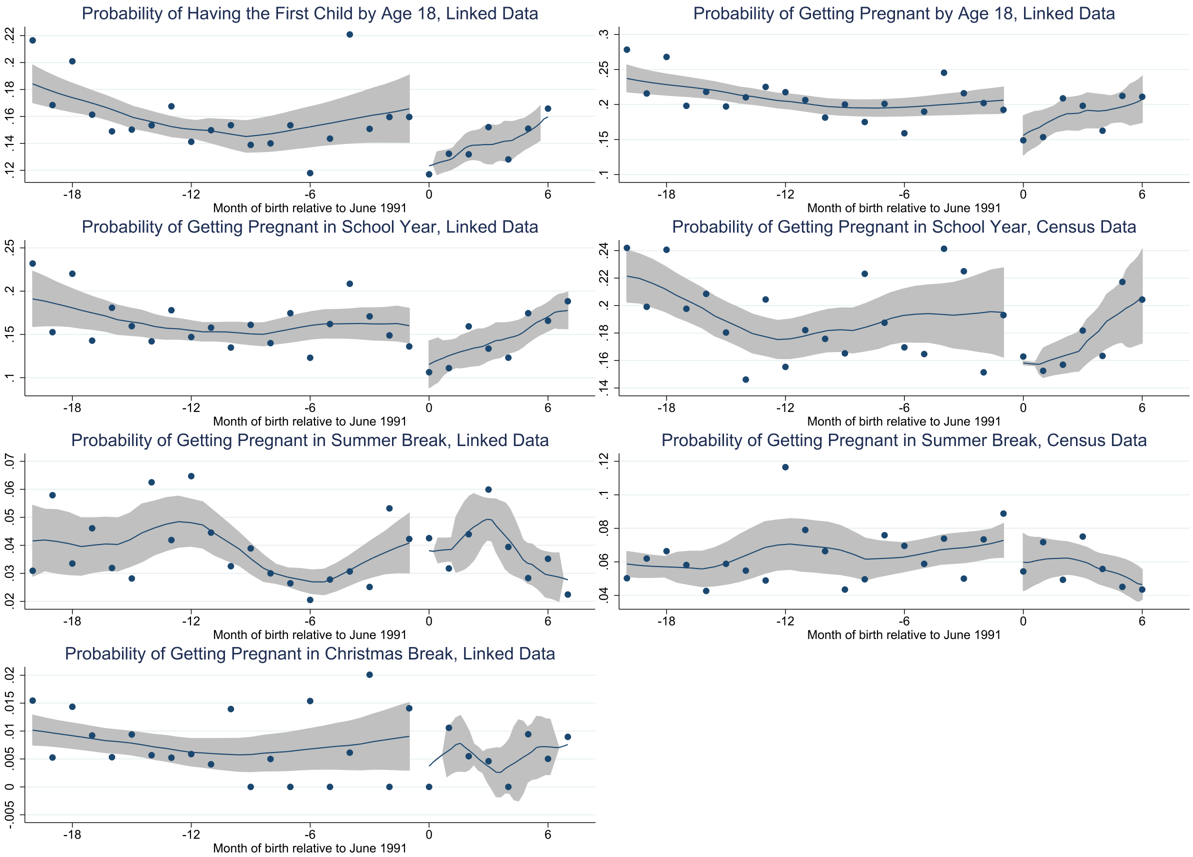


Effects on the probability of giving birth by ages 16-20, Roma women only. Age is a continuous variable; giving birth by age 16, for example, means giving birth either before, or exactly on, the mother's 16th birthday. Local linear regressions fit separately below and above the cutoff on birth year and month averages, using the rule-of-thumb bandwidth of lpoly in Stata with 90% confidence intervals. Month of birth 0 represents June 1991. No. of individual observations: 9,789.

2.5.2 The Mechanics of the Incapacitation Effect

The temporary effects on teenage fertility suggest that the impact among Roma women is due to the incapacitation characteristic of education. If there were any effects through the human capital channel, the impact would not have disappeared by age 20. Our data allows us to examine the conception pattern of these pregnancies as we try to shed light on how such an incapacitation effect is working in practice.

Figure 2.11: The Mechanics of the Incapacitation Effect, Roma Women



Effects on various fertility outcomes of Roma women. Age is a continuous variable; giving birth by age 18 means giving birth either before, or exactly on, the mother's 18th birthday. Local linear regressions fit separately below and above the cutoff on birth year and month averages, using the rule-of-thumb bandwidth of lpoly in Stata with 90% confidence intervals. Month of birth 0 represents June 1991. No. of individual observations: 8,167 in the Census data, 5,785 in the linked data.

As detailed in Section 2.2, we have two methods for identifying the conception time of pregnancies: at monthly precision from the 2011 Hungarian Census data, and at weekly precision on a subsample of the 2011 Hungarian Census data, linked individually to the Vital Statistics database (linked data). Both methods support that the higher CSL age decreased the probability of getting pregnant during school year only, when adolescents had to be physically in school, and that it had no effect on the probability of getting pregnant during summer breaks (Table 2.8 and Figure 2.11). As is Section 5.1, some results are significant on a lower lever in Table 2.8 due to the larger "practical" number of observations.

Results estimated on the linked data show that the probability of getting pregnant decreased as a result of the legislation change (Table 2.8, Block B). We do not see evidence for increased demand for abortions (Block C). In fact, along with less pregnancies, the probability of having an abortion decreased by 2.9 percentage points.

The linked data also suggests that the probability of "unwanted" pregnancies (i.e. those ending with abortion) decreased a little more than the probability of "wanted" (i.e. kept through full term) pregnancies. Although the difference is not significant, the probability of giving birth by age 18 decreased by 5.6 percentage points, and the probability of getting pregnant decreased by 6.5 percentage points. Also, the probability of choosing an abortion to end a pregnancy declined by 8.9 percentage points (Block C), although again, this coefficient is not significant either, and it is only estimated on a small sample (n=442).

Lastly, the linked data offers a possibility to pin the academic year further down to examine the probability of getting pregnant during Christmas breaks separately (Block G). Similarly to summer breaks (Block F), we see small and insignificant coefficients (decrease by 0.8 percentage points). Finding no change in the probability of getting pregnant during school breaks supports the hypothesis that the effect is characterized solely by the incapacitation channel.

Table 2.8: The Mechanics of the Incapacitation Effect (Roma Women)

Outcome	Census data	Linked data
A. Probability of giving birth by age 18	-0.066**	-0.056***
	(0.022)	(0.014)
	[2,788]	[2,327]
B. Probability of getting pregnant	-0.066**	-0.065**
	(0.019)	(0.021)
	[2,788]	[2,327]
C. Probability of abortion	-	-0.029*
		(0.016)
		[2,327]
D. Probability of ending pregnancy with abortion	-	-0.089
		(0.077)
		[442]
E. Probability of getting pregnant during school years	-0.054**	-0.054**
	(0.024)	(0.023)
	[2,788]	[2,327]
F. Probability of getting pregnant during summer breaks	-0.011	-0.003
	(0.015)	(0.009)
	[2,788]	[2,327]
G. Probability of getting pregnant during Christmas breaks	-	-0.008
		(0.007)
		[2,327]

Linear probability models estimated by local linear regressions with rectangular kernel using 180-day bandwidths. Each reported coefficient comes from a separate regression. The coefficients on the intention-to-treat variable are reported in the table. ITT group: born on June 1, 1991, or within 180 days after. Control group: born at most 180 days before June 1, 1991. Control variables: linear functions of the running variable separately below and above the cutoff. Age is a continuous variable; giving birth by age 18 means giving birth either before, or exactly on, the mother's 18th birthday. Robust standard errors clustered by birth year and month are in parentheses, observation numbers in brackets. *** $p<0.01$, **$p<0.05$,*$p<0.1$.

2.6 Robustness Checks

2.6.1 Different Bandwidth Choices

In this subsection we show that our results are not sensitive to bandwidth choice. We do the same local linear estimation as before, using the 50-150% versions of the optimal bandwidth set by the CCT routine. As Table 2.9 shows, our results do not depend on bandwidth choice in magnitude.

Table 2.9: The Effect of the Legislation Change on the Probability of Giving Birth Using Different Bandwidths, Roma Women

	Probability of					
Version of the CCT bandwidth	giving birth by age 18 (Census data)	giving birth by age 18 (Linked data)	getting pregnant by age 18 (Linked data)	getting pregnant during school years (Linked data)	getting pregnant during summer breaks (Linked data)	getting pregnant during Christmas breaks (Linked data)
50%	-0.042	-0.023	-0.051***	-0.001	-0.008	-0.013*
	(0.029)	(0.021)	(0.016)	(0.017)	(0.008)	(0.006)
75%	-0.051*	-0.048***	-0.041*	-0.026	0.006	-0.012
	(0.025)	(0.009)	(0.017)	(0.015)	(0.009)	(0.007)
100%	-0.068***	-0.051***	-0.054**	-0.043**	0.010	-0.007
	(0.022)	(0.012)	(0.018)	(0.017)	(0.009)	(0.006)
	[2,906]	[2,855]	[2,580]	[2,664]	[4,325]	[2,933]
	{188.3}	{221.8}	{191.8}	{205.9}	{329.2}	{226.4}
125%	-0.048**	-0.041**	-0.062***	-0.051**	0.009	-0.008
	(0.022)	(0.014)	(0.018)	(0.019)	(0.009)	(0.006)
150%	-0.050**	-0.034**	-0.045**	-0.038*	0.008	-0.007
	(0.022)	(0.016)	(0.020)	(0.020)	(0.008)	(0.005)

Linear probability models estimated by local linear regressions with rectangular kernel bandwidths set using the bandwidth optimization routine of Calonico, Cattaneo and Titiunik (2014). Each reported coefficient comes from a separate regression. Age is a continuous variable; giving birth by age 18 means giving birth either before, or exactly on, the mother's 18th birthday. Robust standard errors clustered by birth year and month are in parentheses, observation numbers in brackets, bandwidth in days in braces. *** p<0.01, **p<0.05,*p<0.1.

2.6.2 Effect of Cutoffs in 1990-1992

One might be worried that we simply measure the effects of starting elementary school at different ages (a little after age 6 in the control group vs. age 7 in the treated group) instead of measuring the effects of the CSL age increase. To show that this is unlikely to be the case, we replicate our results around the same cutoffs in the year before and after the legislation change. Table 2.10 presents our results. The real, 1991 cutoff is the only one producing significant impacts.

Table 2.10: The Effect of Cutoffs in 1990-1992 (Roma Women)

	Probability of ...					
Cutoff year	giving birth by age 18 (Census data)	giving birth by age 18 (Linked data)	getting pregnant by age 18 (Linked data)	getting pregnant during school years (Linked data)	getting pregnant during summer breaks (Linked data)	getting pregnant during Christmas breaks (Linked data)
1990	0.037	-0.012	0.005	-0.007	0.007	0.005
	(0.029)	(0.013)	(0.018)	(0.022)	(0.011)	(0.003)
1991	-0.066**	-0.056***	-0.065**	-0.054**	-0.003	-0.008
	(0.022)	(0.003)	(0.011)	(0.039)	(0.711)	(0.278)
1992	0.056	0.033	-0.005	0.009	-0.020	0.005
	(0.040)	(0.033)	(0.030)	(0.023)	(0.018)	(0.004)

Linear probability models estimated by local linear regressions with rectangular kernel using 180-day bandwidths. Each reported coefficient comes from a separate regression. The coefficients on the intention-to-treat variable are reported in the table. ITT group: born on June 1, 1991, or within 180 days after. Control group: born at most 180 days before June 1, 1991. Control variables: linear functions of the running variable separately below and above the cutoff. Age is a continuous variable; giving birth by age 18 means giving birth either before, or exactly on, the mother's 18th birthday. Robust standard errors clustered by birth year and month are in parentheses. *** $p<0.01$, **$p<0.05$,*$p<0.1$. Number of observations are around 2,000 in each case.

It might be surprising that we find no effect around the same cutoff in other years. Two things happen around cutoffs in other years: (1) those born in June start school at an older age, and (2) those born in June spend one year less in school before reaching CSL age. Theoretically, starting school at an older age affects teenage fertility in both directions. Black, Devereaux and Salvanes (2011) argue that starting school when older may be beneficial for human capital development because older children are at a more advanced stage of their developmental life. In addition, social development may depend on a child's age relative to the class. If being older than one's peers was beneficial, starting school when older would be beneficial; however, it is not clear whether this is really the case. On the other hand, starting school when older is harmful if children learn more in school than in pre-school (or at home). Furthermore, parental investment in raising their children may depend on school starting age as well. Black, Devereaux and Salvanes (2011) find that starting school at a slightly older age decreases teenage fertility in Norway; however, this effect is not robust across all specifications they use. In the Norwegian school system, schooling is compulsory for 9 years and not until a given age. Thus, not like in the Hungarian case, those starting school later do not spend fewer years in school. In our case, it is likely that the effects of starting school when older and spending fewer years in school before reaching CSL age cancel each other out.

2.6.3 Controlling for Monthly and Yearly Seasonality (Parametric Approach)

Although we find no significant effects around the cutoffs in other years, the estimated coefficients are not significant zeros. One may also worry whether the month of birth is really exogenous (see the dispute about this by Buckles and Hungerman, 2013 vs. Fan et al., 2014). Furthermore, as discussed in Section 2.2.1 and Section 2.4.1, we would like to provide an additional robustness check to our data-linking procedure as well. Therefore, we estimate flexible global polynomial models using a sample of Roma women born in 1980-1993, controlling for these individual characteristics.[12]

Our global polynomial models include the same intention-to-treat variable itti represented as a 1 if individual i was born on June 1, 1991, or after, and as a 0 otherwise, along with a 4th order function of the running variable separately below and above the cutoff, and year of birth, month of birth, day of the week of birth, settlement type, and county fixed effects. If there is a change in the prevalence of teenage pregnancies at any particular year or month of birth, we would be able to control for it with this specification, and would

[12] The choice of the period 1980-1993 is arbitrary. We have tried out several other period lengths and they all give similar results. The only important consideration here is that those born in 1993 are already 18 years old at the time of the Census in October 2011; thus, we can not go further ahead in time. To the other direction, below 1991, we could go as many years as we want. The main consideration was to include a longer period to make room to the estimation of month of birth fixed effects.

estimate the effect that is still remaining. As Table 2.11 shows that the effect of the legislation change in these models is significant and very similar in magnitude to the ones estimated by non-parametric regressions.

Table 2.11: The Effect on the Probability of First Birth-giving (Roma Women, global polynomial models)

	Probability of ...					
	giving birth by age 18 (Census data)	giving birth by age 18 (Linked data)	getting pregnant by age 18 (Linked data)	getting pregnant during school years (Linked data)	getting pregnant during summer breaks (Linked data)	getting pregnant during Christmas breaks (Linked data)
Treatment effect	-0.053***	-0.035**	-0.045**	-0.047***	0.007	-0.004
	(0.019)	(0.015)	(0.018)	(0.015)	(0.011)	(0.005)
No. of obs	35,386	29,521	29,521	29,521	29,521	29,521

Sample: Roma women born in 1980-1993. Intention-to-treat group: born on or after June 1, 1991. Control group: born before June 1, 1991. Control variables: 4th order polynomial function of the running and running*intention-to-treat variable, year of birth fixed effect (FE), month of birth FE, day of the week of birth FE, settlement type FE, county FE. Age is a continuous variable; giving birth by age 18 means giving birth either before, or exactly on, the mother's 18th birthday. Robust standard errors clustered by birth year and month in parenthesis. Significance: *10%, **5%, ***1% .

2.7 The Potential Benefits of Delayed Motherhood

In this section we examine the potential relevance of delaying motherhood from ages 16-18 until ages 18-20. In the Live Statistics database, we observe several health parameters of babies at birth: the week of delivery, their weight, and their APGAR scores. The APGAR score is a compound index of infant health, consisting of five elements: Appearance, Pulse, Grimace, Activity, and Respiration. Its magnitude ranges from 0 to 10, with a higher number indicating better health (Casey et al., 2001).

Table 2.12 compares the health parameters of babies whose mothers were born between 1988-1992. Mothers who gave birth at ages 18-20 have babies who are born after a longer pregnancy and at a slightly higher weight; however, this difference is small.

Table 2.12: Health of Babies at Birth

	All live births in the Live Statistics database		Live births of Roma Women in the linked data	
	Babies whose mothers are age 16-18	Babies whose mothers are age 18-20	Babies whose mothers are age 16-18	Babies whose mothers are age 18-20
Weight	3,024.74	3,073.92	2,981.90	2,997.70
	(5.84)	(4.89)	(11.64)	(11.79)
Week of delivery	38.53	38.67	38.56	38.63
	(0.02)	(0.02)	(0.04)	(0.04)
APGAR score	9.63	9.65	9.66	9.68
	(0.01)	(0.01)	(0.02)	(0.02)
No. of obs.	7,929	12,182	1,662	1,667

Data Source: own estimation from the Vital Statistics database and the 2011 Hungarian Census. Age is a continuous variable. Giving birth at age 16-18 means giving birth between the mothers 16th and 18th birthdays, 18th birthday included. Giving birth at age 18-20 means giving birth between the mothers 18th and 20th birthdays, 20th birthday included. Standard errors of the means are in parentheses.

Comparing the health indicators of babies whose mother was born around the June 1, 1991 cutoff shows a similar picture (see Table 2.13). In fact, there is no health difference among the babies of those mothers born before and after the cutoff. This result is similar to the findings of McCrary and Royer (2011). They examine the effects of longer schooling by exploiting age-at-school-entry policies in the United States. They find that longer education has a generally small, but possibly heterogeneous effect, on infant health. Mothers of the treated and control groups give birth to babies of similar weight and prematurity.

Table 2.13: Health of Babies of Mothers Born Around the Cutoff

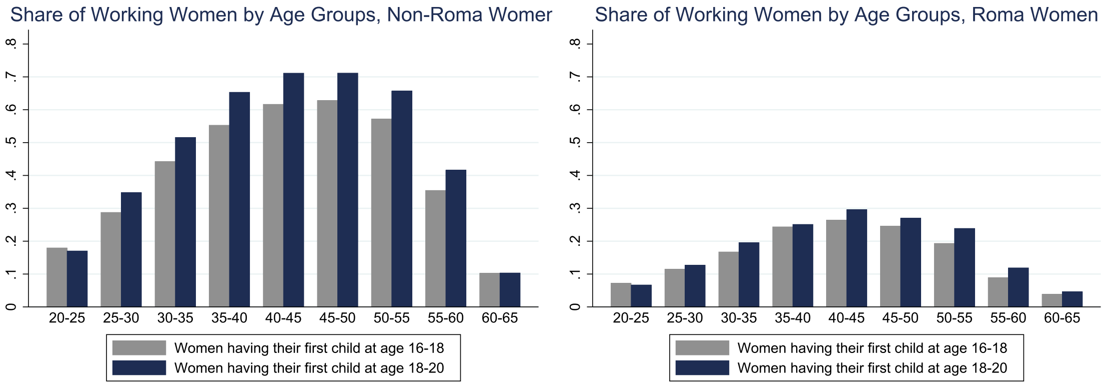
	All live births in the Live Statistics database		Live births of Roma Women in the linked data	
	Babies whose mothers were born before the cutoff	Babies whose mothers were born after the cutoff	Babies whose mothers were born before the cutoff	Babies whose mothers were born after the cutoff
Weight	3061.18	3040.31	3000.6	2993.40
	(11.76)	(11.94)	(24.70)	(27.62)
Week of delivery	38.68	38.60	38.75	38.63
	(0.04)	(0.04)	(0.08)	(0.10)
APGAR score	9.68	9.63	9.66	9.65
	(0.02)	(0.02)	(0.04)	(0.04)
No. of obs.	1,970	1,930	357	317

Data Source: own estimation from the Vital Statistics database and the 2011 Hungarian Census. Born before the cutoff (control group): born at most 180 days before June 1, 1991. Born after the cutoff (intention-to-treat group): born at most 180 days after June 1, 1991. Standard errors of the means are in parentheses.

A better position in the labor market may be the second most important consequence of delayed motherhood. As we observe in the cohort of interest at age 20 in the 2011 Hungarian Census, we cannot tell anything about their long-term labor market outcomes. However, we can compare the labor market outcomes of women having their first child at ages 16-18 with those of women having their first child at ages 18-20. Figure 2.12 shows the share of women ages 20-65 who were working at the time of the 2011 Hungarian Census. On average, non-Roma women are 8 percentage points more likely to work in 2011 if they delayed having their first child from ages 16-18 until ages 18-20. For Roma women, the returns of the same delay in motherhood are 2 percentage points. Although these are not causal effects, they may suggest interesting hypotheses. First, the employment gains of delayed motherhood are much larger for non-Roma women. This finding is in line with the literature on the fertility of ethnic minority women, which argues that teenage birth is less costly for women who suffer from social exclusion. Second, even if the difference in employment gains is smaller, it does exist in the case of Roma women as well. Between ages 40-60, Roma women who have their first child at ages 18-20 are 3 percentage points more likely to work in 2011 than those at ages 16-18. Comparing this to the employment rate of Roma women, which is about 20% in these age groups, this is a 15% increase. Comparing it to the standard deviation of the individual level data, which is 0.40 in the control group of those having their first child between ages 16-18[13], a 3-percentage-point increase in the probability of working is equivalent to an impact of 0.07 standard deviation. For comparison, the average employment effects of active labor market programs in the middle and in the long run are of 0.12-0.20 standard deviation (Card, Kluve and Weber, 2015).

[13]Data Source: own estimation from the 2011 Hungarian Census.

Figure 2.12: Share of Working Women by Age, 2011



Data Source: own estimation from the 2011 Hungarian Census. Age is a continuous variable. Giving birth at age 16-18 means giving birth between the mothers 16th and 18th birthdays, 18th birthday included. Giving birth at age 18-20 means giving birth between the mothers 18th and 20th birthdays, 20th birthday included.

2.8 Discussion

Our analysis finds that the increase in compulsory schooling age from 16 to 18 decreased the probability of having the first child by age 18 among Roma women. Its effect is quite large, at 6.8 percentage points, or a 26% decrease. The impact is temporary and fades out by age 20. Considering the delay only nature of the effect, we formulate a hypothesis that the effect must have been impacted exclusively by the incapacitation channel of education. We find supportive evidence to this hypothesis in the phenomenon that the legislation change decreased the probability of getting pregnant during the school year only, and not during summer and Christmas breaks. We believe that our results have external validity for disadvantaged ethnic minorities living in developed countries. We find no signs of effects through the human capital channel, so we believe that this paper documents the contemporaneous incapacitation effect of being physically in school.

Chapter 3

Integrated Education of Disadvantaged Ethnic Minorities: The Effect of the OOIH Demonstration Program on Roma and Non-Roma Students in Hungary

(Co-authors: Gábor Kézdi and Éva Surányi)

3.1 Introduction

Students from disadvantaged ethnic minorities have lower academic achievement even when they go to schools with better-performing majority students. Segregated education within ethnically mixed schools may be in part responsible for such differences (Card and Rothstein, 2006). While segregation by student characteristics may harm disadvantaged students, fully integrated education is not an obvious choice for schools. Teaching in mixed classes is more difficult than in homogeneous classes, and teachers and school principals may not be prepared for such difficulties.

In this paper we evaluate a program that aimed at transforming the management and teaching practices of schools with a sizable disadvantaged minority to provide high-quality education to all students in a mixed environment. The program was implemented in Hungary, a country with a large disadvantaged Roma minority and high levels of anti-Roma sentiments (Kertesi and Kézdi, 2011; Bernat et al, 2013). The target schools were regular primary schools covering grades 1 through 8, with 20 to 40 percent of Roma students.

The evaluation is non-experimental as the participating schools applied to an open call and were selected by the program administrators. We matched a control school to each of the 30 program schools in the sample, based on pre-program characteristics of student composition and aggregate test scores. We collected survey data on school characteristics, student background, and academic outcomes including standardized tests, socio-emotional skills and inter-group tolerance. The treated and control samples are balanced both in terms of pre-program characteristic and student composition measured by subsequent surveys. Preliminary results of the evaluation were published by Kézdi and Surányi (2009). In this study we take a more systematic approach and address potential problems arising from clustered samples and multiple inference.

We find that the effect of the program was positive overall. The effect is positive on academic development, especially for Roma students, and no negative effects are found on non-Roma students. The effects are positive on socio-emotional skills in both ethnic groups. Anti-Roma sentiments and social distance of non-Roma students are decreased due to the program.

The remainder of the paper is organized the following way. Section 2 provides background to our analysis on the Roma minority, the Hungarian school system and the program. Data is described in Section 3. Section 4 discusses the evaluation method. Section 5 presents the results estimated on school-level and Section 6 on individual-level data. Section 7 concludes.

3.2 Background

3.2.1 The Roma Minority in Hungary

The Roma are one of Europe's largest ethnic minorities. Over 4 million Roma people are estimated to live in Central and Eastern Europe (Bárány, 2002) and around about half million (5–6%) in Hungary (Kemény and Janky, 2006). The Roma are characterized by widespread poverty, low formal employment, low education, poor health and social exclusion (O'Higgins and Ivanov, 2006; Milcher, 2006; FRA-UNDP, 2002). Their low level of education is documented as a major contributor to their low employment and low wages (Kertesi and Kézdi, 2011a). Janky (2004) estimates that two-thirds of the Hungarian Roma belongs to the poorest ten percent of the society.

The Roma are the only significant ethnic minority in Hungary, making up about 6 percent of the population overall and over 10 percent of the population of eighth-grade students (Kemény, 2004; Kertesi and Kézdi, 2015). Most of the Roma of Hungary speak Hungarian and live in neighborhoods that are ethnically mixed, as opposed to segregated settlements (Kemény and Janky, 2006). The vast majority of Roma students complete all eight grades of elementary school in Hungary, although with a substantial achievement gap. While almost all Roma students continue their studies in a secondary school, less than half of them attain a secondary degree in the end (Hajdú, Kertesi, and Kézdi, 2014).

The test score gap between Roma and non-Roma students in grade 8 is estimated to be one standard deviation, similar to the Black-White gap in the early 1980's in the United States (Kertesi and Kézdi, 2011b). Home environment and differential access to schools appear to be the major factors behind this gap, both of which are in turn strongly related to poverty (Kertesi and Kézdi, 2011b). Children in Roma families experience very similar home environments to non-Roma children of the same level of poverty, but they are more likely to study in schools with more disadvantaged students (Kertesi and Kézdi, 2015). The majority of Roma students study in ethnically mixed schools, but there is considerable ethnic segregation across schools (Kertesi and Kézdi, 2013).

3.2.2 The Hungarian School System

Typical students in Hungary go to an 8-grade primary school, followed by a secondary school of 4 years. The secondary school system is composed of three tracks: academic high schools, professional high schools and vocational training schools. Academic and professional high schools end with a maturity exam that is the gateway to higher education. Vocational training schools offer no such exam and are a terminal degree. Admission to secondary schools is based on grades and a centralized placement exam. Labor market prospects differ considerably by the type of secondary school (Kertesi and Varga, 2004; Köllő, 2009), and enrollment in the different types secondary schools is strongly related to test scores in primary school (Hajdú, Kertesi and Kézdi, 2014). Besides academic achievement, ethnicity also plays a role in secondary school enrollment. Compared to non-Roma students with the same test score and grades, Roma students tend to enroll in lower-tier secondary schools (Hajdú, Kertesi and Kézdi, 2014). The most selective academic high schools recruit students in Grade 5 and Grade 7 as well as the usual Grade 9. These schools are not relevant for the student population we study in this paper as they are concentrated in big cities (while most of our schools are in villages and small towns), and they tend to recruit from primary schools with more privileged social background.

School choice is free in Hungary. Primary schools have to admit all students from their administrative catchment area but are free to select from among other applicants. Commuting to schools further away is common, especially in larger towns and cities, and it is strongly positively related to family income. About one half of primary schools have one class per grade. Assignment to classes is fixed in the larger primary schools in Hungary. Students assigned to a particular class stay with that class for most of the day, and class affiliation hardly changes through grade eight.

3.2.3 The OOIH Demonstration Program

The National Educational Integration Network (*Országos Oktatási Integrációs Hálózat, OOIH* in Hungarian) was established by the government of Hungary in 2003, with the aim of promoting the education of disadvantaged and Roma children in elementary schools. The OOIH was established around a demonstration program that is the focus of our paper, and it evolved into an office with a wider mandate. It was re-structured several time since, and the demonstration program stopped slowly and gradually after 2008, without an explicit termination.

The OOIH demonstration program started in 2003, with the participation of 45 regular primary schools. Its stated goals were the following: (1) inclusion and recognition of students regardless their disadvantages, (2) a child and student-friendly environment, (3) new class-management techniques supporting integration in a heterogeneous class, (4) a wide range of activities adjusted to the needs of students, (5) an individualized approach, with an emphasis on individual development, and (6) a focus on partnership and dialog with the students on their individual development needs.

The program provided complex educational, managerial and financial support to the schools and the teachers. The program contained compulsory and optional elements. Compulsory elements included equalizing the ethnic distribution of classes in the same grade (half of the schools had more than one class per grade). Each school was assigned a "coordinator" person, a young Roma who helped the schools with compliance and reported to the OOIH office. Teachers went through intensive training programs with frequent follow-ups, and schools had to choose the type of training program for their teachers (cooperative learning, project methods, differentiated education, etc.). Extra financial help was provided to make the physical environment student- and cooperation-friendly. For further details on the program see Kézdi and Surányi (2009).

The program was implemented gradually within each school, starting with students in grades 1 and 5 and their teachers in 2003/4. Subsequent cohorts entered the program in subsequent years.

Eligible schools had to have a significant Roma minority in the range of 20 to 40 percent. Participation in the program was voluntary. Three was an open call, and participants were chosen by the OOIH staff based on their assessment of the likelihood of success, without assigning a score or other quantitative indicator. Participants in the program were the result of self-selection, and the selection likely favored schools that would have achieved better outcomes without the program. At the same time, many schools applied with no previous experience in teaching aiming explicitly to get extra resources.

Although color-blind in its stated goals, some of the elements of the program aimed at strengthening Roma identity and increasing inter-ethnic tolerance. Perhaps because of no other significant minority, integrated education in Hungary is usually understood as mixing Roma and non-Roma students.

There was a major administrative change in the program in 2004. Of the 45 schools 34 continued to participate, and some additional new schools joined the program. We randomly selected 30 out of the 34 schools for our evaluation study because the original evaluation plan included resources to collect data from exactly 30+30 schools.

3.3 Data

The paper uses three main data sources. Pre-program school characteristics used for the matching strategy are taken from the Public Education Statistics (PES) of the Public Education Information System and the National Assessment of Basic Competencies (NABC) database. Individual and school-level data on the treated and the control schools during and after the program are collected by Kézdi and Surányi (2009), specifically for the purpose of this evaluation.

The PES of the Public Education Information System[1] is the official Hungarian school census database. It collects extensive information on schools, school programs, and students. We use data on school size, location and ethnic minority education from the PES, from 1999 and 2003.

The NABC[2] administrative database registers the results of centrally organized low stake math and reading tests taken each year in Grades 6, 8, and 10. We use data on family background and achievement score results from the 2003 spring wave of the NABC, covering six graders.

The data that were collected for the purpose of this evaluation covers 30 treated and 30 control schools. As the original evaluation plan included 30-30 school only, we choose a random sample of 30 schools from the 34 that had participated in both the first and the second year of the program. To each program school we matched a control school based on pre-program data, geographic proximity and expert evaluation. Section 4 provides more detail on the matching procedure.

[1] *KIR-STAT* in Hungarian.

[2] In English: http://edecon.mtakti.hu/?q=node/15, in Hungarian: http://www.oktatas.hu/kozneveles/meresek/kompetenciameres/alt_leiras

In this evaluation study we focus on upper-grade students: children who were in their fifth grade in 2003/4. They entered eighth grade in the fall of 2006 and graduated from primary school in 2007. We evaluate the effect of the program on their outcomes, skills and attitudes at the end of their primary school career, in the spring of 2007.

Half of our sample consists of schools with one class per grade, while the other half has two or more classes per grade. In the few schools with more than two classes per grade we selected the two classes with the highest and lowest student achievement, as classified by the school principals. We followed all students within the selected classes.

Due to administrative reasons, the first wave of data collection took place in the spring of 2005. By this time the sampled cohort was about to complete sixth grade, one and a half years into the program. Our first data collection was therefore unable to collect pre-program data. Altogether, we collected data in three waves, in the springs of 2005, 2006 and 2007, when sampled students were in their sixth, seventh and eighth grade.

The first survey (2005) measured baseline outcome variables and school characteristics. The second survey (2006) measured student background. We carried out a survey on teaching practices and administered classroom observations in the same year, but the results of those are not used in this analysis. The last survey (2007) measured outcomes, including a standardized reading test, tests on socio-emotional skills, inter-ethnic attitudes. We also obtained data on admission to secondary school, and we measured the ethnicity of the students. See Kézdi and Surányi (2009) for more details on the surveys, and Kézdi and Surányi (2010) on the details of the socio-emotional measures.

We estimate the effect of the program using outcome variables measured in the last year of the program and collected by two surveys: a reading test and a survey measuring personality characteristics and inter-ethnic tolerance. A general concern with integrated education is that it may harm students from the ethnic majority. To address this concern we estimate effects on Roma and non-Roma students separately besides overall average treatment effects.

3.3.1 Measuring Ethnicity

Similarly to other countries of continental Europe, administrative data on students (or any other kind) do not contain ethnic markers. Therefore, we collected data on the ethnic identity of students in a survey, fully complying with the relevant regulations (Act LXIII of 1992 on personal data protection and the recommendations of the data protection ombudsperson). We asked schoolteachers to send a letter with a one-page survey to the parents of all students in the sample after all other surveys were completed, in the spring of 2007. The letter asked them to provide a statement declaring the national and ethnic belonging of the student, and explicitly asked for their consent for the information to be used in the research in an anonymous way. Importantly, the survey allowed for multiple ethnicity. The parents had an opportunity to provide the following answers: my child is Hungarian, non-Roma; Hungarian, Roma; Hungarian, partially Roma; not Hungarian, Roma; Romanian; Slovak; Other.

Twenty-seven percent of the statements were not returned, primarily because the school-teacher refused to send them to the parents. At the same time, for the classes with enough observations, the fraction or Roma students based on this measure was practically the same as the fraction of Roma students indicated by the schoolteacher.

3.3.2 Outcome Measures

Three types of outcome measures were collected in or after 8th grade: academic outcomes, socio-emotional skills, and measures of inter-ethnic relations.

Our most important measure of academic outcomes is a reading test. The results of developmental psychology and education research suggest that the OOIH program might have had a small-scale impact on the students' reading comprehension development at most. Most results in the literature imply that schools, in general, have little influence on students' cognitive processes (Cole, 1990). To a large extent, children's cognitive skills develop well before the beginning of schooling age, because the development of the brain largely finishes by age six (DeBord, 1997; Shore, 1997). It appears that when school education has a large impact on the cognitive performance, the impact is generally realized through information processing strategies and through the shaping of some specific skills or abilities (Cole, 1990).

The fact that underlying cognitive skills are hard to manipulate in school does not mean that the schools have no role in children's skills development in general, and the development of their reading skills in particular. Children normally learn to read at school. A different school can induce different patterns of development in literacy even if underlying cognitive skills are the same. Nevertheless, there is ample evidence for the inability of school-age education to fully make up for deficiencies in early childhood development. Therefore, we can expect limited success at most from the OOIH program in improving students' reading skills. Reading tests were developed following the principles of the PISA test by Kézdi and Surányi (2009).

The other academic outcome measure is enrollment in secondary school after completing primary school. In particular, we evaluate whether students enrolled in higher-tier secondary schools that offer maturity examination (these are the academic and professional high schools). The results of secondary school admissions were collected from the school administrations after the end of the school year.

The second type of outcome measures are related to socio-emotional skills. As documented in labor economics, such skills may be as important in shaping labor market success as academic outcomes (Bowles at al., 2001; Cole, 1990). The most referred socio-emotional skills include self esteem, the locus of control (the belief in one's ability to manage one's own destiny), and coping (the ability to cope with difficult situations). There is evidence that the importance of some non-cognitive skills have significantly increased in the past few decades because of technological change, with the replacement of routine cognitive tasks by the computer (Autor et al., 2003). Contrary to cognitive skills, non-cognitive skills develop during the school years, especially during adolescence (Carneiro and Heckman, 2003). Thus, we can have higher expectations about the impact of the OOIH program on the non-cognitive skills of the students.

We measure locus of control with the help of a four-item questionnaire, a subset of the 13-item test developed by Julian B. Rotter (Rotter, 1996). See Kézdi and Surányi (2010) for the selection procedure that maximized internal consistency reliability. The test is standardized on a national representative sample, together with the rest of the non-cognitive and inter-ethnic tests.

In modern psychology, "self-esteem" reflects a person's overall evaluation or appraisal of his or her own worth. Self-esteem encompasses beliefs (for example, "I am competent/ incompetent") and emotions (for example, triumph/despair, pride/shame)." Besides emotional processes self-esteem involves cognitive and behavioral aspects (Blascovich and Tomaka, 1991). The modern concept of self-esteem originated with the work of Morris Rosenberg (Rosenberg, 1965). The most widely used measure is the Rosenberg Self-Esteem Scale. It is a one-dimensional measure of the comprehensive self-esteem. It contains ten statements that formulate the opinion related oneself, and the respondent simply has to decide whether he/she agrees with the statements. We use a four-item short version of the Rosenberg Self-Esteem Scale. See Kézdi and Surányi (2010) for more details.

A number of other tests have also been developed for self-esteem, among those several specifically for children as well as for adolescents (Blascovich and Tomaka, 1991). We adopted the Harter Self Perception Profile for Children (SPPC) test elaborated by Susan Harter for teenagers (Harter, 1985). The SPPC is widely used for impact assessment studies of programs targeting school-age children. It has sub-scales for various domains of self-esteem, allowing for a more differentiated analysis than the measures focusing on a single component (Harter, 1983). The scale consists of 36 items. Some of its items look at children's judgments related to their specific skills, while other items measure the extent to which they consider themselves to be valuable in general. We use five different sub-scales, four of which measure specific areas: academic competence, social acceptance, external appearance, and behavioral discipline. The fifth scale measures general self-esteem. See Kézdi and Surányi (2010) for details of the shortened version of the SPPC use in the evaluation study.

The effect of the program is also evaluated on the ability of students to cope with stressful and conflict situations. This ability is called "coping" in the psychology literature. Good coping ability enables people to go on without any significant negative consequence even if the root cause of the problem remains (Folkman and Lazarus, 1984). We use a four-item measure formulated in accordance with the coping literature. It is not an adaptation of an existing test.

Our third set of measures aims at inter-ethnic relations. One of the most important goals of integrated education of different ethnic groups is reducing stereotypes, prejudice and social distance. In his classic book, Allport (1954) argued that reduced physical distance, in other words, more contact between otherwise hostile groups, can reduce social distance and prejudice under certain conditions (Allport, 1954). These conditions include equal status of the two groups, inter-group cooperation, common goals, equality of power, and some law or practice that emphasizes equality and supports cooperation. A large literature has emerged since with the aim of analyzing the correlation between inter-group contact and prejudices (Pettigrew, 2006). Despite the nearly half a century of interest among researchers, no agreement has been reached in the subject. Some argue that inter-group contacts result in a significant decrease in prejudice (Jackson, 1993; Pettigrew, 1971), while others argue for a weak impact at most (Amir, 1976; Ford, 1986).

Often-examined forms of distance keeping and conflict are stereotypes, prejudice and discriminating behavior itself. Stereotypes comprise the cognitive component of the attitude towards a group. Prejudice is usually viewed as the affective or emotional aspects of intergroup contact. Measuring behavior is the third direction of the research. Discriminating behavior is usually viewed as the result of stereotypes and prejudice in conflict situations. While the three aspects originate from a common theoretical stem, research fields on the different aspects are often isolated from each other. In this evaluation study, we analyze all three dimensions. Integrated education increases inter-group contact, and some elements of the OOIH program can be understood as instruments for strengthening the conditions of contact to reduce social distance as conceptualized by Allport. Besides measures on stereotypes and prejudices, we examine social anxiety and social dominance attitudes that are apparently connected with the development of prejudices.

As described above in Section 3.3.1, ethnicity was measured separately from the other variables, due to Hungarian data protection regulations. All questions related to stereotypes, prejudice and social distance were asked as opinions about both the Roma and non-Roma ethnic group. The respondents' ethnicity was merged to these data afterwards, allowing for identifying opinions about the other ethnic group.

A stereotype is usually described as a simplistic, exaggerated and overgeneralized judgment made about the members of a social group. In this study stereotypes are measured using a bipolar classification of characteristics expressing relative judgments. The questionnaire presented five characteristics in a series of opposite statements. In each case the respondent had to decide which statement of the opposing pair of statements described the Roma or the non-Roma the best way, measured on a scale of 1-10. For example, they had to express their opinion regarding to which extent the Roma or the non-Roma have all the abilities to perform well at school, etc.

Prejudice is defined as the affective aspects of inter-group relations. Measures of ethnic prejudices try to map a large range of emotions related to a given ethnic group. We use an adapted version of the Bogardus social distance scale which is a standard tool in the research of inter-ethnic relations (Bogardus, 1928). The scale measures the social distance that the individual wishes to keep from the members of a given social group, at various levels of intimacy. The questions of this adapted version touch upon marriage, friendship and relationships with classmates and neighbors.

Social dominance orientation (SDO) is often thought as an important factor of stereotypes and prejudice. SDO summarizes the extent to which individuals accept or believe in social hierarchy and ethnic inequality (Sidanius and Pratto, 1999). A strong SDO means that the individual thinks that some groups are more valuable than others, the hierarchy among social groups is unavoidable and even desirable, and social dominance is a necessary fact of life. We developed a six-item scale adapting adult tests to adolescents. The items include questions like Do you think some people are more valuable than others?, or, Do you think winning is more important than the way we play games?, etc.

Social anxiety is an important indicator of an individual's behavior in social (interpersonal) relationships, and, as such, it is important in the development of children's later career (Leary and Meadows, 1991). Social anxiety may be an important mediation in inter-group conflicts, and it may be interesting in itself as well. Besides the fear of aliens or outsiders, little research has been done in the subject of children's social anxiety. The important exception is the social anxiety test developed by La Greca and colleagues (La Greca et al., 1988). We narrowed the original 10-item test to a shorter 5-item version, containing three case-relevant dimensions: fear of a negative judgment, fear of interaction, and retreat.

3.3.3 Controlling for Social Desirability

Measuring the attitudes towards oneself and others by self-administered questionnaires may be problematic if students project attitudes that are positively biased in order to make a good impression. The question here is not simply the extent of the bias but whether it is different in program versus control schools. The tendency of people to present themselves in a more positive way in their statements than in reality is often considered the result of social desirability. Survey researchers have addressed such biases since the 1950s (see Reynolds, 1982).

We have adopted the Children's Social Desirability Scale (CSDS) of Crandall et al. (1965). The CSDS describes children's everyday actions (e.g. "Sometimes I don't like to share things with my friends", "I never shout when I am angry"). It has been used in a number of studies that targeted elementary and secondary schoolchildren. The reduced, five question version of the CSDS test was developed by recording of the original 48 question test jointly with the Rosenberg test on positive self-esteem (see later) and the questions of the RCMAS (Revised Children's Manifest Anxiety Scale) lie detector test (the latter is a 37 item self-filled questionnaire that measures the children's manifest anxiety and is used primarily for clinical and diagnostic purposes). The selected items have been named as a „making a good impression" test. Among the questions there are statements such as: "I never say things that may hurt someone's feelings." and "I always behave respectfully with older people." See Kézdi and Surányi (2010) for more details.

3.4 Evaluation Framework

3.4.1 Matching Control Schools

This is a non-experimental evaluation that aims to estimate the Average Treatment Effect on the Treated (ATET) of the program. Identification relies on the unconfoundedness assumption, i.e., the assumption that after conditioning on a set of control variables, students in program and control schools would have achieved the same results without the program. We use a matched sample of schools based on pre-program school characteristics. In the analysis we further condition on a set of variables on student background and school characteristics that were pre-determined to program participation.

We matched control schools in a three-step procedure. First, we estimated a propensity score on all primary schools of Hungary based on the following pre-program school characteristics: geographical location, ethnic composition, school size, ethnic minority education, family background and achievement score results of the students. In the second step a 1:1 nearest neighbor matching was conducted for each program schools of the same region and town size; this was therefore exact matching on geographics and propensity score matching on other characteristics. In the third step expert opinion was asked to verify the selection (Gábor Havas sociologist with extensive local knowledge of Hungarian elementary schools). In 6 of the 30 cases the expert suggested that the matched control school was not a good choice (e.g., because of errors in some of our measures or changes in school composition). In those cases we replaced the nearest neighbor match with the second nearest neighbor. The second nearest neighbor was selected to be the control school in two additional cases where the first nearest neighbor declined participation in the study.

Table 3.1: Covariate Balance of Pre-Program School Characteristics Used for Matching (1999, 2003)

	National average	Treated schools	Control schools	Balance statistics
Population in town/village of school	225,992	87,363	83,399	0.008
No of students in school	286	320	311	0.041
Fraction of students eligible for Roma minority support in 1999	0.05	0.31	0.26	0.155
Fraction of students at risk	0.09	0.16	0.14	0.103
Fraction of students with				
Mother's education less than 8 grades	0.02	0.09	0.07	0.151
Mother's education exactly 8 grades	0.21	0.35	0.33	0.112
Father's education less than 8 grades	0.01	0.03	0.05	-0.159
Father's education exactly 8 grades	0.14	0.26	0.25	0.047
No working parents	0.20	0.29	0.30	-0.069
Number of books at home 0 to 50	0.12	0.29	0.26	0.129
Number of books at home about 50	0.12	0.16	0.15	0.072
Competence scores at school level, Grade 6, Spring 2003				
Mathematics: school average	489	456	446	0.163
Reading: school average	488	449	436	0.198
Mathematics: school standard deviation	87	88	88	-0.003
Reading: school standard deviation	89	92	95	-0.111
No. of obs (schools)		30	30	60

Source: Kézdi and Surányi, 2009, and own estimation from the data. Balance statistics are calculated based on Imbens and Wooldridge (2009, Section 5, page 19.).

Table 3.1 shows the covariate balance of the matched sample, together with national averages for comparison. The last column shows the balance statistic proposed by Imbens and Wooldridge (2009, Section 5, page 19.):

$$\Delta x = \frac{\bar{X_1} - \bar{X_2}}{\sqrt{VAR_1 + VAR_2}},$$

where $\bar{X_1}$ and $\bar{X_2}$ are treated and control group means, and VAR_1 and VAR_2 are treated and control group variances. The advantage of this balance statistic with respect to the usual t-statistic is its independence from sample size. As a rule of thumb, Imbens and Wooldridge (2009) suggest that a value above 0.25 indicates that linear regression methods are likely to be sensitive to model specification. All of our statistics are below 0.25. Some characteristics suggest that program schools have a slightly more disadvantaged student body but the differences are negligible.

Note that our sample is not representative from a national point of view. The program and control schools in our sample are located in smaller settlements, they scored lower on average on the national competence tests both in mathematics and reading comprehension. Students in our sample are significantly more likely to be Roma and come from disadvantaged families. This is a natural consequence of the program's focus.

3.4.2 Selection into Program and Control Groups

Assignment into the program was not random. School applications implied self selection while the goal of the decision making mechanism of OOIH was to choose explicitly the potentially most successful applicants. Thus, we can only estimate the ATET of the program.

We are worried about the presence of unobserved characteristics affecting both program assignments and student outcomes. If there is a bias from unobservables, it is likely to point upwards. For example, schools that were more motivated towards integrated education or those that already had experience with integrated education were probably more likely to apply to participate in the program, and it is also probable that those schools could have been able to provide higher quality learning environment independently from program participation.

Table 3.2: Covariate Balance, Individual Characteristics

	Treated	Control	Balance statistics
Individual characteristics in Grade 6			
Fraction of girls	0.496	0.506	-0.014
Age	12.849	12.963	-0.126
Fraction of Roma students	0.270	0.253	0.027
Fraction of students eligible to child protection subsidy	0.442	0.529	-0.123
Fraction of students eligible to free schoolbooks	0.545	0.655	-0.160
Fraction of disadvantaged students	0.387	0.414	-0.039
Fraction of students at risk	0.057	0.087	-0.082
Fraction of students eligible to free lunch	0.465	0.521	-0.080
Average No. of siblings	4.821	4.148	0.028
Fraction of students with			
No working parent	0.156	0.173	-0.033
Mother's education less than 8 grades	0.057	0.057	0.001
Mother's education exactly 8 grades	0.283	0.267	0.025
Father's education less than 8 grades	0.020	0.030	-0.048
Father's education exactly 8 grades	0.229	0.214	0.026
No. of obs	608	565	1,173

Source: Kézdi and Surányi, 2009, and own estimation from the data. The data of students with complete data from both Grade 6 and Grade 8. Balance statistics are calculated based on Imbens and Wooldridge (2009, Section 5, page 19.).

Table 3.3: Sample Size and Response Rates in Grades 6 and 8

	Treated	Control
No. of surveyed students in Grade 6	1,099	1,081
No. of responses in Grade 6 - cognitive test score	905	934
No. of responses in Grade 6 - family background	982	1,028
No. of surveyed students in Grade 8	865	927
No. of responses in Grade 8 - survey on non-cognitive outcomes	816	810
No. of responses in Grade 8 - survey on cognitive outcomes	825	876
No. of responses - ethnicity	559	583
No. of students with all responses from both Grade 6 and Grade 8 (except for ethnicity)	608	565

Source: Kézdi and Surányi, 2009, and own estimation from the data.

As detailed in Section 3.3, the first phase of data collection took place two years into the program, at the end of Grade 6. The collected individual-level data supports that student composition and survey response rates are very similar in treated and control schools (see Table 3.2 and 3.3). However, treated and control schools may differ with respect to teacher quality. We measure teacher quality from two data sources. First, Kézdi and Surányi (2009) interviewed 195 teachers in 59 schools in 2005. They collected data on the teaching career and education background of their interviewees. Table 3.4 compares interviewed teachers in treated and control schools. The average share of teachers teaching with no appropriate certification is very similar in the two groups. However, the share of teachers holding an MA degree is much higher among treated schools.

Table 3.4: Covariate Balance, Teacher Characteristics from the Interview Data (2005)

	All schools				Schools with no experience with integrated education prior the program			
	Treated	Control	Diff	Balance Statistics	Treated	Control	Diff	Balance Statistics
Share of teachers with no appropriate degree, Grade 1-4	0.05	0.03	0.02	0.071	0.04	0.04	0.00	0.000
Share of teachers with no appropriate degree, Grade 5-8	0.09	0.09	-0.00	-0.007	0.04	0.05	-0.01	-0.041
Share of teachers with MA degree	0.12	0.02	0.10	0.480	0.08	0.02	0.07	0.443
Fraction female	0.92	0.89	0.03	0.138	0.94	0.89	0.05	0.226
Average age	43.72	41.81	1.90	0.258	43.83	42.90	0.93	0.125
Average tenure (years since in present school)	18.14	16.31	1.84	0.214	18.50	17.35	1.15	0.118
Share of teachers with a degree in special education of ethnicities	0.30	0.12	0.18	0.567	0.27	0.12	0.15	0.411
Share of teachers participating in teachers training	0.96	0.91	0.05	0.236	0.94	0.93	0.01	0.064
Average share of Roma students in class	0.44	0.44	-0.00	-0.003	0.44	0.41	0.03	0.095
Average No. of students in class	20.95	21.05	-0.10	-0.018	21.62	21.27	0.34	0.062
No of schools	30	30	60		14	14	28	

Data source: collected by Kézdi and Surányi (2009) through interviews with teachers in 2005. Balance statistics are calculated based on Imbens and Wooldridge (2009, Section 5, page 19.).

The second data source to measure teacher quality is the National Assessment of Basic Competencies (NABC) database. The NABC data measure the share and number of teachers teaching either with no teaching degree in general, or teaching a subject (i.e. math) without having a degree to teach that subject in particular. The first NABC wave we to be used is the one taken in Spring 2006, which is the end of the third year of the program evaluated in this paper. Table 3.5 shows that treated and control schools are very similar to each other in these measures on average.

Table 3.5: Covariate Balance, Teacher Characteristics from the NABC Data (2006)

	All schools				Schools with no experience with integrated education prior the program			
	Treated	Control	Diff	Balance statistics	Treated	Control	Diff	Balance Statistics
Whether teachers with no teachers degree teach in school	0.21	0.24	-0.03	-0.045	0.33	0.21	0.12	0.183
No. of teachers teaching with no teachers degree, per 100 students	0.15	0.24	-0.09	-0.129	0.30	0.26	0.05	0.064
Whether teachers teaching a subject with no degree for that particular subject	0.36	0.36	0.00	0.000	0.33	0.38	-0.05	-0.073
No. of teachers teaching a subject with no degree for that particular subject , per 100 students	0.39	0.53	-0.14	-0.118	0.21	0.40	-0.19	-0.274
No of schools	28 (30)	29 (30)	57 (60)		12 (14)	14		26

Data source: own estimation from the school (telephely) level data of the National Assessment of Basic Competencies (NABC) database taken in 2006. This is a last-minute table and we could link 57 schools into their NABC data, but this will be corrected to 60. Balance statistics are calculated based on Imbens and Wooldridge (2009, Section 5, page 19.).

The data on teacher quality show more teachers with an MA degree in treated schools, and besides, there is a possibility of unobserved heterogeneity as well. We apply the following robustness checks to control the selection problem. First, in addition to the matching procedure, we control for all observable pre-program school-level characteristics in our regressions that could signal teaching quality differences. Pre-program test results in students achievements may partially control for imbalances in teacher quality.

Second, Kézdi and Surányi (2009) collected data on whether the schools had already had experience with integrated education before entering the program. This is true for 14 schools out of the 30. In addition to that, two years after it was launched, 2 control schools joined the program. We defined these two types of schools and their matched pairs as always-takers and we control for this behavior in our regressions.

Third, we control for grade-6 test results when we estimate the effect of the program on grade-8 outcomes of the students in a differences-in-differences fashion in all sets of regressions. This method allows us to control for non-observable differences among the treated and control schools as long as we believe that these unobservable differences are constant in time. The baseline results of the students were not measured; the earliest outcomes available come from Grade 6. Using Grade-6 outcomes instead of baseline results is not straightforward as they may reflect the impact of the program already causing a bad-control problem (Angrist and Pischke, 2009). However, we will show that controlling for second year test results does not change the magnitude of the coefficients; thus, any potential bias coming from bad control problem must be small.

Fourth, we apply a clustered version of the Rosenbaum bounds matching sensitivity analysis test in Section 5.2. The test evaluates how stable our results are in the presence of an unobservable variable that affects both the assignment to the program and the outcome variables in pre-set correlation structures.

3.4.3 Multiple Outcomes and Hypothesis Testing

We evaluate the causal effect of the program on cognitive and non-cognitive skills and inter-ethnic attitudes of the students. Testing the effect of a program on several outcome variables raises concerns about multiple inference (Anderson, 2008). Considering a set of statistical inferences simultaneously, the probability of a type-I. error is higher, or equivalently, hypothesis tests that incorrectly reject the null are more likely to occur, than originally intended by a single test at a time.

There are two solutions to the multiple inference problem and they both have their pros and cons. Multiple-test procedures correct downwards the originally intended p-value criteria for statistical significance taking into account the number of hypotheses tested. They calculate a new corrected overall critical p-value, which has the feature that an individual null hypothesis is considered to be acceptable if and only if its corresponding p-value is greater than its corrected critical p-value. However, these procedures lower the power of the tests by each additional hypothesis, hence increase the probability of type-II errors or false non-rejections.

The second approach is to lower the number of hypotheses tested by aggregating single outcome variables into summary indexes. This is beneficial because it lowers the number of simultaneous tests. However, it may reduce the scope of the analysis: we can only test, for example, whether the program had any effect, or whether it affected cognitive outcomes in general, but not reading skills in particular. Thus, the trade-off is between gaining test power and sacrificing the specificity of the analysis. Based on Anderson (2008) we apply the combination of these two methods to solve the multiple inference problem. In this section, we reduce the number of tests, i.e., the number of outcome variables, to one. We construct one summary index from the single outcome variables and test the general effect of the program. Then, in Section 6.2 we break down the general program effect and test it separately on the cognitive, non-cognitive and tolerance outcomes of the students. We aggregate the original outcome variables to three summary indexes: to a summary index of cognitive, non-cognitive and tolerance measures. We also adjust our hypothesis tests according to the number of tests performed using two types of multiple test procedures. This strategy allows us to unfold the fact that the general positive effect of the program comes from its positive impact on non-cognitive skills and tolerance measures, rather then on the cognitive skills of the students.

The procedure, in which we reduce the number of hypothesis tests to be performed by constructing one summary index from all outcome variables, has several advantages. First, it ensures that including an additional outcome measure does not increase the number of tests, thus the probability of false rejection, while it absorbs new information from all possible sources. Second, it provides a statistical test whether the program has a general effect on a set of outcomes. Third, by giving more weight for new information, it reduces random errors in the outcome measures (Anderson, 2008). See the exact primary outcomes in Table C.5 in Appendix C.

The summary index is constructed based on Anderson (2008) as the weighted mean of standardized primary outcome variables, where the weights are set to maximize the information captured by the index by giving higher weights for new, uncorrelated information. Formally, we are constructing an index from K outcome variables. Let's denote the primary outcome variables by y_k. The summary index is referred as $\bar{s}_i$, which is the weighted average of the primary outcome values y_{ki} of individual i. In general,

$$\bar{s}_i = (\mathbf{1}'\hat{\Sigma}^{-1}\mathbf{1})^{-1}(\mathbf{1}'\hat{\Sigma}^{-1}\tilde{y}_i),$$

where $\mathbf{1}$ is a column vector of ones, $\hat{\Sigma}^{-1}$ is the inverted covariance matrix of the primary outcome variables with each other, and $\tilde{y}_i$ is a column vector of all primary outcomes included in the summary index for individual i. In other words, this is an efficient generalized least square (GLS) estimates of the mean of outcomes. The vector of weights (i.e., the inverted covariance matrix) we constructed from our data is found in Appendix C.

Practically, the weight of every primary outcome variable is proportionate to the sum of its row in the inverted variance-covariance matrix. Thus, the more correlated a variable is with the rest of the variables, the lower weight it gets. Therefore, the procedure gives higher weights for new information and reduces measurement errors. In addition to that, the procedure ignores missing values when it estimates the mean: if at least one of the weighted variables has a non-missing value for individual i, it calculates an index value $\bar{s}_i$; therefore, it uses all available information. However, it gives higher weights for variables with less missing values (Anderson, 2008).

3.4.4 Estimation

As the treatment itself was given on school level, we start with estimating its effect on school-average achievements of the students in the sample. We estimate the following linear models by OLS:

$$index_s = \alpha + \beta * program_s + \delta X_s + u_s, \; s = 1...60,$$

where $index_s$ stands for school-level averages of the summary index constructed as detailed above, $program_s$ is a binary variable indicating whether school s participated in the program or not, and X_s is a matrix of school characteristics and school-level averages of student characteristics.

After the school-level estimates, we turn to individual-level estimation that allows to estimate the heterogeneous effects of the program with respect to ethnicity. We estimate the following linear models again by OLS:

$$index_i = \alpha + \beta * program_i + \delta X_i + u_i,$$

where $index_i$ stands for the summary index of individual i constructed as detailed above, $program_i$ is binary variable indicating whether individual i attended a program school or not, and X_i is a matrix of individual student characteristics. In addition to individual characteristics, we control for school characteristics as well.

3.5 School-Level Estimates of Overall Effects

Table 3.6 shows the overall, school-level effects of the program on our summary index constructed from cognitive, non-cognitive and inter-ethnic attitude measures. To address potential problems with small sample inference we present test results based on two simulation-based test procedures in Appendix C.

All models are estimated using linear regressions. The summary index measured in Grade 8 is the dependent variable, and a treatment dummy is on the right hand side. In Model 1 we use no additional control variables besides the program variable, while in Model 2 we control for several school-level characteristics, geographical information, Grade 6 test results, individual characteristics and the share of Roma students in school as well.

The program has a significantly positive effect on school performance. The magnitude of this effect is about 7.5-9.0% of control group standard deviation. Adding individual and school-level control variables or Grade 6 test results to the model does not change the magnitude of the estimated coefficients.

Table 3.6: School-Level Estimates of Overall Effects

	Model 1	Model 2
Program	0.090	0.075
Robust SE	0.029	0.035
P-values	0.003	0.044
School-level controls	NO	YES
Control for always takers	NO	YES
Control for school-average of individual characteristics	NO	YES
Control for school-average social desirability	NO	YES
Share of Roma students - school average	NO	YES
No. of obs	60	58

Linear regressions. Dependent variable: summary index of cognitive, non-cognitive and inter-ethnic attitude measures, measured in Grade 8.

3.5.1 Sensitivity Test for Unobserved Characteristics

This sections conducts a matching sensitivity analysis to support the validity of our identification strategy. Sensitivity tests are used to examine the robustness of matching-based results in the presence of unobserved heterogeneity. In particular, we use the clustered version of the Rosenbaum bounds test (Hansen et al, 2014). In our identification framework, i.e. under conditional independence, if two schools have the same observed characteristics their probability of entering the program is equal: if $X_i = X_j$, $p_i = p_j$. However, in the presence of such unobserved characteristics that affect both the probability of getting treatment and students' skills at the same time, we may face a selection bias and even if $X_i = X_j$, and $p_i \neq p_j$. A sensitivity test approaches to answer the following question: how strongly should an unobserved characteristic influence the assignment mechanism in order to take away the significance of the estimated program effect? Let's define a Γ sensitivity parameter such that

$$\frac{1}{1+\Gamma} \leq Pr(w_{s1} = 1) \leq \frac{\Gamma}{1+\Gamma}, \text{ and}$$

$$w_{s0} = 1 - w_{s1}.$$

If $\Gamma = 1$ there is no selection bias, $Pr(w_{s1} = 1) = 1/2$ and $\frac{p_j}{p_i} = 1$. The bigger is Γ, the higher is the odds ratio $\frac{p_j}{p_i}$ of the assignment probabilities of a treated and a control school pair with the same X's. For example, if $\Gamma = 2$, treated school j is two times as likely to participate in the program then control school i even though $X_i = X_j$.

The test derived by Hansen et al. (2014) does the following procedure. It defines a test statistics T as the weighted sum of the differences of the ranks of students with respect to the outcome measure within each treated and control pair of schools. Using ranks instead of the direct values of outcome measures is advantageous as it is resistant to outlier figures. In the case of random assignment,

$$Pr(w_{s1} = 1) = Pr(w_{s0} = 1) = 1/2, \text{ and}$$

$$T = \sum_{s=1}^{S} w_s Z_{s1}\left(\frac{1}{n_{s1}} \sum_{i=1}^{n_{s1}} q_{s1i} - \frac{1}{n_{s2}} \sum_{i=1}^{n_{s2}} q_{s2i}\right) + w_s Z_{s2}\left(\frac{1}{n_{s2}} \sum_{i=1}^{n_{s2}} q_{s2i} - \frac{1}{n_{s1}} \sum_{i=1}^{n_{s1}} q_{s1i}\right) =$$

$$= \sum_{s=1}^{S} w_s (2Z_{s1} - 1)\left(\frac{1}{n_{s1}} \sum_{i=1}^{n_{s1}} q_{s1i} - \frac{1}{n_{s2}} \sum_{i=1}^{n_{s2}} q_{s2i}\right) =$$

$$= \sum_{s=1}^{S} B_s Q_s, \text{ where}$$

$B_s = 2Z_{s1} - 1 = +/-1$, and

$Q_s = \frac{w_s}{n_{s1}} \sum_{i=1}^{n_{s1}} q_{s1i} - \frac{w_s}{n_{s2}} \sum_{i=1}^{n_{s2}} q_{s2i}$.

In the randomization case under the null hypothesis of no treatment effect, T is a sum of S independent random variables taking values $+/- Q_s$ with probability $1/2$, thus $E(T) = 0$ and $Var(T) = \sum_{s=1}^{S} Q_s^2$. If $S \to \infty$, the central limit theorem implies that $\frac{T}{\sqrt[2]{var(T)}}$ converges in distribution to the standard Normal distribution, $\Phi()$.

In the case of a selection bias the uneven odds of assignment , i.e. Γ, is incorporated into the test statistic the following way. Let $\theta = \frac{1}{1+\Gamma}$ and define $\pi_s = \theta$ if $Q_s > 0$ and $\pi_s = 1 - theta$ otherwise.

Then, for each pre-set, theoretical value of $\Gamma = 1...n$, recalculates T to give the lower and upper bound of treatment probability. Then it assumes that the distribution of T and its bounds converges to a normal distribution and estimates the lower and upper bounds of p-values. As long as the upper bound of p-values from a two-tailed test stays under 0.05, results are insensitive to an omitted variable affecting the probability of treatment by a factor of Γ.

Table 3.7: Clustered Rosenbaum Bounds Matching Sensitivity Test

Gamma	Upper Bound
1.0	.0051572
1.1	.00986059
1.2	.01692729
1.3	.02673975
1.4	.03956087
1.5	.05553146
1.6	.07468042
1.7	.09694171
1.8	.12217389
1.9	.15017928
2.0	.18072131

Clustered Rosenbaum bounds matching sensitivity test by Hansen et al. (2010). Number of pairs: 30. Number of schools: 60.

Table 3.7 shows that the estimated program effect is insensitive to an omitted variable up to a $\Gamma < 1.4$. For $\Gamma = 1.4$ or over, the upper bound of p-value from a two-tailed test rises above 0.05. Thus, the matching procedure and the estimated effects seem to be insensitive to small biases but they are sensitive to moderately large biases.

3.6 Individual Estimates: Effects on Roma and Non-Roma Students

In this section we estimate the effects of the program on individual-level data. Table 3.8 supports our earlier results estimated on school-level data (see Table 3.6). It shows significantly positive program effects at 11-13% standard deviation.

Table 3.8: Individual-Level Estimates of Overall Effects

	Model 1	Model 2
Program	0.113	0.125
Clustered robust SE	0.033	0.028
P-values	0.001	0.000
Diff-in-diffs setup	NO	YES
School-level controls	NO	YES
Control for non-compliance	NO	YES
Control for individual characteristics	NO	YES
Control for social desirability	NO	YES
Control for ethnicity	NO	YES
Control for teacher quality in Grade 5-8	NO	YES
No. of obs	1,173	1,173

Linear regressions. Dependent variable: summary index of cognitive, non-cognitive and inter-ethnic attitude measures, measured in Grade 8.

3.6.1 Academic Achievement, Socio-emotional Skills and Anti-Roma Sentiments

In this section we try to identify what types of skills are the most affected by the program: cognitive, non-cognitive or inter-ethnic relations skills. We construct three summary indexes: an index of cognitive measures, an index of non-cognitive measures, and an index of inter-ethnic relations. Then, we test the effect of the program on these three outcome indexes at a time by controlling for the number hypotheses tested. The summary indexes are constructed just as before. We are constructing $J = 3$ indexes from K_j ($j = 1...3$) outcome variables in case of each j index. Let's denote the primary outcome variables by y_{kj} indicating that all y_k variables are assigned to exactly one of the J summary indexes. For example, in the case of our first ($J = 1$) index, the cognitive outcome index, $K_1 = 2$, because it incorporates two primary outcome variables. The summary index of cognitive skills is referred as $\bar{s}_{1i}$, which is the weighted average of the primary outcome values y_{11i} and y_{21i} of individual i. Then,

$$\bar{s}_{ji} = (\mathbf{1}'\hat{\Sigma}_j^{-1}\mathbf{1})^{-1}(\mathbf{1}'\hat{\Sigma}_j^{-1}\tilde{y}_{ji}),$$

where $\mathbf{1}$ is a column vector of ones, $\hat{\Sigma}_j^{-1}$ is the inverted covariance matrix of the primary outcome variables with each other, and $\tilde{y}_{ji}$ is a column vector of all primary outcomes included in summary index j for individual i. The three matrices of weights are found in Appendix C.

To conduct multiple inference with respect to the three constructed indexes we use two multiple test procedures: we control for the Familywise Error Rate (FWER) and the False Discovery Rate (FDR). FWER control is a more conservative procedure and it is to be used when committing type-I. errors may be extremely costly; FDR control tests have more power thus they are more applicable for exploratory analysis and in case of testing a large number of hypothesis. In our case of three hypotheses they both lead to the similar conclusions.

Let's assume that 3 hypotheses, H_1, H_2,...,H_M are tested, out of which J are true ($J \leq M$). FWER is the probability that at least one of the J true hypotheses is falsely rejected, and it increases with M. FWER control techniques adjust the p-values of the individual tests upwards (or the critical p-value for statistical inference downwards). We use the Holm-Bonferroni step-down resampling method by Holm (1979). This procedure is more powerful than the most popular Bonferroni correction which simply multiples p-values by the number of hypothesis tested. [3] Also, the Holm-Bonferroni method does not need assumptions like other, more powerful procedures do.

The procedure goes as follows. Let's denote $p_1, p_2, ..., p_M$ the p-values corresponding to the M hypotheses above, and α the significance level. The first step is to rank the p-values starting with the smallest. Let k be the place of each p-values in the new order. Then, investigate every single hypothesis in the rank with respect to the following individually corrected p-value:

$$p_k^* = \tfrac{\alpha}{M+1-k}.$$

In the case of three hypothesis and $\alpha = 0.05$, the smallest individual p-value will be investigated with respect to $0.05/3$, the second smallest to $0.05/2$, etc. Investigation stops when the first hypothesis in the rank is left non-rejected. For example, if the smallest p-value (thus the first hypothesis in the rank) is not smaller than $0.05/3$, no hypothesis can be rejected. This method ensures that FWER$\leq \alpha$.

While FWER limits the probability of committing any Type-I errors, FDR controls the expected share of rejections that are Type-I errors. We use the Benjamini and Hochberg (1995) FDR correction method as follows. Again, we rank the p-values $p_1, p_2, ..., p_M$ corresponding to the M hypotheses in an increasing order, but this time we investigate them one-by-one starting with the highest (thus, this is a step-up procedure). Let τ be the place of each p-values in the new order. Then, we are looking for the largest τ for which $p_\tau < \frac{\alpha\tau}{M}$. In case of three hypotheses and $\alpha = 0.05$, the largest p-value in the rank will be investigated against $p^* = 0.05 * 3/3$, the second one against $p^* = 0.05 * 2/3$, and the first one against $p^* = 0.05 * 1/3$. When we find the largest p-value that meets the applicable criteria, we reject the corresponding hypothesis and all the rest with smaller p-values. The procedure controls FDR at level α for independent or positively dependent p-values (Anderson, 2008).

Results are shown in Table 3.9 and 3.10 for Model 2, and in Table C.6 and C.7 of Appendix C.2 for Model 1 as before.[4] The effect of the program is significant on a 5% level in the case of the indexes of non-cognitive outcomes and tolerance measures with both p-value correction procedures in Model 2. However, in Model 1 the impact on the cognitive skills of the students is not significant on conventional levels.

[3]For more on this please see Anderson (2008), Section 3.2.2

[4]The regressions of Table 3.9on school-level data are in Table C.9 in Appendix C.2.

Table 3.9: Effects on Academic Achievement, Socio-emotional Skills and Anti-Roma Sentiments

	Index of cognitive outcomes	Index of non-cognitive outcomes	Index of tolerance measures
	Model 2		
Program	0.147	0.125	0.143
Cl robust SE	0.043	0.039	0.049
Uncorrected p-values	0.001	0.002	0.005
FDR-control test rejects on...	5%	5%	5%
FWER-control test rejects on...	5%	5%	5%
Diff-in-diffs setup	YES	YES	YES
School-level controls	YES	YES	YES
Control for non-compliance	YES	YES	YES
Control for individual characteristics	YES	YES	YES
Control for social desirability	YES	YES	YES
Control for ethnicity	YES	YES	YES
Control for teacher quality in Grade 5-8	YES	YES	YES
No. of obs	1173	1173	1173

Linear regressions. Dependent variables: the summary index of cognitive outcomes in the first column, the summary index of non-cognitive index in the second column and the summary index of inter-ethnic attitudes in the third columns, measured in Grade 8. The same results from Model 1 are reported in Table C.6 in Appendix C. FDR and FWER-control multiple testing procedures are used to test 12 hypotheses together in Tables 3.9, C.6, 3.10, and C.7.

3.6.2 Effects on Roma and Non-Roma Students

In this section we examine whether the impact of the program differs on Roma and non-Roma students. We estimate the following models:

$$index_i = \alpha + \beta_1 * program_i + \beta_2 * Roma_i + \beta_3 * non - Roma_i * program_i + \delta X_i + u_i,$$

where the interaction term $non - Roma_i * program_i$ captures the heterogeneity of the program effect with respect to ethnicity. As we are interested in whether majority, non-Roma students could have been harmed by integrated education, we defined the interaction term with an emphasis on the non-Roma. Coefficient β_1 captures the effect of the program on Roma students, $\beta_1 + \beta_3$ captures the effect of the program on non-Roma students.

Table 3.10 shows that the interaction term is not significant in the case of the cognitive and non-cognitive outcomes. Even if these coefficients were significant, the effect of the program would still have been positive for both Roma (0.202 and 0.139) and non-Roma students (0.108 and 0.120). The program, however, increased the inter-ethnic tolerance of non-Roma students significantly more (-0.022 vs. 0.200).

Table 3.10: Effects on Roma and Non-Roma Students

	Index of cognitive outcomes	Index of non-cognitive outcomes	Index of tolerance measures
	Model 2		
Program (effect on the Roma)	0.202	0.139	-0.022
Cl robust SE	0.098	0.066	0.081
P-values	0.044	0.040	0.783
Roma	-0.235	-0.086	0.499
Cl robust SE	0.104	0.062	0.067
P-values	0.028	0.103	0.000
Program*non-Roma interaction	-0.074	-0.019	0.222
Cl robust SE	0.112	0.079	0.086
Uncorrected p-values	0.512	0.807	0.013
FDR-control test rejects on...	-	-	10%
FWER-control test rejects on...	-	-	-
Effect on the non-Roma	0.108	0.120	0.200
Diff-in-diffs setup	YES	YES	YES
School-level controls	YES	YES	YES
Control for non-compliance	YES	YES	YES
Control for individual characteristics	YES	YES	YES
Control for social desirability	YES	YES	YES
Control for ethnicity	YES	YES	YES
Control for teacher quality in Grade 5-8	YES	YES	YES
No. of obs	1173	1173	1173

Linear regressions. Dependent variables: the summary index of cognitive outcomes in the first column, the summary index of non-cognitive index in the second column and the summary index of inter-ethnic attitudes in the third columns, measured in Grade 8. The same results from Model 1 are reported in Table C.7 in Appendix C. FDR and FWER-control multiple testing procedures are used to test 12 hypotheses together in Tables 3.9, C.6, 3.10 and C.7.

In Table 3.10 we find that the effect of the program on the index of cognitive skills is not significantly, but lower on non-Roma students. In Table 3.11 we show that the difference of the impact between Roma and non-Roma students is also not significant for the two factors of the index of cognitive skills: reading scores and secondary school acceptance. Again, even if the coefficients of the interaction terms were significant, the effect of the program would still have been positive for both Roma (0.173 and 0.155) and non-Roma (0.083 and 0.091) students.

Table 3.11: Effects on Roma and Non-Roma Students - Reading Scores and Secondary School Acceptance

	Reading score	Attending academic high school
	Model 2	
Program (effect on the Roma)	0.173	0.155
Cl robust SE	0.131	0.062
P-values	0.195	0.016
Roma	-0.337	-0.136
Cl robust SE	0.128	0.073
P-values	0.011	0.063
Program*non-Roma interaction	-0.090	-0.064
Cl robust SE	0.142	0.078
P-values	0.529	0.417
Effect on the non-Roma	0.083	0.091
Diff-in-diffs setup	YES	YES
School-level controls	YES	YES
Control for non-compliance	YES	YES
Control for individual characteristics	YES	YES
Control for social desirability	YES	YES
Control for ethnicity	YES	YES
Control for teacher quality in Grade 5-8	YES	YES
No. of obs.	1140	954

Linear regressions. Dependent variables: standardized reading test score in Grade 8 in the first column, and the probability of starting an academic high school degree in the second column. The same results from Model 1 are reported in Table C.8 in Appendix C.

3.7 Conclusions

We examine the effects of an education program offering integrated, student- and cooperation-friendly learning environment for Roma and non-Roma elementary school students in Hungary. We use a non-experimental method. We find that the program had a positive general effect captured by a composite index of cognitive, non-cognitive, and inter-ethnic measures. Most of this positive effect was realized on the non-cognitive outcomes and inter-ethnic attitudes of the students. The program had a positive effect on the outcomes of both Roma and non-Roma students.

In an ideal world, evidence on the impacts of integrated education of ethnic minority and majority students would come from randomized controlled trials. To our best knowledge, no similar programs were evaluated in experimental setting. Our identification strategy relies on the assumption that having a rich dataset of observable school- and individual-level characteristics, no unobservables affected both program participation and school outcomes. We know that treated schools were selected into the program; thus, we aim at estimating ATET. To support our results, we have completed several robustness checks. Using a Rosenbaum bounds sensitivity analysis, we show that our identification strategy is not sensitive to unobservable characteristics affecting the assignment mechanism up to a small-medium level. However, our results are sensitive to a potential unobservable force affecting program participation at a moderate or high level.

The data collected by Kézdi and Surányi (2009) enables us to control for interim test results in a semi diff-in-diffs fashion. Controlling for time-invariant unobserved differences among treated and control schools does not change our results. Even if the observed positive phenomenon is not 100 percent due to the program, we can still conclude that the appropriate sensitive approach in integrated education is beneficial for all parties involved.

Chapter 4

References

4.1 References for Chapter 1

Acemoglu, D. and Angrist, J. D., 2001. "How large are human capital externalities? Evidence from compulsory schooling laws." NBER Macroannual Pages 9–59.

Adato, M. and Hoddinott, J. (ed.), 2010. "Conditional Cash Transfers in Latin America." IFPRI books, International Food legislation Research Institute (IFPRI), No. 978-0-8018-9498-5.

Anderson, M. L, 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." Journal of the American Statistical Association, 103(484):1481_1495, 2008.

Anderson, D. M., Hansen B. and Walker, M. B., 2013. "The minimum dropout age and student victimization." Economics of Education Review, Volume 35, August 2013, Pages 66-74.

Angrist, J. D., 1990. "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records." American Economic Review, American Economic Association, vol. 80(3), pages 313-36, June 1990.

Angrist, J. D. and Krueger, A. B., 1991. "Does Compulsory Schooling Attendance Affect Schooling and Earnings?" Quarterly Journal of Economics, 106(1), Pages 979-1014.

Angrist, J. D. and Pischke, J., 2009. "Mostly Harmless Econometrics: An Empiricist's Companion" Princeton University Press.

Benjamini, Y. and Hochberg, Y, 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." Journal of the Royal Statistical Society. Series B (Methodological) Vol. 57, No. 1 (1995), pp. 289-300.

Bjorklund, A. and Jantti, M., 1997. "Intergenerational Income Mobility in Sweden Compared to the United States." American Economic Review, American Economic Association, vol. 87(5), December 1997, Pages 1009-18.

Black, S., Devereux, P. and Salvanes, K. G., 2008. "Staying in the Classroom and out of the Maternity Ward? The Effect of Compulsory Schooling Laws on Teenage Births." Economic Journal, 118(530): 1025-54.

Black, S., Devereux, P. and Salvanes, K. G., 2011. "Too Young to Leave the Nest? The Effects of School Starting Age." The Review of Economics and Statistics, MIT Press, vol. 93(2), pages 455-467, May.

Buckles, K. S. and Hungerman, D. M., 2013. "Season of Birth and Later Outcomes: Old Questions, New Answers." The Review of Economics and Statistics, MIT Press, vol. 95(3), pages 711-724, July 2013.

Büttner, B. and Thomsen, S. L., 2010. "Are we spending too many years in school? Causal evidence of the impact of shortening secondary school duration" ZEW Discussion Papers 10-011, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Cabus, S. J. and De Witte, K., 2011. "Does school time matter?—On the impact of compulsory education age on school dropout." Economics of Education Review, Elsevier, vol. 30(6), Pages 1384-1398.

Calonico, S., Cattaneo M.,D. and Titiunik, R., 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." Econometrica Volume 82, Issue 6, November 2014., Pages 2295–2326.

Calonico, S., Cattaneo, M.,D. and Titiunik, R., 2014a. "Robust Data-Driven Inference in the Regression-Discontinuity Design." Stata Journal, 14, Number 4, Pages 909–946.

Cheng, M.-Y., Fan, J. and Marron, J. S,. 1997. "On automatic boundary corrections." Ann. Statist. 25 (1997), no. 4, 1691--1708.

Cygan-Rehm, K. and Maeder, M., 2013. "The effect of education on fertility: Evidence from a compulsory schooling reform." Labour Economics Volume 25, December 2013, Pages 35–48.

d'Addio, A. C., 2007. "Intergenerational Transmission of Disadvantage: Mobility or Immobility across Generations? A Review of the Evidence for OECD Countries." Technical Report 7, OECD Social, Employment and Migration Working Papers.

Devereux, P. and Hart, R., 2010. "Forced to be Rich? Returns to Compulsory Schooling in Britain". Economic Journal, 120(549), pp. 1345–1364.

Dobos, V., 2014. "Magasabb korhatás - több érettségi?" MA Thesis, ELTECON.

Education Office of the Ministry of Human Resources, 2014. "Országos Jelentés. A magántanulóvá nyilvánítással kapcsolatos eljárás szakmai ellenőrzése." Oktatási Hivatal Tanügy-igazgatási és Ellenőrzési Osztály. Budapest, October 18, 2014.

Fan, J. and Gijbels, I., 1996. "Local polynomial modelling and its applications." Chapman & Hall

Fan, E., Liu, J-T. and Chen, J. C., 2014. "Is the 'Quarter of Birth' Endogenous? Evidence From One Million Siblings in Taiwan." National Bureau of Economic Research Working Paper Series, No. 20444, August 2014.

Freeman, J. and Simonsen, B., 2015. "Examining the Impact of legislation and Practice Interventions on High School Dropout and School Completion Rates A Systematic Review of the Literature" Review of Educational Research June 2015 vol. 85 no. 2, Pages 205-248.

Green, C. and Navarro Paniagua, M., 2012. "Does Reducing the School Leaving Age Reduce Teacher Effort? Evidence from a legislation Experiment." Economic Inquiry, Volume 50, Issue 4, pp 1018–1030. September, 2012.

Grenet, J., 2013. "Is Extending Compulsory Schooling Alone Enough to Raise Earnings? Evidence from French and British Compulsory Schooling Laws." Scandinavian Journal of Economics, Wiley Blackwell, vol. 115(1), 01, Pages 176-210.

Gulesci, S. and Meyersson, E., 2013. ""For the love or the Republic." Education, Secularism, and Empowerment." Working Papers 490, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.

Hablicsek, L., 2007. "Kísérleti számítások a roma lakosság területi jellemzőinek alakulására és 2021-ig történő előrebecslésére." Demográfia, L. évf. 2007/1: 5–54.

Hahn, J., Todd, P., and Van Der Klaauw, W., 2001. "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design." Econometrica, 69(1), Pages 201-209.

Hajdú, T., Hermann, Z., Horn, D., Kertesi, G., Kézdi, G. and Varga, J, 2015. "Az érettségi védelmében" Budapest Working Papers On The Labour Market, The Hungarian Academy of Sciences. 2015/1. Pages 1-27.

Hámori, Sz. and Köllő, J., 2011. "Kinek használ az évvesztés?" Közgazdasági Szemle, LVIII. évf., 2011. február, Pages 133–157.

Harmon, C., and Walker, I., 1995. "Estimates of the Economic Return to Schooling for the United Kingdom." American Economic Review, vol. 85(5), pages 1278-86, December.

Haveman, R. and Wolfe, B., 1995. "The Determinants of Children's Attainments: A Review of Methods and Findings." Journal of Economic Literature 33 (4), Pages1829–1878.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N. and Verashchagina, A., 2007. "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends." The B.E. Journal of Economics Analysis & legislation 7 (2), Article 10.

Imbens, G, W. and Kalyanaraman, K., 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator" Review of Economic Studies, Oxford University Press, vol. 79(3), 2012, Pages 933-959.

Imbens, G. W. and Lemieux, T., 2008. "Regression discontinuity designs: A guide to practice." Journal of Econometrics, Elsevier, vol. 142(2), February 2008, Pages 615-635.

Kézdi, G. and Surányi, É., 2008. "Egy sikeres iskolai integrációs program tapasztalatai" Educatio Társadalmi Szolgáltató Közhasznú Társaság Budapest, 2008.

Kézdi, G., Köllő J., and J. Varga, 2009. "The Failures of 'Uncertified' Vocational Training", In: Fazekas, K; Köllő, J. (eds.): Munkaerőpiaci Tükör 2008. Budapest: Magyar Tudományos Akadémia, Közgazdaságtudományi Intézet, pp. 87-136.

Landis, R. N., and Reschly, A. L., 2011. "An examination of compulsory school attendance ages and high school dropout and completion." Educational Policy, 25, 719– 761.

Lee, D. S. and Lemieux, T., 2010. "Regression Discontinuity Designs in Economics" Journal of Economic Literature 48 (June 2010), Pages 281–355.

Lleras-Muney, A., 2005. "The Relationship between Education and Adult Mortality in the United States," Review of Economic Studies, 21(1), Pages 189-221.

Lochner, L. and Moretti, E., 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." American Economic Review Vol. 94, No. 1, March 2004, Pages 155-189.

Mártonfi, Gy. (ed.), 2011a. "Hány évig tartson a tankötelezettség?" Oktatáskutató és Fejlesztő Intézet, Budapest, 2011, Pages 1-12.

Mártonfi, Gy. (ed.), 2011b. “A 18 éves korra emelt tankötelezettség teljesülése és (mellék)hatásai.” Oktatáskutató és Fejlesztő Intézet, Budapest, 2011, Pages 1-47.

McCrary, J. and Royer, H., 2011. “The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth." American Economic Review, 101(1), Pages 158-95.

Meghir, C. and Palme, M. 2005. "Educational Reform, Ability, and Family Background." American Economic Review, 95(1): 414-424.

Meghir, C., Rivkin, S., 2011. “Econometric Methods for Research in Education”, In: Eric A. Hanushek, Stephen Machin and Ludger Woessmann, Editor(s), Handbook of the Economics of Education, Elsevier, 2011, Volume 3, Pages 1-87.

Milligan, K., Moretti, E. and Oreopoulos, P., 2004. “Does education improve citizenship? Evidence from the United States and the United Kingdom.” Journal of Public Economics 88 (2004), Pages 1667 – 1695.

Ministry of Human Resources, 2013. “Statistical Yearbook of Education, 2012/2013”. Ministry of Human Resources, Section of Statistics. Budapest, 2013.

National Institute for Public Education, 2006. “Report on the Hungarian Public Education (Jelentés a magyar közoktatásról, in Hungarian.” Országos Közoktatási Intézet, Budapest, 2006.

National Institute for Public Education, 2010 “Report on the Hungarian Public Education (Jelentés a magyar közoktatásról, in Hungarian).” Országos Közoktatási Intézet, Budapest, 2010.

OECD, 2014. “PISA 2012 Results in Focus. What 15-year-olds know and what they can do with what they know.” OECD, 2014. Pages 1-41.

OECD, 2015. “Education Policy Outlook: Hungary. OECD, November 2015 Pages 1-27.

Oosterbeek, H. and Van Ophem, H., 2000. “Schooling choices: Preferences, discount rates, and rates of return.” Empirical Economics, 2000, 25, Pages 15-24.

Oosterbeek, H. and Webbink D., 2007. “Wage effects of an extra year of basic vocational education.” Economics of Education Review Volume 26, Issue 4, August 2007, Pages 408–419.

Oreopolous, P., 2006. “Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter” American Economic Review, 96(1), Pages 152-175.

Oreopolous, P., 2007. “Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling.” Journal of Public Economics, Volume 91, Issues 11–12, December 2007, Pages 2213–2229.

Pischke, J-S. and Wachter, T. von, 2008. “Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation.” The Review of Economics and Statistics, MIT Press, vol. 90(3), August, Pages 592-598.

Stephens, M. and Yang, D-Y., 2014. “Compulsory Education and the Benefits of Schooling.” American Economic Review Vol. 104, No. 6, June 2014, Pages 1777-92.

4.2 References for Chapter 2

Alam, Andaleeb, Javier Baez and Ximena V. Del Carpio, 2011. “Does cash for school influence Young Women’s behavior in the Longer Term? Evidence from Pakistan.”, IZA Working Paper 5703.

Angrist, J. D., Krueger, A. B., 1991. “Does Compulsory Schooling Attendance Affect Schooling and Earnings?” Quarterly Journal of Economics, 106(1), Pages 979-1014.

Bean, Frank D., and Swicegood Gray, 1985. "The Idea of Opportunity Costs and Minority/Majority Fertility Differences." In Mexican American Fertility Patterns, 34-51. University of Texas Press, 1985.

Béki, O., 2013: “A serdülıkori gyermekvállalás lehetséges okai egy cigánytelepen.” Belvedere Meridionale. XXV. 3. 117–129. pp

Berthelon, M. E., and Kruger, D. I., 2011. “Risky behavior among youth: Incapacitation effects of school on adolescent motherhood and crime in Chile.” Journal of Public Economics, 95(1), 41-53.

Bifulco, R., Lopoo, L. M., and Oh, S. J., 2015. “School Desegregation and Teenage Fertility”. Educational Evaluation and Policy Analysis December 2015 37: 591-611.

Black, S. E., Devereux, P. J., and Salvanes, K. G., 2008. “Staying in the Classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births.” The Economic Journal, 118(530), 1025-1054.

Blum RW, Beuhring T, Shew ML, Bearinger LH, Sieving RE, Resnick MD, 2000. “The effects of race/ethnicity, income, and family structure on adolescent risk behaviors.” Am J Public Health. 2000 Dec; 90(12):1879-84.

Bonell, C., 2004. “Why is teenage pregnancy conceptualized as a social problem? A review of quantitative research from the USA and UK. Culture, health and sexuality.”, 6(3), 255-272.

Breierova, L. and Esther, D., 2004. “The Impact of education on fertility and child mortality: Do fathers matter less than mothers?” NBER Working Paper 10503.

Buckles, K. S., Hungerman, D. M., 2013. "Season of Birth and Later Outcomes: Old Questions, New Answers." The Review of Economics and Statistics, MIT Press, vol. 95(3), pages 711-724, July 2013.

Budapest Institute, 2013. “Evaluation of the Roma inclusion impact of two mainstream EU funded employment programmes.” Research Report. Budapest Institute, 2013. p 92.

Calonico, S., Cattaneo, M.,D., Titiunik, R., 2014. “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs.” Econometrica Volume 82, Issue 6, November 2014., Pages 2295–2326.

Calonico, S., Cattaneo, M.,D., and Titiunik, R., 2014a. “Robust Data-Driven Inference in the Regression-Discontinuity Design.” Stata Journal, 14, Number 4, Pages 909–946.

Card, D., Kluve, J. and Weber, A., 2015. “What works? A meta-analysis of recent active labor program evaluations.” NBER Working Paper 21431.

Casey, B. M, McIntire, D.D. Ph.D., and Leveno, K.J., 2001. “The Continuing Value of the Apgar Score for the Assessment of Newborn Infants.” New England Journal of Medicine, 2001; 344:467-471February 15, 2001.

Cheng, Ming-Yen; Fan, Jianqing; Marron, J. S., 1991. “On automatic boundary corrections.” Ann. Statist. 25 (1997), no. 4, 1691--1708.

Chevalier, A. and Viitanen, T., 2003. “The long-run labor market consequences of teenage motherhood in Britain.” Journal of Population Economics. EPDP, 516. Centre for Economic Performance, London School of Economics and Political Science, London, UK.

Clark, D., Geruso, M., Royer, H., 2014. "The Impact of Education on Family Formation: Quasi-Experimental Evidence from the UK" Working Paper, Version July 21, 2014. Pages: 1-57.

Cygan-Rehm, K. and Maeder, M., 2013. "The effect of education on fertility: Evidence from a compulsory schooling reform.", Labour Economics, Volume 25, December 2013, Pages 35–48

Durst, J., 2007. "Több a kára, mint a haszna: születésszabályozás a gettóban." Demográfia, 2007. 50. évf. 1. szám 74–103.

Education Office of the Ministry of Human Resources, 2014. "Országos Jelentés. A magántanulóvá nyilvánítással kapcsolatos eljárás szakmai ellenőrzése." Oktatási Hivatal Tanügy-igazgatási és Ellenőrzési Osztály. Budapest, October 18, 2014.

Fan, E., Liu, J-T., Chen, J. C., 2014. "Is the 'Quarter of Birth' Endogenous? Evidence From One Million Siblings in Taiwan." National Bureau of Economic Research Working Paper Series, No. 20444, August 2014.

Fletcher, J. and Wolfe, 2008. "Education and labor market consequences of teenage childbearing: evidence using the timing of pregnancy outcomes and community fixed effects." NBER working paper number 13847

Fryer, Roland G. and Paul Torelli, 2010. "An empirical analysis of 'acting white"', Journal of Public Economics, 94(5–6), pp. 380–396

Hablicsek, L., 2007. "Kísérleti számítások a roma lakosság területi jellemzőinek alakulására és 2021-ig történő előrebecslésére." Demográfia, L. évf. 2007/1: 5–54.

Hahn, J., Todd, P., and Van Der Klaauw, W., 2001. "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design." Econometrica, 69(1), Pages 201-209.

Hajdú, T., 2015. "How can young Roma achieve success in Hungary? Roma in Secondary Schools and in the Labor Market." TÁRKI Social Research Institute, Budapest, January 2015.

Hámori, Sz., Köllő, J., 2011. "Kinek használ az évvesztés?" Közgazdasági Szemle, LVIII. évf., 2011. február, Pages 133–157.

Imbens, G. W., Lemieux, T., 2008. "Regression discontinuity designs: A guide to practice." Journal of Econometrics, Elsevier, vol. 142(2), February 2008, Pages 615-635.

Jacob, B. A. and Lefgren, L. 2003. "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration and Juvenile Crime." American Economic Review Vol. 93, No. 5, December 2003 (pp. 1560-1577)

Janky B., 2005. "A cigány nők helyzete és termékenysége." In: Nagy Ildikó – Pongrácz Tiborné – Tóth István György (szerk.): Szerepváltozások. Jelentés a nők és a férfiak helyzetéről. SZCSEM-Tárki, Budapest, 136-148.

Janky, B., 2007. "A korai gyermekvállalást meghatározó tényezők a cigány nők körében." Demográfia, 50. évf. 1. szám, 55-73

Jones, E. F., Forrest, J. D., Goldman, N., Henshaw, S. K., Lincoln, R., Rosoff, J. I., Westoff, C. F. and Wulf, D., 1985. "Teenage pregnancy in developed countries: determinants and policy implications." Family Planning Perspectives, 17(2), 53-63.

Kamarás, F., 2009. "Tizenéves terhességek demográfiai jellemzői." KSH, 2009.

Kearney, M. N. and Levine, P. B., 2012. “Why is the Teen Birth Rate in the United States So High and Why Does It Matter?” Journal of Economic Perspectives, Volume 26, Number 2. Spring 2012, Pages 141–166

Kemény I. and Janky B., 2003. “A cigány nemzetiségi adatokról.” Kisebbségkutatás, 2. sz. 309–315. p

Kertesi G., 2005. “A társadalom peremén: Romák a munkaerőpiacon és az iskolában.” Budapest: Osiris Kiadó, 2005. 531 p.

Kertesi G. and Kézdi G., 2006. “A hátrányos helyzetű és roma fiatalok eljuttatása az érettségihez. (Egy különösen nagy hosszú távú költségvetési nyereséget biztosító befektetés).” Budapest Working Papers on the Labour Market, 2006.

Kertesi G. and Kézdi G., 2009. “Roma és nem roma fiatalok középiskola továbbtanulása.” In: Fazekas K. (szerk.): Oktatás és foglalkoztatás. MTA Közgazdaságtudományi Intézet, Budapest, 2009. 122-136. old.

Kertesi G. and Kézdi G. 2010. “Iskolázatlan szülők gyermekei és roma fiatalok a közép-iskolában. Beszámoló az Educatio Életpálya-felvételének 2006 és 2009 közötti hullámaiból.” Társadalmi Riport, 2010. Tárki, Budapest, 2010

Kertesi, G. and Kézdi, G., 2011. ‘The Roma/non-Roma test score gap in Hungary’, American Economic Review, 101(3), pp. 519–525.

Kertesi G. and Kézdi G., 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 BWP – 2014/1

Kézdi, G., Köllő J., and J. Varga, 2009. “The Failures of ’Uncertified’ Vocational Training”, In: Fazekas, K; Köllő, J. (eds.): Munkaerőpiaci Tükör 2008. Budapest: Magyar Tudományos Akadémia, Közgazdaságtudományi Intézet, pp. 87-136.

Kézdi, G. and Simonivits, G., 2016. “Economic Hardship Triggers Identification with Disadvantaged Minorities." The Journal of Politics, 2016.

Ladányi J. – Szelényi I., 2002. “Cigányok és szegények Magyarországon, Romániában és Bulgáriában.” Szociológiai Szemle, 4. sz. 72–94. p.

Lee D. and Lemieux, T., 2010. “Regression Discontinuity Designs in Economics.” Journal of Economic Literature, American Economic Association, vol. 48(2), pages 281-355

Mártonfi, Gy. (ed.), 2011a. “Hány évig tartson a tankötelezettség?” Oktatáskutató és Fejlesztő Intézet, Budapest, 2011, Pages 1-12.

Mártonfi, Gy. (ed.), 2011b. “A 18 éves korra emelt tankötelezettség teljesülése és (mellék)hatásai.” Oktatáskutató és Fejlesztő Intézet, Budapest, 2011, Pages 1-47.

McCrary, J., and Royer, H., 2006. “The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth” (No. w12329). National Bureau of Economic Research.

McCrary, Justin and Heather Royer, 2011. “The Effect of Female Education on Fertility and Infant Health: Evidence from School Entrey Policies Using Exact Date of Birth”, American Economic Review, 101(1): 158-95

Messing, V., 2011. “Megjegyzések egy roma projekt margójára.” socio.hu, MTA Társadalomtudományi Kutatóközpont Szociológiai Intézet. 2011/2. p 8.

National Institute for Public Education, 2006. “Report on the Hungarian Public Education (Jelentés a magyar közoktatásról, in Hungarian.” Országos Közoktatási Intézet, Budapest, 2006.

Navarro Panaiagua, M. and Walker, I., 2012. "The Impact of Teenage Motherhood on the Education and Fertility of their Children: Evidence for Europe." IZA
Discussion Papers 6995, Institute for the Study of Labor (IZA).

OECD, 2014. “PISA 2012 Results in Focus. What 15-year-olds know and what they can do with what they know.” OECD, 2014. Pages 1-41.

OECD, 2015. “Education Policy Outlook: Hungary. OECD, November 2015 Pages 1-27.

Poh Lin Tan, 2014. “Re-examining the Impact of Education on Teenage Motherhood: Evidence from North Carolina.” Working Paper presented at the Population Association of America Annual Meeting 2014, Session 146. Adolescent Sexual and Reproductive Health.

Silles, M. A., 2011. “The effect of schooling on teenage childbearing: evidence using changes in compulsory education laws.” Journal of Population Economics, 24(2), 761-777.

Szikra, D., 2010. “Serdülőkori terhességek – egy jövőbeli kutatást megalapozó tanulmány és összefoglalás.” MTA Gyermekszegénység elleni program.

Tárki, 2013. "Tárki Monitor Jelentések 2012. Egyenlőtlenség és polarizálódás a magyar társadalomban." TÁRKI Social Research Institute, Budapest, April 2013. p 144.

UNFPA, 2013. “ADOLESCENT PREGNANCY: A Review of the Evidence.” United Nations Population Fund, New York, 2013.

Wilson, T.J., 2012. “Compulsory Education and Teenage Motherhood.” July 30, 2012. Working paper presented at the Royal Economic Society’s 2013 annual conference

4.3 References for Chapter 3

Allport, G. W., 1954. “The nature of prejudice.” Cambridge, MA: Addison-Wesley.

Amir, Y., 1976. “Toward the elimination of racism.” Pergamon Press, New York (1976)

Anderson, M. L., 2008. “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects.” Journal of the American Statistical Association, 103(484):1481_1495, 2008.

Angrist, J. D. and Pischke, J., 2009. “Mostly Harmless Econometrics: An Empiricist’s Companion” Princeton University Press.

Autor, D. H. , F. Levy, and R. J. Murnane, 2003. “The Skill Content of Recent Technological Change: An Empirical Exploration.” Quarterly Journal of Economics, 118(4):1279_1333, 2003.

Bárány, Z., 2002. “The East-European Gypsies: Regime Change, Marginality and Ethnopolitics.” Cambridge: Cambridge University Press.

Benjamini, Y. and Hochberg, Y, 1995. “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.” Journal of the Royal Statistical Society. Series B (Methodological) Vol. 57, No. 1 (1995), pp. 289-300.

Bernát A., Juhász A., Krekó P. and Molnár, Cs., 2013. "A radikalizmus és cigányellenesség gyökerei a szélsőjobboldal szimpatizánsai körében. In: Kolosi Tamás and Tóth István György (szerk.): Társadalmi riport 2012. Budapest: TÁRKI: 355–376

Bertrand, M., Duflo E. and Mullainathan, S., 2002. "How much should we trust differences-in-differences estimates?" NBER Working Paper 8841. Pages 1-39.

Blascovich J. and Tomaka, J., 1991. "Measures of personality and social psychological attitudes."

Bogardus, E., 1928. "Immigration and Race Attitudes."

Bowles, S., Gintis, H. and Osborne, M., 2001. "The determinants of earnings: A behavioral approach." Journal of Economic Literature, 2001/39(4), 1137–1176.

Card, D. and Rothstein, J., 2006. "Racial segregation and the black–white test score gap." National Bureau of Economic Research: Working Paper #12078.

Carneiro, P. and Heckman, J. J., 2003. "Human Capital Policy." IZA Discussion Paper No. 821

Cole, M., 1990. "Cognitive development and formal schooling: The evidence from cross-cultural research."In L. C. Moll (Ed.), Vygotsky and education: Instructional implications of sociohistorical psychology (pp. 89-110). New York: Cambridge University Press.

Crandall, V. C., Crandall, V. J., and Katkovsky, W., 1965. "A children's social desirability questionnaire." Journal of Consulting Psychology, 29(1), 27-36.

DeBord, K., 1997. "Brain development". [Extension Publication]. Raleigh, NC: North Carolina Cooperative Extension Service.

Flabbi, L. – Paternostro, S. and Tiongson, E. R.: "Returns to education in the economic transition: a systematic assessment using comparable data." World Bank Policy Research Working Paper Series, 2007/4225.

Folkman, S. and Lazarus, R. S., 1984. "Stress, appraisal, and coping." New York: Springer

Ford, W. S, 1986. "Favorable intergroup contact may not reduce prejudice: Inconclusive journal evidence, 1960 –1984." Sociology and Social Research, 70, 256 –258.

FRA-UNDP, 2012. "The situation of Roma in 11 EU Member States", European Agency of Fundamental Rights and UNDP.

Ben B. Hansen, Paul R. Rosenbaum, and Dylan S. Small, 2014. "Clustered Treatment Assignments and Sensitivity to Unmeasured Biases in Observational Studies". Journal of the American Statistical Association, 109(505): 133-144.

Hajdú T., Kertesi, G. and Kézdi G., 2014. "Roma fiatalok a középiskolában. Budapest Working Papers On The Labour Market. BWP 2014/7. Pages 1-39.

Hansen, B. B., Rosenbaum, P. R., and Small, D. S., 2014. "Clustered Treatment Assignments and Sensitivity to Unmeasured Biases in Observational Studies," Journal of the American Statistical Association, 109, 133–144.

Harter, S., 1983. "Developmental perspectives on the self-system." In: Hetherington EH, editor. Handbook of Child Psychology. Vol. 4. Socialization, personality and social development. New York: John Wiley, 275-385.

Harter S., 1985. “Manual of the Self-Perception Profile for Children.” University of Denver: Denver Co.

Heckman, J. J. and Rubinstein, Y., 2001. “The Importance of Noncognitive Skills: Lessons from the GED Testing Program.” American Economic Review, 2001/91(2), 145–149.

Heckman, J. J., Stixrud, J. and Urzua, S. S., 2006. “The Effects of Cognitive and Non-cognitive Abilities on Labor Market Outcomes and Social Behavior.” NBER Working Paper, 12006.

Holm, S., 1979. "A simple sequentially rejective multiple test procedure." Scandinavian Journal of Statistics. 6 (2): 65–70.

Imbens, G. W. and J. M. Wooldridge, 2009. “Recent Developments in the Econometrics of Program Evaluation.” Journal of Economic Literature, 47:1, 5–86.

Janky B., 2004. “The Income Situation of Gipsy Families.” In: Kolosi Tamás, Tóth István György, Vukovich György Social Report 2004. 487 p. p. Budapest: TÁRKI, pp. 386-399.

Jackson, J. W., 1993. “Contact theory of intergroup hostility: A review and evaluation of the theoretical and empirical literature.” International Journal of Group Tensions, 23, 43– 65.

Kemény, I., 2004. „A magyarországi cigány népesség demográfiája” (The demography of the Hungarian Gypsy population.). Demográfia 3-4.

Kemény, I. and Janky, B., 2006. “Roma Population of Hungary 1971-2003.” In: Kemény, I. (ed.) Roma of Hungary: East European Monographs. New York: Social Science Monographs; Atlantic Research and Publ., pp. 70-225.

Kertesi, G. and Varga, J. 2004. “Iskolázottság és foglalkoztatás.” In: Fazekas Károly (szerk.), Munkaerőpiaci tükör 2004., MTA Közgazdaságtudományi Intézet, Budapest, 2004.

Kertesi, G. and Varga, J., 2015. “Foglalkoztatás és iskolázottság Magyarországon.” Közgazdasági. Szemle, 52. 2005. 7-8.sz. pp. 633-662.

Kertesi, G. and Kézdi, G., 2011a. “Roma employment in Hungary after the postcommunist transition.” Economics of Transition, 19(3), pp. 563-610.

Kertesi, G. and Kézdi, G., 2011b. “The Roma/non-Roma test score gap in Hungary.” American Economic Review, 101(3), pp. 519–525.

Kertesi, G. and Kézdi, G., 2015. “On the Test Score Gap between Roma and non-Roma Students in Hungary and its potential causes.” Economics of Transition 23(3).

Kertesi, G. and Kézdi, G., 2013. “Ethnic segregation between Hungarian schools: Long-run trends and geographic distribution.” Hungarian Statistical Review. Special number 16.

Kézdi G., 2004. “Iskolázottság és keresetek.” In: Fazekas, K. (szerk.), Munkaerőpiaci tükör 2004., MTA Közgazdaságtudományi Intézet, Budapest, 2004.

Kézdi G. and Surányi, É., 2009. “A Successful School Integration Program. An Evaluation of the Hungarian National Government’s School Integration Program 2005–2007.” Research Report. Roma Education Fund, Budapest, 2009.

Kézdi G. and Surányi, É., 2010. "Measuring Non-Cognitive Skills in the Evaluation of the National Education Integration Program of Hungary." Budapest Working Papers On The Labor Market, BWP 2010/1.

Köllő J., 2006. "Workplace Literacy Requirements and Unskilled Employment in East-Central and Western Europe. Evidence from the International Adult Literacy Survey (IALS)." Budapest Working Papers On The Labor Market, BWP 2006/7.

Köllő, J., 2009. "A pálya szélén : iskolázatlan munkanélküliek a posztszocialista gazdaságban." Budapest, Osiris. 245 p.

La Greca, A. M., Kraslow Dandes, S., Wick P. P., Shaw K. and Stone, L. W., 1988. "Development of the Social Anxiety Scale for Children: Reliability and Concurrent Validity." Journal Journal of Clinical Child Psychology Volume 17, 1988 - Issue 1. Pages 84-91.

Leary, M. R. and Meadows, S., 1991. "Predictors, elicitors, and concomitants of social blushing." Journal of Personality and Social Psychology, Vol 60(2), Feb 1991, 254-262.

Milcher, S., 2006. "Poverty and the determinants of welfare for Roma and other vulnerable groups in South-eastern Europe", Comparative Economic Studies, Vol. 48, No.1, pp. 20-35.

O'Higgins, N. and Ivanov, A., 2006. "Education and Employment Opportunities for the Roma." Comparative Economic Studies, 2006, 48, (6–19)

Pettigrew, T. F., 1971. "Racially separate or together?" New York: McGraw-Hill

Pettigrew, T. F. and Tropp, L. R., 2006. "A Meta-Analytic Test of Intergroup Contact Theory." Journal of Personality and Social Psychology, 2006, Vol. 90, No. 5, 751–783

Reynolds, W. M., 1982. "Development of reliable and valid short forms of the Marlowe-Crowne social desirability scale." Journal of Clinical Psychology, 38(1), 119-125.

Rosenberg, M., 1965. "Society and the adolescent self-image." Princeton, NJ: Princeton University Press.

Rotter, J. B., 1996. "Generalized expectancies of internal versus external control of reinforcements." Psychological Monographs, 1996/80 (609).

Shore, R., 1997. "Rethinking the brain." New York: Families and Work Institute.

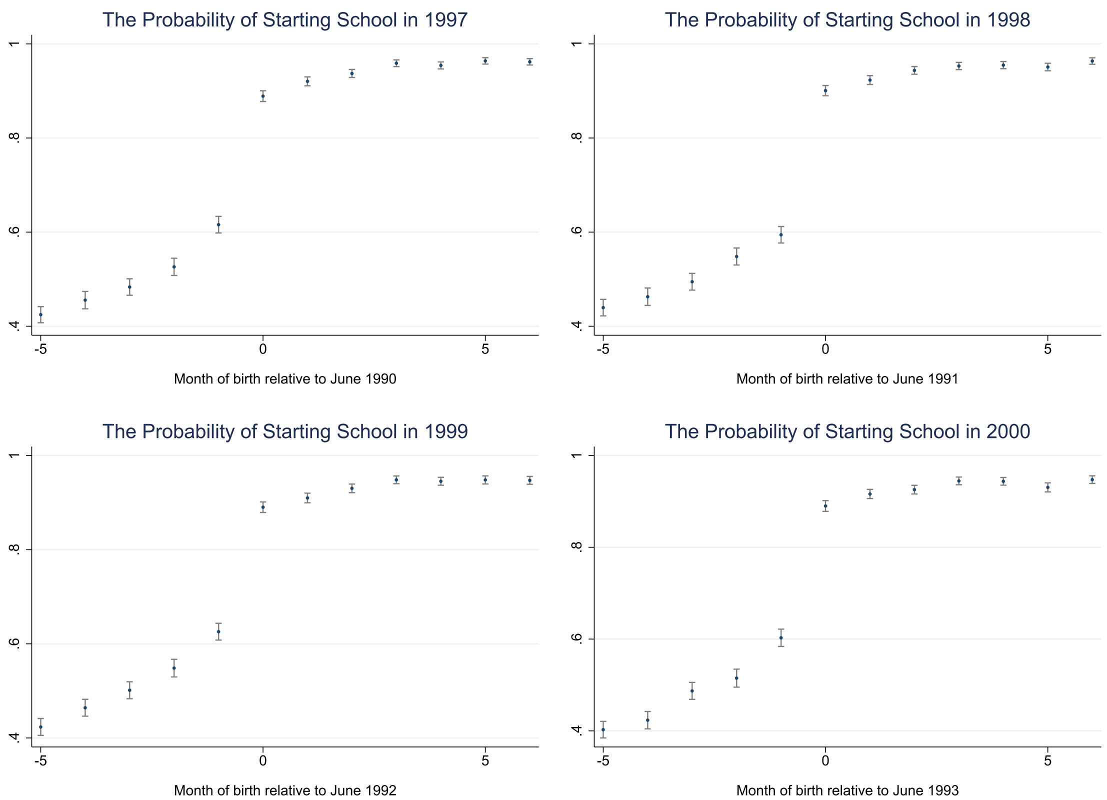
Sidanius, J. and Pratto, F., 1999. "Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression." Cambridge, MA: Cambridge University Press.

Appendix A

Appendix for Chapter 1

A.1 Appendix

Figure A.1: The Probability of Starting School in 1997-2000, Children of Mothers with a Primary Degree



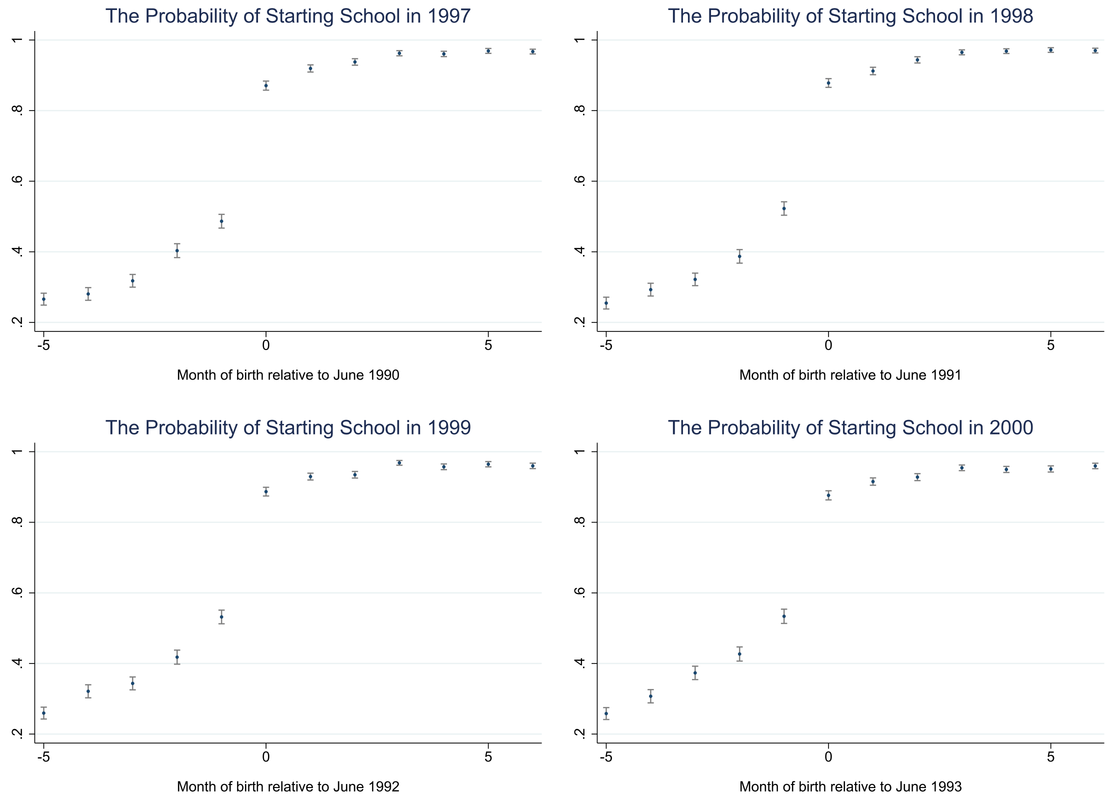
The average probability of starting school in 1997 for those born in 1990, in 1998 for those born in 1991, in 1999 for those born in 1992 and in 2000 for those born in 1993, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June. No. of individual observations: 36,053, 35,399, 34,266 and 33,045, respectively.

Table A.1: The 10-Grade Waves of the National Assessment of Basic Competencies (NABC) Survey Data Used in this Paper

Wave	Birth dates of the main cohort covered by each wave	Coverage	No. of students	No. of students with complete birth data	No. of students born June 90 - May 91	No. of students born June 91 - May 92	No. of students born in May 1991	No. of students born in June 1991	Sample taken
2006	June 89- May 90	30 students per school programs	43,602	35,115	629	4	1	0	no
2007	June 90 - May 91	30 students per school programs	43,775	36,605	26,085	664	1,421	284	no
2008	June 91 - May 92	all students	112,409	111,343	17,502	60,153	3,350	5,516	37,654*
2009	June 92 - May 93	all students	108,960	108,907	5,312	26,402	854	1,176	37,289*
2010	June 93 - May 94	all students	107,274	107,142	1,270	6,194	198	221	36,935*
Total			401,620	399,112	50,798	93,417	5,824	7,197	

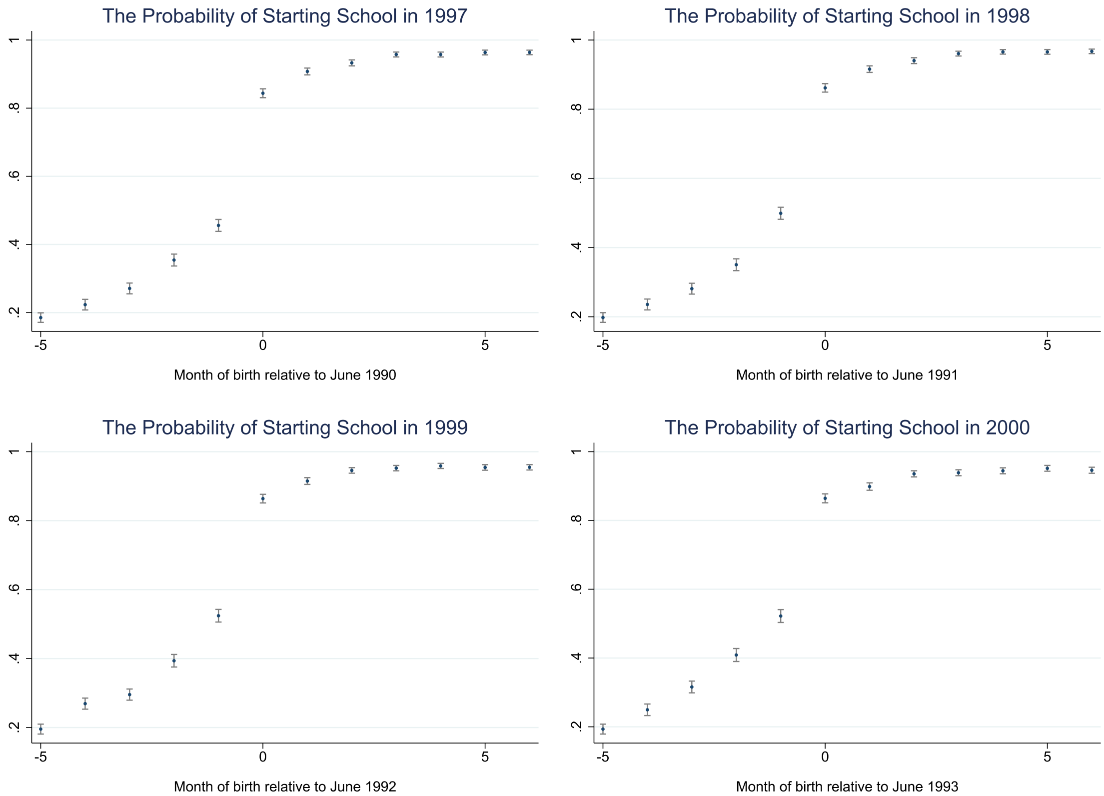
Data source: own calculations from the 10-grade waves of the 2006-2010 NABC data. *The 2006-2007 waves cover a subsample of 10-graders while the 2008-10 samples cover the full student population. Those born right before the cutoff are covered by the earlier waves, while those born after covered by the later ones. Thus, random samples of 22 students are taken from each program in each school to match the sampling strategy of waves 2006-2007, taken into consideration the fact that in the waves 2006-2007 the share of observations with complete information about year and month of birth is much lower. The sampling number 22 was chosen to match the number of observations in the samples of waves 2008-10 to the number of observations with date of birth information in waves 2006-2007.

Figure A.2: The Probability of Starting School in 1997-2000, Children of Mothers with a Vocational Training School Degree



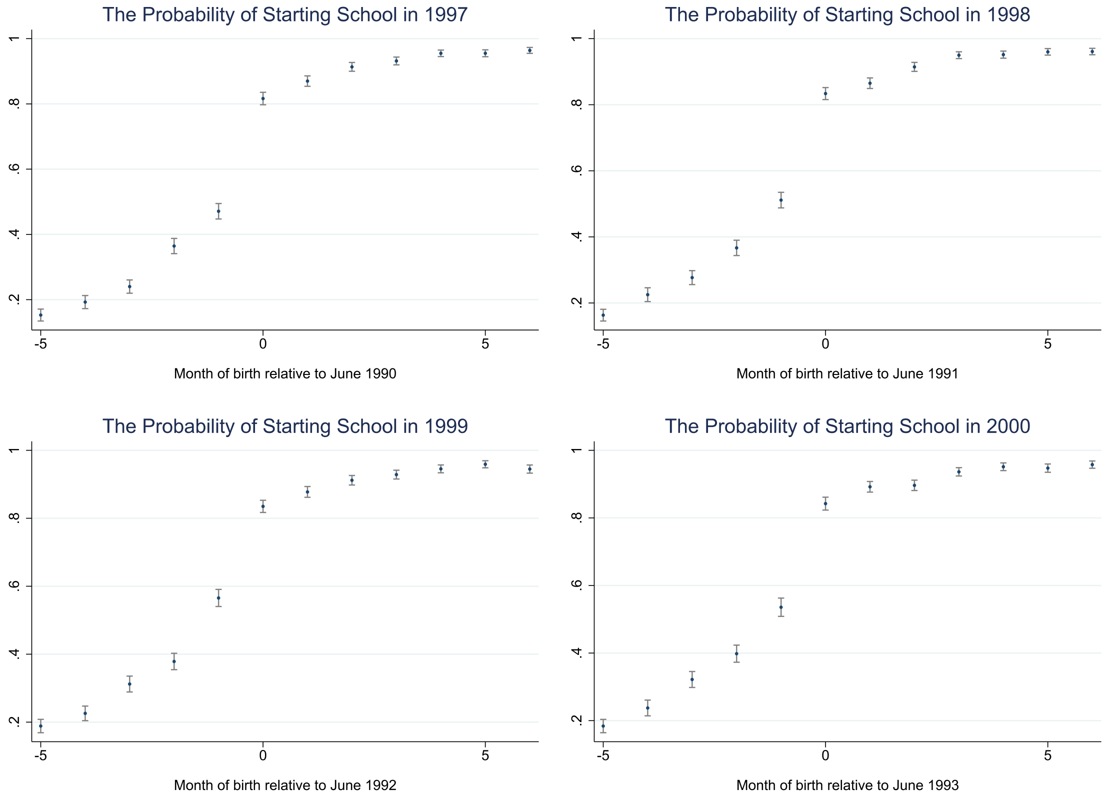
The average probability of starting school in 1997 for those born in 1990, in 1998 for those born in 1991, in 1999 for those born in 1992 and in 2000 for those born in 1993, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June. No. of individual observations: 30,784, 30,979, 30,410 and 29,894, respectively.

Figure A.3: The Probability of Starting School in 1997-2000, Children of Mothers with a High School Degree



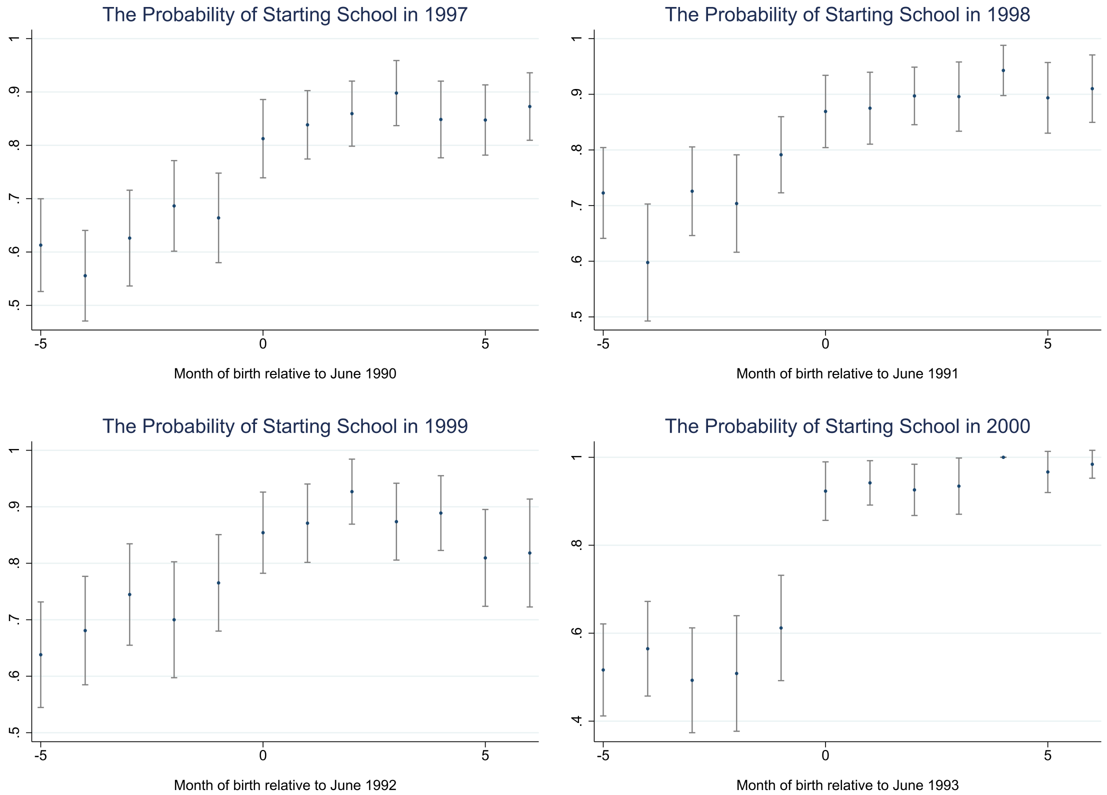
The average probability of starting school in 1997 for those born in 1990, in 1998 for those born in 1991, in 1999 for those born in 1992 and in 2000 for those born in 1993, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June. No. of individual observations: 35,993, 36,197, 34,557 and 32,843, respectively.

Figure A.4: The Probability of Starting School in 1997-2000, Children of Mothers with a Tertiary Degree



The average probability of starting school in 1997 for those born in 1990, in 1998 for those born in 1991, in 1999 for those born in 1992 and in 2000 for those born in 1993, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June. No. of individual observations: 19,477, 19,603, 18,190 and 16,942, respectively.

Figure A.5: The Probability of Starting School in 1997-2000, Information on Maternal Education is missing



The average probability of starting school in 1997 for those born in 1990, in 1998 for those born in 1991, in 1999 for those born in 1992 and in 2000 for those born in 1993, plotted with the 95% confidence intervals of the means. Data Source: own estimation from the 2001 Hungarian Census. 0 on the x-axis refers to being born in June. No. of individual observations: 1,412, 1,308, 1,077 and 858, respectively.

Table A.2: Literature on the Wage Returns of Higher CSL Age

Paper	Country and year of the reform	CSL age increase from ... to ...	Other elements of the reform?	Identification	Main result	Main conclusion
			Wage returns			
Harmon and Walker (1995)	England & Wales in 1947 and 1972	Age 14 to 15 and	supply side expansion	comparing pre- and post-reform cohorts*	10-15% wage return	wage returns are large
Meghir and Palme (2005)	Sweden, 1940's	Grade 7-8 to 9 (age 14-15 to 16)	no streaming, means-tested subsidies	Diff-in-diffs	positive (-) wage return on children of unskilled (skilled) fathers	abolishment of streaming might have reduced school quality
Oreopoulos (2007)	England,1947	Age 14 to 15	supply side expansion	Diff-in-diffs	15% lifetime wealth returns	teenagers must be myopic
Oosterbeek and Webbink (2007)	The Netherlands, 1975	Age 15 to 16	vocational education extended from 3 to 4 years	Diff-in-diffs	no effects on wages	general education in the extra year, one year less work experience
Pischke and Wachter (2008)	Germany, 1940's-1950's	Grade 8 to 9 (age 14 to 15)		Diff-in-diffs	no effect on wages	important skills are learn in earlier Grades
Devereux and Hart (2010)	England,1947	Age 14 to 15	supply side expansion	RDD	3% wage return	no effect on qualifications, heterogeneous LATE
Grenet (2013)	France, 1967 and Britain, 1972	Age 15 to 16		RDD	6-7% wage return in GB, no return in FR	effect on qualifications is key
Leonard and Sweetman (2013)	Newfoundland, 1983	Grade 11 to 12	(no change in the curriculum)	RDD	little to no general equilibrium impact on wages	too early measurement, economic crisis

Source: own collection. Papers investigating the wage effects of CSL age raise directly or using it as instrument only. Papers using school entry policies (i.e. Angrist and Kruger, 1991), several types of CSL legislation changes at the same time (i.e. Acemoglu and Angrist, 2001), or school time shortening measures (i.e. Büttner and Thomsen, 2010) are excluded. *Callan and Harmon (1999), Brunello and Miniaci (1999), Brandolini and Cipollone (2002), Levine and Plug (1999), Vieira (1999) and Pons and Gonzalo (2002) are using the same pre- vs. post-cohort methodology, criticized later by Card (1999) and Oreopoulos (2006) for not controlling for cohort fixed effects (Grenet, 2013).

A.2 Appendix

Table A.3: The Definition of Outcome Variables Constructed from the 2011 Hungarian Census

Outcome variable	Definition	Unit of measurement
	Length of schooling	
Number of successfully completed years in the school system	The number of academic years one completed in each school types together. It does not include grade retentions, and unfinished or not incomplete (failed) school years. It is smaller or equal to the no. of years one spent in school, which is not measured by the data.	Academic years
	Highest degree obtained	
Any secondary school	Earned at least a secondary degree from any track.	binary variable
Vocational training school	Earned a vocational training school degree (but not a high school degree).	binary variable
Professional high school	Earned a high school degree (érettségi) in a professional high school.	binary variable
Academic high school	Earned a high school degree (*érettségi*) in an academic high school.	binary variable
	Secondary school track choice	
Any secondary school	Finished at least one academic year successfully in any secondary school above Grade 8.	binary variable
Vocational training school	Finished at least one academic year in a vocational training school.	binary variable
Professional high school	Finished at least one academic year in a professional high school.	binary variable
Academic high school	Finished at least one academic year in an academic high school above Grade 8.	binary variable
	Dropping out of secondary school	
Any secondary school	Finished at least one academic year in a secondary school but did not earn any secondary degree and was not in school at the time of the Census.	binary variable
Vocational training school	Finished at least one academic year in a vocational training school but have not earned any secondary degree and was not in school at the time of the 2011 Hungarian Census.	binary variable
Professional high school	Finished at least one academic year in a professional high school but did not earn any secondary degree and was not in school at the time of the 2011 Hungarian Census.	binary variable
Academic high school	Finished at least one academic year in an academic high school but did not earn any secondary degree and was not in school at the time of the 2011 Hungarian Census.	binary variable

Table A.4: Education Outcomes of Students Born in 1990

	Total sample	Mother's highest education at giving birth				
		Primary	Vocational	High school	Tertiary	Missing
Number of years completed successfully						
Primary, secondary and tertiary education	12.5	11.4	12.7	13.2	13.5	12.3
Probability of starting secondary school						
Any secondary	0.907	0.796	0.959	0.986	0.993	0.869
Vocational training school	0.225	0.380	0.286	0.123	0.035	0.203
Professional high school	0.326	0.275	0.425	0.405	0.215	0.268
Academic high school	0.447	0.240	0.371	0.550	0.799	0.471
Probability of earning a secondary degree						
Any secondary degree	0.837	0.671	0.898	0.953	0.979	0.789
Vocational training school	0.158	0.270	0.214	0.084	0.020	0.137
Professional high school degree	0.266	0.210	0.353	0.345	0.172	0.214
Academic high school degree	0.414	0.192	0.331	0.525	0.787	0.440
Probability of dropping out of secondary school						
Any secondary	0.045	0.102	0.031	0.026	0.005	0.055
Vocational training school	0.115	0.149	0.06	0.055	0.471	0.150
Professional high school	0.027	0.054	0.019	0.013	0.007	0.037
Academic high school	0.016	0.046	0.015	0.008	0.002	0.017
No. of obs.	118,987	27,474	21,055	28,827	10,302	31,329

Data source: 2011 Hungarian Census.

A.3 Appendix

Table A.5: Effects on School Choice (ITT effects, 100-day bandwidth)

Tracks	Total sample	Mother's highest education at giving birth: Primary	Vocational	High school	Tertiary	Missing
Effect on the probability of finishing at least the first year in a secondary school						
Any sec school	-0.002	0.003	-0.003	-0.005*	-0.000	0.000
	(0.003)	(0.007)	(0.005)	(0.002)	(0.001)	(0.007)
Corrected p-values	0.544	0.806	0.727	0.059	0.960	0.958
Vocational training	-0.015*	-0.050***	-0.028	-0.009*	0.011***	0.005
school	(0.005)	(0.006)	(0.018)	(0.003)	(0.002)	(0.009)
Corrected p-values	0.055	0.002	0.250	0.053	0.005	0.742
Professional high	-0.008	0.003	-0.008	-0.031***	-0.051***	0.023
school	(0.004)	(0.011)	(0.012)	(0.006)	(0.009)	(0.006)
Corrected p-values	0.156	0.859	0.634	0.006	0.006	0.024
Academic high school	0.021***	0.050***	0.038***	0.036***	0.022**	-0.023**
	(0.004)	(0.009)	(0.008)	(0.006)	(0.013)	(0.010)
Corrected p-values	0.005	0.006	0.007	0.005	0.218	0.119
No. of obs.	67,259	15,588	12,097	16,697	5,950	16,928

Local linear kernel regressions using 100-day bandwidth. Robust standard errors clustered by year-and-month-of-birth are in parentheses. P-values are corrected by the number of hypothesis tests (72) done together in Tables A.5 - A.7 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after multiple testing correction. The same results estimated by CCT bandwidths are in Table 1.11 in Section 6.

Table A.6: Effects on Dropout Rates (ITT effects, 100-day bandwidth)

Tracks	Total sample	Mother's highest education at giving birth: Primary	Vocational	High school	Tertiary	Missing
Effect on the probability of dropping out of ...						
Any secondary school	0.009***	0.017*	0.005	0.004*	-0.002*	0.013
	(0.002)	(0.006)	(0.003)	(0.002)	(0.001)	(0.006)
	[61,314]	[12,532]	[11,574]	[16,405]	[5,904]	[14,899]
Corrected p-values	0.005	0.054	0.262	0.077	0.069	0.126
Vocational training	0.029***	0.052***	-0.002	0.028*	0.061***	0.016
school	(0.005)	(0.012)	(0.007)	(0.012)	(0.012)	(0.013)
	[14,319]	[5,841]	[3,210]	[1,841]	[189]	[3,238]
Corrected p-values	0.005	0.009	0.843	0.082	0.005	0.364
Professional high	0.004	0.002	0.006	0.001	0.001	0.003
school	(0.003)	(0.005)	(0.004)	(0.002)	(0.004)	(0.010)
	[23,387]	[4,578]	[5,480]	[7,011]	[1,333]	[4,985]
Corrected p-values	0.268	0.808	0.274	0.734	0.856	0.792
Academic high school	0.003	-0.005	0.012***	-0.001	-0.005	0.006
	(0.002)	(0.009)	(0.002)	(0.003)	(0.001)	(0.004)
	[30,878]	[3,732]	[4,585]	[9,460]	[11,081]	[8,289]
Corrected p-values	0.424	0.745	0.007	0.836	0.006	0.330

Local linear kernel regressions using 100-day bandwidth. Robust standard errors clustered by year-and-month-of-birth are in parentheses, number of observations are in brackets. P-values are corrected by the number of hypothesis tests (72) done together in Tables A.5 - A.7 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after multiple testing correction. The same results estimated by CCT bandwidths are in Table 1.12 in Section 6.

Table A.7: Effects on School Completion (ITT effects, 100-day bandwidth)

	Total sample	Mother's highest education at giving birth: Primary	Vocational	High school	Tertiary	Missing
	Effect on the probability of gaining a secondary degree					
Any sec degree	-0.017***	-0.014	-0.022**	-0.014***	-0.007	-0.017**
	(0.0013)	(0.009)	(0.005)	(0.002)	(0.012)	(0.005)
Corrected p-values	0.008	0.252	0.013	0.003	0.705	0.026
Vocational training	-0.019***	-0.050***	-0.029	-0.015***	0.004	0.002
school degree	(0.003)	(0.003)	(0.014)	(0.003)	(0.002)	(0.005)
Corrected p-values	0.005	0.000	0.141	0.006	0.161	0.795
Professional high	-0.015**	-0.017*	-0.020*	-0.028**	-0.037***	0.010
school degree	(0.004)	(0.007)	(0.007)	(0.007)	(0.008)	(0.009)
Corrected p-values	0.028	0.087	0.059	0.012	0.009	0.407
Academic high school	0.018***	0.055***	0.028*	0.030**	0.025	-0.030*
degree	(0.003)	(0.005)	(0.010)	(0.007)	(0.018)	(0.010)
Corrected p-values	0.005	0.001	0.053	0.012	0.310	0.053
No. of obs.	67,259	15,588	12,097	16,697	5,950	16,928

Local linear kernel regressions using 100-day bandwidth. Robust standard errors clustered by year-and-month-of-birth are in parentheses. P-values are corrected by the number of hypothesis tests (72) done together in Tables A.5 - A.7 using the FDR multiple testing procedure by Benjamini and Hochberg (1995). A */**/*** indicates significance on 10%/5%/1% level after multiple testing correction. The same results estimated by CCT bandwidths are in Table 1.10 in Section 6.

A.4 Appendix

The bandwidth optimization routine of Calonico, Cattaneo and Titiunik (2014)

The CCT bandwidth optimization procedure works as follows. Let's denote the p^{th}-order local polynomial reduced form estimator along a series of bandwidths h_n, as $\hat{\beta}_p(h_n)$. Then,

$$\hat{\beta}_p(h_n) = \hat{\beta}_{+,p}(h_n) - \hat{\beta}_{-,p}(h_n)$$

where $\hat{\beta}_{+,p}(h_n)$ and $\hat{\beta}_{-,p}(h_n)$ and denote the intercept at the cutoff of a weighted p^{th}-order local polynomial regression for treated (above-the-cutoff) and control (below-the-cutoff) observations only. More precisely, $\hat{\beta}_{+,p}(h_n)$ and $\hat{\beta}_{-,p}(h_n)$ are the solutions of the minimization problems of minimizing the following sum of squared errors at bandwidths h_n:

$$\hat{\beta}_{+,p}(h_n) = argmin \sum_{i=1}^{n} 1(x_i \geqslant 0) \left\{y_i - r_p(x_i)'\beta\right\}^2 K_{h_n}(x_i)$$

$$\hat{\beta}_{-,p}(h_n) = argmin \sum_{i=1}^{n} 1(x_i < 0) \left\{y_i - r_p(x_i)'\beta\right\}^2 K_{h_n}(x_i)$$

where
$1()$ is an indicator function;
$r_p(x_i)$ is a p^{th}- degree polynomial function of the running variable;
$K_h(u) = K(u/h)/h$ with $K()$ being a Kernel function; and
$h_n$is a positive bandwidth sequence.

The $\hat{\beta}_p(h_n)$ local polynomial regression estimator offers a series of consistent estimators of the treatment effect at the cutoff. The estimator is consistent if in its MSE expansion as

$$MSE_p(h_n) = E\left[\left\{\hat{\beta}_p(h_n) - \beta\right\}^2 |R_n\right] \approx h_n^{2(p+1)} B_{n,p}^2 + \frac{1}{n * h_n} V_{n,p}$$

where $B_{n,p}$and $V_{n,p}$ represent the leading asymptotic bias and variance of $\hat{\beta}_p(h_n)$, $h_n \to 0$ and $n * h_n \to \infty$.

Appendix B

Appendix for Chapter 2

B.1 Appendix

Table B.1: The Balance of the Linked Data

	Mean, unlinked observations	No. of unlinked observations	Mean, linked observations	No. of linked observations	T-test p-value
All events					
Born after June 1, 1991	0.48	2,234	0.48	1,480	0.88
Age	16.71	2,234	16.69	1,480	0.48
Completed elementary school	0.76	2,234	0.83	1,480	0.00
Lives in a settlement with less than 10,000 ppl	0.67	2,234	0.72	1,480	0.00
Live births					
Born after June 1, 1991	0.50	924	0.50	701	0.94
Age	16.77	924	16.69	701	0.09
Completed elementary school	0.72	924	0.76	701	0.05
Lives in a settlement with less than 10,000 ppl	0.71	924	0.77	701	0.01
Abortions					
Born after June 1, 1991	0.47	1089	0.47	670	0.84
Age	16.65	1089	16.69	670	0.42
Completed elementary school	0.79	1089	0.89	670	0.00
Lives in a settlement with less than 10,000 ppl	0.64	1089	0.67	670	0.26
Miscarriages					
Born after June 1, 1991	0.48	203	0.47	106	0.85
Age	16.74	203	16.62	106	0.32
Completed elementary school	0.82	203	0.89	106	0.12
Lives in a settlement with less than 10,000 ppl	0.65	203	0.69	106	0.50
Stillbirths					
Born after June 1, 1991	0.44	18	0.67	3	0.50
Age	16.92	18	17.69	3	0.20
Completed elementary school	0.89	18	1.00	3	0.57
Lives in a settlement with less than 10,000 ppl	0.83	18	1.00	3	0.47

Sample: women born within 180 days before or after June 1, 1991. Women living in settlements with less than 50,000 people only.

Table B.2: The Effects of Personal Characteristics on Vital Statistics Database Events Being Linked to the Census

	All events	Live births	Abortions	Miscarriages	Stillbirths
Age	-0.025***	-0.035**	-0.022*	-0.044	0.198
at the event	(0.006)	(0.013)	(0.008)	(0.028)	(0.140)
Educational status at	0.124***	0.094***	0.188***	0.210*	0.195
the event	(0.015)	(0.024)	(0.026)	(0.078)	(0.259)
Settlement	-0.006**	-0.001	-0.008	-0.003	0.011
size	(0.002)	(0.005)	(0.004)	(0.009)	(0.056)
No. of obs	3,714	1,625	1,756	309	21

Sample: women born within 180 days before or after June 1, 1991. Women living in settlements with less than 50,000 ppl only. Linear probability models, one regression per event type. Dependent variable: whether event i could be linked to the Census. Additional control variables: number of days of birth relative to June 1, 1991, county fixed effects. Robust standard errors clustered by month of birth are in parenthesis.

Table B.3: Treatment Effect on the Probability of Linking Pregnancy Events to the Census

Events in the Vital Statistics Database	Local linear model	Global polynomial model
All events	-0.049**	-0.009
	(0.017)	(0.027)
	[3,702]	[18,999]
Live birth only	-0.026	0.032
	(0.060)	(0.048)
	[1,625]	[8,338]
Abortion only	-0.061**	-0.084*
	(0.025)	(0.042)
	[1,759]	[9,020]
Miscarriage only	-0.098	-0.320**
	(0.091)	(0.092)
	[309]	[107]
Stillbirth only	0.113	0.196
	(0.334)	(0.390)
	[21]	[1,534]
Sample	Women born within 180 days before or after June 1, 1991	Women born in 1988-92
Control variables	Linear function of the running variable separately below and above the cutoff.	4th order function of the running variable separately below and above the cutoff; FE: year of birth, month of birth, day of the week at birth, county, settlement type.

Linear probability models. Left hand side variable: whether observation i in the Vital Statistics database could be linked to the Census. Women living in settlements with less than 50,000 pax only. Robust standard errors clustered by birth year and month in parenthesis, observation numbers in brackets. A negative treatment effect indicates that the pregnancy and birth-related events of those born after June 1, 1991, are less likely to be linked to the Census.

Appendix C

Appendix for Chapter 3

C.1 Appendix

Table C.1: The Inverse Covariance Matrix and the Weights of the Index of Cognitive Measures

	Reading test score	Secondary school acceptance
Reading test score	1.724047	-0.60052375
Secondary school acceptance	-0.60052375	2.1362163
Weights	.42250173	.57749827

Table C.2: The Inverse Covariance Matrix and the Weights of the Index of Non-cognitive Measures

	Locus of control	Rosenberg	SPPC	Coping
Locus of control	1.8498615	-0.18055095	-0.25961007	-0.11797495
Rosenberg	-0.18055095	2.1670802	-0.9262326	-0.21138207
SPPC	-0.25961007	-0.9262326	2.5069004	-0.36711616
Coping	-0.11797495	-0.21138207	-0.36711616	2.0290039
Weights	0.29177609	0.19175356	0.21547715	0.3009932

Table C.3: The Inverse Covariance Matrix and the Weights of the Index of Inter-ethnic Relations

	Social anxiety	Social dominance	Social distance	Stereotypes
Social anxiety	1.7044797	-0.06261108	-0.05462732	0.12311984
Social dominance	-0.06261108	1.646431	-0.0485324	-0.00653806
Social distance	-0.05462732	-0.0485324	2.2569782	-1.0978078
Stereotypes	0.12311984	-0.00653806	-1.0978078	2.8086074
Weights	0.27935653	0.24969356	0.17248023	0.29846968

Table C.4: The Inverse Covariance Matrix and the Weights of the Index of the Summary Index

	Reading test score	Secondary school acceptance	Locus of control	Rosenberg	SPPC	Coping	Social anxiety	Social dominance	Social distance	Stereotypes
Reading	1.784	-0.546	0.027	0.082	-0.182	-0.091	-0.101	-0.190	-0.031	0.120
Sec sch. a.	-0.546	2.230	-0.005	0.107	-0.220	-0.096	-0.002	-0.176	0.232	0.128
Locus of c.	0.027	-0.005	1.865	-0.202	-0.260	-0.115	0.032	-0.155	-0.028	0.013
Rosenberg	0.082	0.107	-0.202	2.278	-0.887	-0.159	-0.403	0.171	0.021	0.040
SPPC	-0.182	-0.220	-0.260	-0.887	2.648	-0.271	-0.378	-0.029	-0.093	-0.076
Coping	-0.091	-0.096	-0.115	-0.159	-0.271	2.166	-0.378	-0.113	0.113	0.200
Social a.	-0.101	-0.002	0.032	-0.403	-0.378	-0.378	2.037	-0.041	-0.068	0.065
Social do.	-0.190	-0.176	-0.155	0.171	-0.029	-0.113	-0.041	1.729	-0.074	-0.051
Social di.	-0.031	0.232	-0.028	0.021	-0.093	0.113	-0.068	-0.074	2.290	-1.070
Stereot.	0.120	0.128	0.013	0.040	-0.076	0.200	0.065	-0.051	-1.070	2.853
Weights	0.075	0.143	0.101	0.090	0.022	0.108	0.066	0.092	0.111	0.192

Table C.5: Outcome Measures

Outcome	Interpretation - what is measured?	Methodology
	Cognitive outcomes	
Literacy test score	reading comprehension skills	reading tests were developed following the principles of the PISA test
Direction of further studies after completing elementary school	whether one continues his/her studies in a higher-tier secondary school ending with a graduating examination	administratively observed binary variable: 1 if accepted to a higher-tier secondary school and 0 otherwise.
	Non-cognitive outcomes	
Locus of control	the extent to which individuals believe that they can control events that affect them	four-item version of the Rotter locus of control test based on Rotter (1996)
Rosenberg Self-Esteem Scale (RSE)	a person's overall evaluation or appraisal of his or her own worth	four-item version of the Rosenberg Self-Esteem Scale test based on Rosenberg (1965)
Harter Self Perception Profile for Children (SPPC)	general self-esteem and self-esteem in specific areas (academic competence, social acceptance, external appearance, behavioral discipline)	ten-item version of the Harter Self Perception Profile for Children (SPPC) based on Harter (1985, 1983)
Coping ability	the ability that helps individuals to get over with difficult or conflict situations	four-item test in the spirit of the SPPC test
	Inter-ethnic relations	
Ethnic stereotypes	a simplistic, exaggerated and overgeneralized judgment made about the members of a social group	five-item scale adapted from adult tests based on focus group conversations
Bogardus social distance	social distance that one wishes to keep from the members of a given social group	eight-item scale; adapted from the adult Bogardus SD test based on focus group conversations
Social Dominance Orientation (SDO)	the extent to which one believes in social hierarchy and ethnic inequality	six-item scale adapted from adult tests
Social anxiety	discomfort or a fear in social interactions that involves fear of a negative judgment, fear of interaction, and retreat	five-item scale adapted from children's test by La Greca et al. (1988)

*All noncognitive and inter-ethnic measures were standardized on a nationally representative sample. Thus, they mean deviation from the national average, and the unit of measurement is the national standard deviation. Reading test score was standardized within sample: its mean is zero and its standard deviation is one. **All measures were collected in 8th grade by tests and surveys completed by the students, except for the direction of further studies which was collected from school administration after the end of the school year.

C.2 Appendix

Table C.6: Effects on Academic Achievement, Socio-Emotional Skills and Anti-Roma Sentiments

	Index of cognitive outcomes	Index of non-cognitive outcomes	Index of tolerance measures
	Model 1		
Program	0.119	0.117	0.119
Cl robust SE	0.068	0.048	0.053
Uncorrected p-values	0.088	0.019	0.030
FDR-control test rejects on...	-	10%	10%
FWER-control test rejects on...	-	-	-
Diff-in-diffs setup	NO	NO	NO
School-level controls	NO	NO	NO
Control for non-compliance	NO	NO	NO
Control for individual characteristics	NO	NO	NO
Control for social desirability	NO	NO	NO
Control for ethnicity	NO	NO	NO
Control for teacher quality in Grade 5-8	NO	NO	NO
No. of obs	1173	1173	1173

Linear regressions. Dependent variables: the summary index of cognitive outcomes in the first column, the summary index of non-cognitive index in the second column and the summary index of inter-ethnic attitudes in the third columns, measured in Grade 8. FDR and FWER-control multiple testing procedures are used to test 12 hypotheses together in Tables 3.9, C.6, 3.10, and C.7.

Table C.7: Effects on Roma and Non-Roma Students

	Index of cognitive outcomes	Index of non-cognitive outcomes	Index of tolerance measures
	Model 1		
Program*non-Roma interaction	-0.067	-0.067	0.182
CL robust SE	0.137	0.084	0.076
Uncorrected p-values	0.628	0.425	0.020
FDR-control test rejects on...	-	-	5%
FWER-control test rejects on...	-	-	-
Diff-in-diffs setup	NO	NO	NO
School-level controls	NO	NO	NO
Control for non-compliance	NO	NO	NO
Control for individual characteristics	NO	NO	NO
Control for social desirability	NO	NO	NO
Control for ethnicity	NO	NO	NO
Control for teacher quality in Grade 5-8	NO	NO	NO
No. of obs	1173	1173	1173

Linear regressions. Dependent variables: the summary index of cognitive outcomes in the first column, the summary index of non-cognitive index in the second column and the summary index of inter-ethnic attitudes in the third columns, measured in Grade 8. FDR and FWER-control multiple testing procedures are used to test 12 hypotheses together in Tables 3.9, C.6, 3.10 and C.7.

Table C.8: Impact heterogeneity with respect to ethnicity-Reading scores and secondary school acceptance

	Reading score	Secondary school acceptance
	Model 1	
Program*non-Roma interaction	-0.092	-0.030
Cl robust SE	0.176	0.092
Uncorrected p-values	0.606	0.749
Diff-in-diffs setup	NO	NO
School-level controls	NO	NO
Control for non-compliance	NO	NO
Control for individual characteristics	NO	NO
Control for social desirability	NO	NO
Control for ethnicity	NO	NO
Control for teacher quality in Grade 5-8	NO	NO
No. of obs	1140	954

Linear regressions. Dependent variables: standardized reading test score in Grade 8 in the first column, and the probability of starting an academic high school degree in the second column.

Table C.9: Multiple testing of the impact's channel - school-level regression

	Index of cognitive outcomes	Index of non-cognitive outcomes	Index of tolerance measures
		Model 1	
Program	0.093	0.142	0.117
Robust SE	0.065	0.051	0.051
P-values	0.157	0.007	0.026
		Model 2	
Program	0.090	0.052	0.125
Robust SE	0.066	0.053	0.070
P-values	0.186	0.334	0.085
No. of obs	58	58	58

C.3 Appendix

This section tests the ATET effects of the OOIH Demonstration Program on school level data using small sample inference methods.

Small Sample Inference

Considering the empirical distributions of the summary index in the treated and the control group of schools, our school level sample should be more than 10% larger in order to be able to reject the null hypothesis of $\beta = 0$ on a 5% significance level, conditional on that the alternative hypothesis is true ($\beta \neq 0$), with a 90% probability. In other words, based on the empirical distribution of our outcome variable, to be able to conduct a t-test with 90% power, sample size of both groups should be at least 34 schools each. (REF) Moreover, in addition to testing power, results of conventional tests based on asymptotic theory may be misleading due to this low sample size. To support that our results still hold even if we do not rely on asymptotic assumptions, in addition to regression-based t-tests, we conduct two randomization-based exact testing procedures:

- Fisher's exact p-values (Imbens and Wooldridge, 2009)
- Bertrand, Duflo and Mullainathan (2002) - BDM-method

Both procedures have the following advantages:

- they do not rely on large sample assumptions (exact), work irrespective of sample size

- both use the empirical distribution of estimated effects for placebo laws
- they don't rely on any assumptions regarding the error term

They are similar procedures but test different hypotheses:

- Fisher's: whether the program had any effects?
- BDM: whether the program has a non-zero average effect?

Fisher's exact p-values

This procedure compares real, observed average outcome differences between the treated and the control group to an empirical distribution of simulated outcomes "caused" by randomly assigned placebo treatments assigned to random halves of 60-school sample. Let's have the following hypotheses:

- H_0: the program has no effect for any school in the sample:

$Y_s(1) = Y_s(0)$, $\forall s = 1, ..., S$

- H_a: the program has a nonzero effect for at least one school:

$\exists s$ such that $Y_s(1) \neq Y_s(0)$

Then, we form the following test statistics:

$T(W, Y) = \bar{Y}_1 - \bar{Y}_0$

We make $n = 400$ draws of placebo treatments. Under H_0 we can deduce the value of the test statistic from each round:

$T(\tilde{W}, Y) = \tilde{Y}_1 - \tilde{Y}_0$

Then we compare the value of the balance statistics to the actual observed difference of treated-control outcomes. The p-value of the statistic is going to be the probability that the simulated statistic is at least as large, in absolute value, as the observed difference. In this framework, we know the exact values of all the missing potential outcomes under H_0. Thus, there are no nuisance parameters to estimate, and as a result, we can simply deduce the distribution of any statistic. In short, we simply calculate exact p-values as the ratio of the number of the cases when the placebo difference is at least as large as the observed difference, and the number of draws.

The BDM (2002) Method

The second procedure we perform is based on Bertrand et al. (2002) and it goes as follows. Again, we compare the difference of average outcomes in the treated and control groups to empirical distribution of simulated such differences "caused" by randomly assigned placebo treatments. Again, placebo treatments are randomly assigned to schools to maintain the clustered nature of the data. We perform 400 random draws of placebo effect and we real difference is out of the 95$(or90$) of the empirical distribution of the difference under the placebo treatments. Formally, we have the following hypotheses:

- H_0: the program has no effect on average:

$\bar{Y}_1 - \bar{Y}_0 = 0$

- H_A: the program has a non-zero average effect effect:

$\bar{Y}_1 - \bar{Y}_0 \neq 0$

- We reject the null if $\bar{Y}_1 - \bar{Y}_0$ is outside of the 95$(or90$) of the simulated distribution of differences under placebo rules $\bar{Y^*}_1 - \bar{Y^*}_0$:

$\bar{Y}_1 - \bar{Y}_0$ <cutoff($p_{0.025}$)

$\bar{Y}_1 - \bar{Y}_0$ >cutoff($p_{0.975}$)

The Estimation of School Level Effects Using Small Sample Inference

Table C.10: School-Level Estimates of Overall Effects

	Model 1	Model 2
Program	0.090	0.075
Robust SE	0.029	0.035
P-values	0.003	0.044
Fisher's exact p-values	0.003	0.073
BDM test rejects on 5%	+	-
BDM test reject on 10%	+	+
Diff-in-diffs setup	NO	YES
School-level controls	NO	YES
Control for always takers	NO	YES
Control for school-average of individual characteristics	NO	YES
Control for school-average social desirability	NO	YES
Share of Roma students - school average	NO	YES
No. of obs	60	58

Linear regressions. Dependent variable: summary index of cognitive, non-cognitive and inter-ethnic attitude measures, measured in Grade 8.