Economics of Education Review 97 (2023) 102487

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Longer schooling with grade retention: The effects of increasing the school leaving age on dropping out and labour market success

Anna Adamecz

Institute of Economics Centre for Economic and Regional Studies (HUN-REN KRTK KTI), 1097 Budapest, Tóth Kálmán utca 4 and UCL Social Research Institute, University College London (UCL), 27 Woburn Square London WC1H 0AA

ARTICLE INFO

JEL classification:
I2

Keywords:
Compulsory school leaving age
Differences-in-regression-discontinuities
Dropping out

ABSTRACT

This paper examines the effects of increasing the compulsory school leaving age from 16 to 18 in Hungary using a difference-in-regression-discontinuities design empirical strategy. While the reform increased the length of schooling, it did not decrease the probability of dropping out of secondary school, either on average or among the most at-risk group of Roma ethnic minority students. Due to grade retentions, marginal students were older than their peers and couldn't have reached the final grade of secondary school by age 18 to earn a degree. The reform also did not affect the probability of employment, hours worked, wages and the probability of working in low-skilled occupations at ages 20 and 25. In education systems that allow grade retention, compulsory education should have the explicit goal of keeping students in school until they earn a secondary degree, rather than just until a certain age.

1. Introduction

Compulsory school leaving (CSL) age policies constrain decision-making about the time and effort individuals invest in attending school. A substantial body of literature has investigated the effects of increasing the CSL age on various social and economic outcomes or has used such reforms as instrumental variables for education[1]. However, the evidence of this literature is mixed. In some cases, for example, increasing the CSL age resulted in positive wage returns (Oreopoulos, 2007; Devereux & Hart, 2010), while in other cases, no wage returns were found (Oosterbeek & Webbink, 2007; Pischke & von Wachter, 2008). Grenet (2013) attempted to explore why longer schooling did not necessarily increase labour market success by comparing the wage effects of similar reforms in Britain and France. He found that due to the reform, school completion increased sharply in Britain but not in France, hence raising the CSL age brought positive wage returns only in Britain. Increasing the CSL age might also have unintended negative effects, such as reducing the effort that teachers put into teaching (Green & Paniagua, 2012) and increasing the criminal behaviour of students within schools (Anderson, Hansen & Walker, 2013).

This paper investigates the effects of increasing the CSL age from 16 to 18 in Hungary on schooling and labour market outcomes. The Hungarian reform offers important lessons for three reasons. First, it was implemented in an education system with early ability tracking, strong selection mechanisms and substantial between-school differences (OECD, 2015). Raising the CSL age in such an environment may have different effects than in countries like the US, the UK and France, which have been primarily examined by the existing literature. Second, the Hungarian system allows for grade retention. By the time students reach the end of elementary school in Grade 8, more than 15 % of them are already at least 16 years old. Therefore, even if they are required to stay in school until age 18, some might still not complete Grade 12, which is necessary to earn a secondary degree. Third, the Hungarian data capture student characteristics and enable the presentation of descriptive evidence on how the distribution of students changed in schools. Understanding what happened in schools after the reform may enhance our comprehension of why CSL age reforms succeed in one context but not in another.

I identify the causal effects of the reform using a difference-in-regression-discontinuities design (DRDD) strategy (Grembi, Nannicini & Troiano, 2016; Hong, Dragan & Glied, 2019) that exploits the elementary school enrolment rule. While compliance with the enrolment rule is not perfect, there is a jump in the probability of being exposed to the reform around a cutoff date of birth. However, being born

E-mail address: a.adamecz-volgyi@ucl.ac.uk.

[1] Effects on wages: Meghir and Palme, 2005; Grenet, 2013; mortality: Lleras-Muney, 2005; fertility: Black, Devereux and Salvanes, 2011; criminal behaviour: Anderson, 2014; Lochner and Moretti, 2004; voting behaviour: Milligan, Moretti and Oreopoulos, 2004.

https://doi.org/10.1016/j.econedurev.2023.102487
Received 21 November 2021; Received in revised form 23 October 2023; Accepted 23 October 2023
Available online 31 October 2023

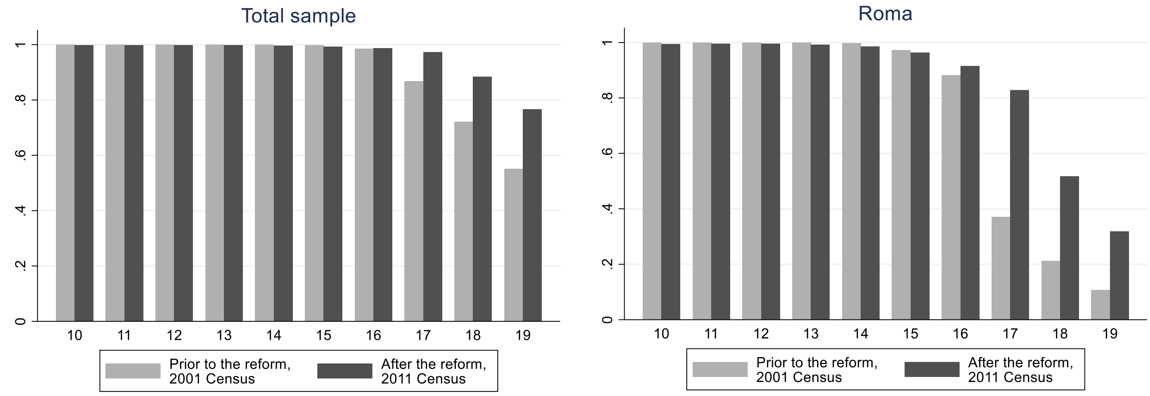

Fig. 1. The share of those still in school before and after the reform, by age.

Table 1
Comparison cohorts and the construction of the DRDD estimates.

	Outcomes at age 20	Outcomes at age 25
Reform cohort	Those born in 1991 (aged 20 in 2011)	Those born in 1991 (aged 25 in 2016)
Contemporary comparison cohorts	Those born in 1990 (aged 21 in 2011) Those born in 1992 (aged 19 in 2011)	Those born in 1990 (aged 26 in 2016) Those born in 1992 (aged 24 in 2016)
Past comparison cohorts	Those born in 1980-1982 (aged 19-21 in 2001 in the Census and aged 29-31 in 2011 in Admin3)	Those born in 1975-77 (aged 24-26 in 2001 in the Census and aged 34-36 in 2016 in Admin3)
DRDD estimates	(RDD_1991-RDD_1981)-[(RDD_1990-RDD_1980)+(RDD_1992-RDD_1982)]/2	(RDD_1991-RDD_1976)-[(RDD_1990-RDD_1975)+(RDD_1992-RDD_1977)]/2

Note that while the Census data were collected in 2001, 2011 and 2016, and Admin3 database covers the period between 2003 and 2017. Thus, when using the Admin3 data, I cannot look at the outcome variables of the past comparison cohorts in 2001, at the same age as the reform and the contemporary comparison cohorts. Hence, I'll look at the outcome variables of the past comparison cohorts at the same time when the reform cohort, either in 2011 or 2016. As the results will show, the effect of the reform on employment will be similarly zero using both data sources/strategies; consequently, this choice should not affect the conclusions.

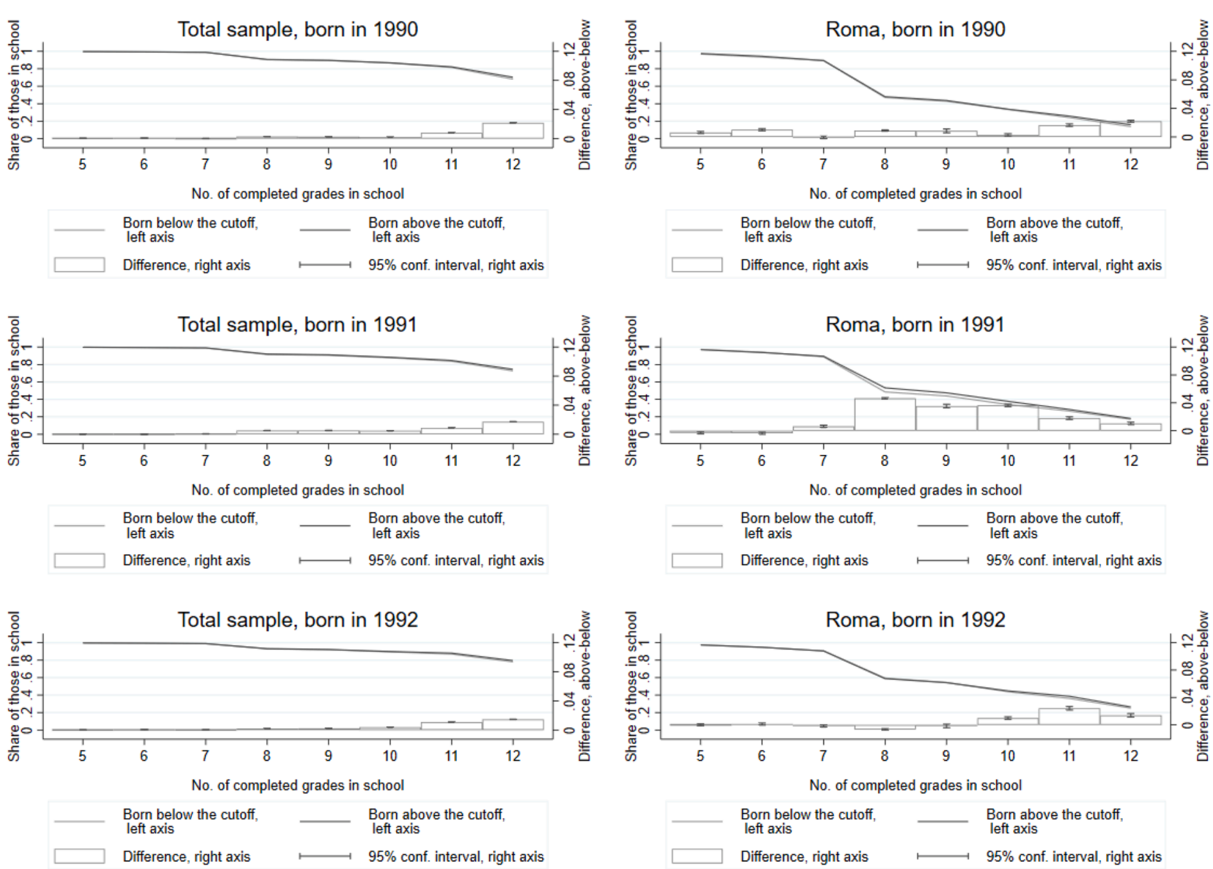

Fig. 2. The probability of still being in school after completing 5-12 grades in school.

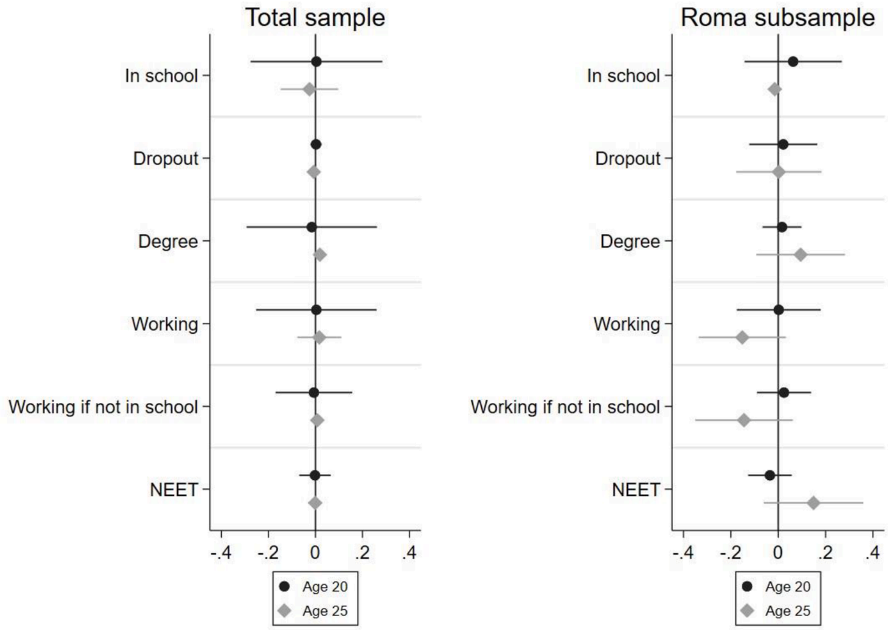

Fig. 3. The effects of the reform on schooling and labour market outcomes (Census data).

before or after the enrolment cutoff, or in other words, starting elementary school at a younger or older age, might impact schooling outcomes without a reform as well. Thus, I estimate the differences in schooling and labour market outcomes around the cutoff of the reform cohort and the cutoffs of several comparison cohorts and interpret the difference of these regression discontinuity estimates as the intention to treat (ITT) effect of the reform.

Using the Hungarian Census as well as administrative data on employment and wages at ages 20 and 25, I find that the reform had no effects on the schooling and early labour market outcomes of the first treated cohort. While it increased the length of schooling on average by half a year, as marginal students were older than their peers due to grade retention, the reform did not keep potential dropouts long enough in school to help them to earn a secondary degree. Neither did the reform have any effect on the probability of employment, hours worked, hourly wages, and the probability of working in a low-skilled occupation. These results suggest that staying in school longer without earning a degree was not enough to increase the employment prospects, and very likely, the human capital of students.

Dropping out of school is an especially huge problem for Roma ethnic minority young people. While on average, one in ten students dropped out of school at age 17 before the reform, this ratio was over 60 % among Roma students. I find that although Roma students became more likely to complete Grade 9 and 10, the reform did not increase their probability to complete Grade 12 and earn a secondary degree. In education systems that allow grade retention, such reforms should explicitly aim for keeping students in school until they earn a degree rather than just until a certain age.

Fig. 4. The effects of the reform on employment, hours worked, wages and the probability of working in low-skilled occupations (Total sample, Admin 3 data, DRDD coefficients).

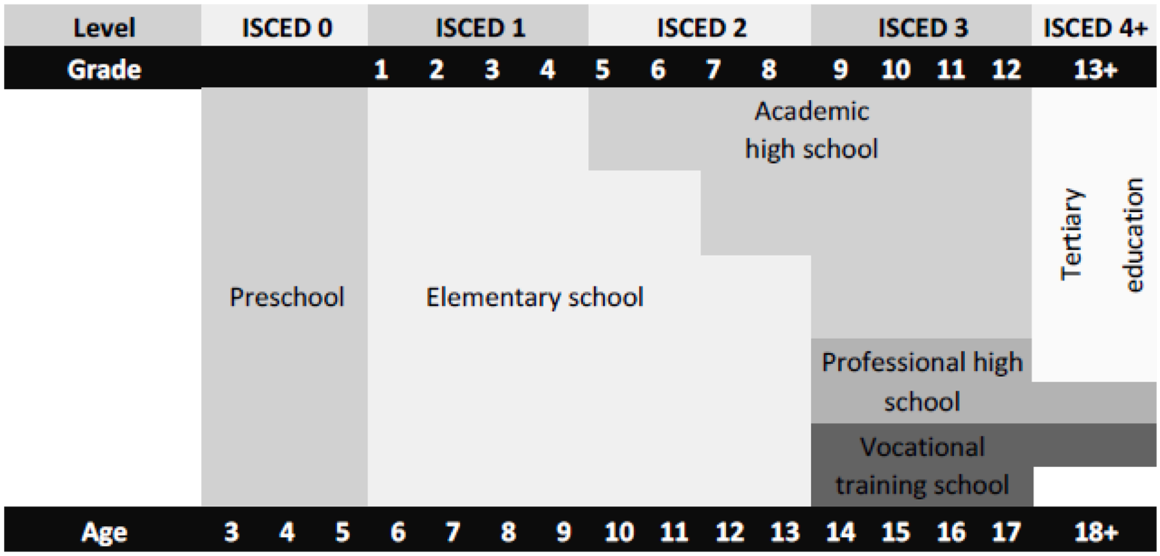

Fig. A1. The structure of the Hungarian education system
Source: Horn, 2014. For the cohort of interest, preschool was compulsory from age 5. The CSL age was 16 for those starting elementary school in September 1997 or earlier. The CSL age was 18 for those starting elementary school in September 1998 or later.

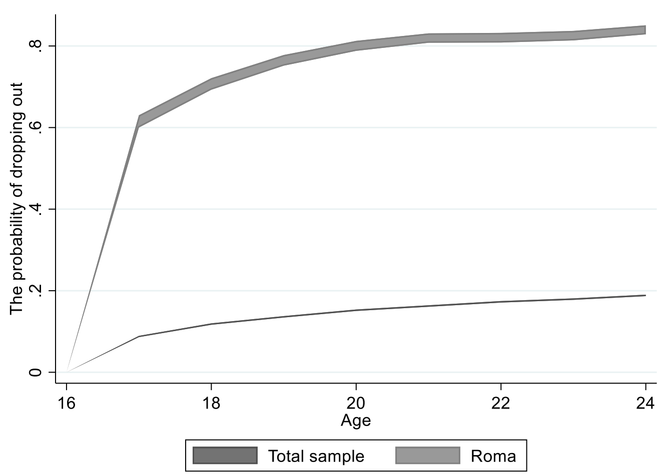

Fig. A2. The probability of dropping out of school without earning a degree by age before the reform
Source: own estimation from the 2001 Hungarian Census. No. of observations: 1,561,429 and 45,212.

While the reform could have affected secondary school choice as it was known since the first treated cohort enrolled in elementary school, it did not affect secondary school enrolment. About 22 % of students, including those most at-risk of dropping out, attended vocational training schools. The data suggest that as lower-standing, lower-performing students stayed in vocational training schools longer, the distribution of students shifted to the left and potentially crowded out resources in vocational training schools. I find that the probability of dropping out might have even increased in vocational training schools, especially among Roma students. While this latter effect is large in magnitude at around 10 percentage points, it is not significant due to the small sample size.

The last takeaway from this analysis is that increasing the CSL age may not always be a good instrumental variable (IV) for education. It violates the monotonicity assumption of the instrument if the reform negatively affects the quality of education for some students (Cygan-Rehm & Maeder, 2013). The monotonicity assumption requires students to be impacted by the instrument in the same direction (Angrist & Pischke, 2008). In this context, it would assume that the reform induced at least some individuals to have more education and no one to have less education, both in terms of length and quality. This assumption would be violated as the reform decreased the quality of teaching in vocational training schools.

The remainder of the paper unfolds as follows. Section 2 introduces the Hungarian education system and the reform, while Section 3 presents the data sources. Section 4 details the identification strategy and the empirical methods and Section 5 shows the main results along with their robustness checks. Section 6 presents suggestive evidence on the potential mechanisms behind the findings and Section 7 summarizes and discusses the results.

2. The Hungarian education system and the reform

2.1. Compulsory education in Hungary

The Hungarian education system has long faced challenges in providing high-quality education for students of differing backgrounds (OECD, 2015). Free elementary school choice and early tracking have hindered equity and have caused a high variance of student achievements between schools. This has been a long-standing problem that is still not resolved. In the 2012 Program for International Student Assessment (PISA) study, average math test scores were similar to those of the United States, somewhat below the OECD average (OECD, 2014).

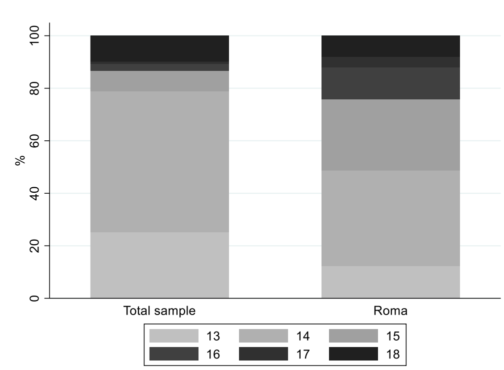

Fig. A3. The distribution of students in Grade 8 by age before the reform
Source: own estimation from the 2001 Hungarian Census. No. of observations: 127,035 and 3,758.

Table A1
Compliance with the school enrolment rule in aggregate administrative data

	Early starters	Compliers	Late starters	No. of students in grade 1	CSL age
Compliance by academic years, share of those starting elementary school at the given year					
1997/1998	0.02	0.80	0.18	127,214	16
1998/1999	0.02	0.78	0.20	125,875	18
1999/2000	0.01	0.78	0.21	121,424	18
Compliance by cohorts, share of cohort size					
Last cohort before the reform (born between June 90-May 91)	0.02	0.79	0.19	129,489	16
First cohort after the reform (born between June 91-May 92)	0.02	0.78	0.20	126,294	18

Source: Public Education Statistics (PES) of the Public Education Information System. *Early starters:* those who enrol in elementary school earlier than expected based on the enrolment rule. *Compliers:* those who enrol in elementary school according to the enrolment rule. *Late starters:* those who enrol in elementary school later than expected based on the enrolment rule. Individual-level data are not available for this period.

Table A2
The definition of outcome variables

Outcome variable	Definition	Unit of measurement
A. Outcome variables from the Censuses (all measured either in October 2011 (age 20 outcomes) or October 2016 (age 25 outcomes))		
In school	Attends school at the time of observation	binary variable
Dropout	Did not earn a secondary degree and does not attend school at the time of observation.	binary variable
Degree	Earned a secondary degree.	binary variable
Works	Works (is employed) at the time of observation; based on self-reporting.	binary variable
Works if not in school	Works (is employed) at the time of observation, conditional on not attending school.	binary variable
NEET	Neither in school nor in employment.	binary variable
Secondary school enrolment		
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
High school	Finished at least one academic year in a (professional or academic) high school.	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 observation.	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 observation.	binary variable
High school	Finished at least one academic year in a high school but did not earn any secondary degree and was not in school at the time of observation.	binary variable
B. Outcome variables from the administrative data (Admin3) (all measured either in October 2011 (age 20 outcomes) or October 2016 (age 25 outcomes))		
Employed	Officially employed according to the tax register	binary variable
Hours worked	Contractual working hours	continuous variable
Log hourly wage	Official monthly wage according to the tax register divided by contractual working hours.	continuous variable
Works in a low-skilled occupation	ISCO Major Group 9: Elementary occupations	binary variable
Disappeared from administrative data	Equals 1 if someone became invisible in the administrative data, i.e., they did not die, did not work, did not get registered as unemployed, did not go to school and did not see a family doctor (GP) between January 2009 and the time of observation for at least one consecutive year. Once someone disappeared, the variable stays 1.	binary variable

Source: own collection from the 2011 Census, the 2016 Microcensus and the Admin3 database.

The effects of socioeconomic background on test scores, however, were among the largest in Hungary, suggesting that the education system does a poor job in counterbalancing social inequalities.

About 10 % of school-aged young people belong to the Roma ethnic minority, which is the largest ethnic minority in Hungary. Being Roma is highly correlated with poverty, social exclusion, long term unemployment, and access to low-quality public services, including health care and education (Ladányi & Szelényi, 2002; Kemény & Janky, 2005; Kertesi, 2005; Gábos et al., 2006). The Roma-non Roma gap in standardized test scores at the end of elementary school (Grade 8) is around one standard deviation, similar to the Black-White test score gap in the 1980s in the US (Kertesi & Kézdi, 2011).

For most students, elementary school starts at age 6 or 7 and has eight grades (Grade 1-8). Most secondary schools have four grades[2] (Grade 9-12) and end with the possibility of earning a secondary degree. There are two main tracks in secondary school: the vocational training school track and the high-school track (Fig. A1 in Appendix A). The core programs in all tracks lasted for four years at that time, both right before and right after the reform. Admission to secondary schools is merit-based, and there is a strong sorting of students across the tracks. In terms of ability, on average, vocational training school students lag behind high school students by almost a full standard deviation in their math and reading test scores in Grade 8 (Hermann, 2013).

Table A3
Grade retention in grades 1-4, % of students in grade

Academic year	Grade 1	Grade 2	Grade 3	Grade 4	CSL age
1995/1996	4.0	1.9	1.6	1.7	16
1996/1997	3.9	2.0	1.5	1.6	16
1997/1998	3.9	1.9	1.5	1.7	16
1998/1999	4.0	1.8	1.5	1.6	18
1999/2000	3.9	1.9	1.4	1.5	18

Source: National Institute of Public Education, 2006. Table 4.28 in the Appendix, page 478.

[2] Some highly selective elite academic high schools recruit top-talent students already in Grade 4 and Grade 6, but these students are not likely to be affected by a CSL age reform.

Table A4
Descriptive statistics: school enrolment at age 7

	Year of birth	Census 2001					
		Born below the cutoff			Born above the cutoff		
		Obs	Mean	SD	Obs	Mean	SD
	Total sample						
Starts school at age 7	1990	50658	0.38	0.49	52647	0.93	0.25
Starts school at age 7	1991	52208	0.39	0.49	52269	0.93	0.25
Starts school at age 7	1992	49675	0.4	0.49	50406	0.93	0.25
	Roma subsample						
Starts school at age 7	1990	1989	0.7	0.46	1923	0.95	0.22
Starts school at age 7	1991	1966	0.7	0.46	1930	0.95	0.22
Starts school at age 7	1992	1933	0.67	0.47	1956	0.94	0.23

Source: own estimation from the 2001 Hungarian Census. The average probability of starting school at age 7. Born below the cutoff: born 150 days (5 months) before 1 June. Born above the cutoff: born 150 days (5 months) after 1 June.

Table A5
Descriptive statistics: sample balance and outcome variables at age 20 by year of birth (Census data, total sample)

	Census 2001						Census 2011					
	Born below the cutoff			Born above the cutoff			Born below the cutoff			Born above the cutoff		
	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD
	1980						1990					
Year of birth	63834	1980	0	63618	1980	0	49694	1990	0	51936	1990	0
Month of birth	63834	3	1.42	63618	7.97	1.4	49694	3.01	1.42	51936	8	1.4
Roma	63834	0.03	0.17	63618	0.03	0.16	47667	0.05	0.22	49957	0.05	0.22
Female	63834	0.49	0.5	63618	0.49	0.5	49694	0.49	0.5	51936	0.49	0.5
Age	63834	20.21	0.4	63618	20	0	49694	21	0	51936	20.8	0.4
In school	63834	0.4	0.49	63618	0.44	0.5	49694	0.49	0.5	51936	0.56	0.5
Dropout	63834	0.04	0.19	63618	0.04	0.2	49694	0.14	0.34	51936	0.13	0.34
Degree	63834	0.8	0.4	63618	0.78	0.41	49694	0.84	0.37	51936	0.83	0.38
Works	63834	0.45	0.5	63618	0.4	0.49	49694	0.38	0.48	51936	0.31	0.46
Works of not in school	38395	0.68	0.47	35470	0.65	0.48	25519	0.62	0.49	23074	0.58	0.49
NEET	63834	0.19	0.4	63618	0.19	0.39	49694	0.24	0.42	51936	0.22	0.42
	1981						1991					
Year of birth	61317	1981	0	60533	1981	0	51892	1991	0	52670	1991	0
Month of birth	61317	3.03	1.41	60533	7.94	1.41	51892	3.02	1.42	52670	7.98	1.4
Roma	61317	0.03	0.17	60533	0.03	0.16	49909	0.05	0.22	50711	0.05	0.22
Female	61317	0.49	0.5	60533	0.48	0.5	51892	0.49	0.5	52670	0.49	0.5
Age	61317	19.2	0.4	60533	19	0	51892	20	0	52670	19.81	0.39
In school	61317	0.51	0.5	60533	0.58	0.49	51892	0.61	0.49	52670	0.69	0.46
Dropout	61317	0.03	0.17	60533	0.03	0.17	51892	0.13	0.33	52670	0.12	0.33
Degree	61317	0.77	0.42	60533	0.72	0.45	51892	0.81	0.39	52670	0.78	0.41
Works	61317	0.32	0.47	60533	0.26	0.44	51892	0.26	0.44	52670	0.19	0.39
Works of not in school	30292	0.6	0.49	25645	0.57	0.49	20184	0.54	0.5	16461	0.49	0.5
NEET	61317	0.2	0.4	60533	0.18	0.39	51892	0.21	0.41	52670	0.19	0.39
	1982						1992					
Year of birth	56766	1982	0	57240	1982	0	49950	1992	0	51795	1992	0
Month of birth	56766	3	1.42	57240	7.97	1.4	49950	2.98	1.41	51795	7.97	1.4
Roma	56766	0.03	0.17	57240	0.03	0.16	48001	0.05	0.22	49822	0.05	0.22
Female	56766	0.49	0.5	57240	0.48	0.5	49950	0.48	0.5	51795	0.49	0.5
Age	56766	18.2	0.4	57240	18	0	49950	19	0	51795	18.81	0.39
In school	56766	0.64	0.48	57240	0.75	0.44	49950	0.76	0.43	51795	0.82	0.38
Dropout	56766	0.02	0.15	57240	0.02	0.14	49950	0.11	0.31	51795	0.1	0.31
Degree	56766	0.64	0.48	57240	0.36	0.48	49950	0.68	0.47	51795	0.52	0.5
Works	56766	0.2	0.4	57240	0.13	0.34	49950	0.14	0.35	51795	0.09	0.28
Works of not in school	20182	0.51	0.5	14545	0.47	0.5	12071	0.45	0.5	9214	0.38	0.49
NEET	56766	0.18	0.38	57240	0.13	0.34	49950	0.16	0.36	51795	0.13	0.34

Source: own estimation from the 2001 and 2011 Hungarian Censuses. The average probability of starting school at age 7. Born below the cutoff: born 150 days (5 months) before 1 June. Born above the cutoff: born 150 days (5 months) after 1 June.

Most students completed elementary school even before the reform, and the gap in completing elementary school between Roma and non-Roma students is negligible (Hajdú, Kertesi & Kézdi, 2014). On average, 80 % of students earned at least a secondary degree, while among the Roma, this ratio was less than 20 % (Fig. A2 in Appendix A). The share of dropouts was small at age 16 (as one had to complete the academic year in which they turned 16), but by age 17, more than 60 %, and by age 18, almost 70 % of Roma students dropped out.

As already mentioned, the Hungarian system allows grade retention. Students who are not able to fulfil the requirements of a grade would stay in that grade for one more year (or even more, until they fulfil the requirements). Thus, by the time they reach the end of elementary school in Grade 8, serial grade repeaters could be 2-4 years older than their peers. On average, 20 % of Grade 8 students were already at least 15 years old before the reform, while among the Roma, this ratio was around 50 % (Fig. A3 in Appendix A). Note that practically, in terms of keeping them in school until earning a degree at the end of Grade 12, the reform could only be binding for students who were at most 14 years old in Grade 8; otherwise, they could still drop out at age 18 before earning a degree.

Table A6
Descriptive statistics: sample balance and outcome variables at age 20 (Census data, Roma subsample)

	Census 2001						Census 2011					
	Born below the cutoff			Born above the cutoff			Born below the cutoff			Born above the cutoff		
	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD
	1980						1990					
Year of birth	1796	1980	0	1737	1980	0	2483	1990	0	2539	1990	0
Month of birth	1796	2.95	1.44	1737	7.92	1.4	2483	2.96	1.42	2539	8.02	1.38
Roma	1796	1	0	1737	1	0	2483	1	0	2539	1	0
Female	1796	0.48	0.5	1737	0.49	0.5	2483	0.49	0.5	2539	0.5	0.5
Age	1796	20.22	0.42	1737	20	0	2483	21	0	2539	20.8	0.4
In school	1796	0.05	0.22	1737	0.06	0.24	2483	0.1	0.3	2539	0.13	0.33
Dropout	1796	0.05	0.22	1737	0.06	0.24	2483	0.67	0.47	2539	0.67	0.47
Degree	1796	0.18	0.38	1737	0.15	0.36	2483	0.29	0.46	2539	0.29	0.45
Works	1796	0.44	0.5	1737	0.42	0.49	2483	0.46	0.5	2539	0.42	0.49
Works of not in school	1702	0.45	0.5	1633	0.44	0.5	2241	0.5	0.5	2216	0.47	0.5
NEET	1796	0.52	0.5	1737	0.53	0.5	2483	0.68	0.47	2539	0.67	0.47
	1981						1991					
Year of birth	1789	1981	0	1577	1981	0	2471	1991	0	2542	1991	0
Month of birth	1789	2.99	1.39	1577	8.01	1.44	2471	3.02	1.42	2542	8.02	1.41
Roma	1789	1	0	1577	1	0	2471	1	0	2542	1	0
Female	1789	0.49	0.5	1577	0.49	0.5	2471	0.49	0.5	2542	0.48	0.5
Age	1789	19.19	0.39	1577	19	0	2471	20	0	2542	19.79	0.4
In school	1789	0.09	0.29	1577	0.11	0.31	2471	0.17	0.37	2542	0.2	0.4
Dropout	1789	0.05	0.21	1577	0.04	0.21	2471	0.67	0.47	2542	0.63	0.48
Degree	1789	0.18	0.38	1577	0.17	0.37	2471	0.25	0.44	2542	0.26	0.44
Works	1789	0.39	0.49	1577	0.38	0.49	2471	0.4	0.49	2542	0.35	0.48
Works of not in school	1620	0.42	0.49	1409	0.42	0.49	2061	0.47	0.5	2024	0.43	0.49
NEET	1789	0.52	0.5	1577	0.51	0.5	2471	0.65	0.48	2542	0.65	0.48
	1982						1992					
Year of birth	1709	1982	0	1592	1982	0	2545	1992	0	2638	1992	0
Month of birth	1709	3.06	1.41	1592	7.95	1.44	2545	2.98	1.42	2638	7.92	1.4
Roma	1709	1	0	1592	1	0	2545	1	0	2638	1	0
Female	1709	0.47	0.5	1592	0.47	0.5	2545	0.49	0.5	2638	0.5	0.5
Age	1709	18.19	0.39	1592	18	0	2545	19	0	2638	18.82	0.39
In school	1709	0.17	0.38	1592	0.22	0.41	2545	0.32	0.47	2638	0.37	0.48
Dropout	1709	0.04	0.19	1592	0.03	0.18	2545	0.58	0.49	2638	0.55	0.5
Degree	1709	0.16	0.36	1592	0.12	0.32	2545	0.2	0.4	2638	0.16	0.37
Works	1709	0.31	0.46	1592	0.28	0.45	2545	0.29	0.45	2638	0.25	0.43
Works of not in school	1416	0.37	0.48	1245	0.35	0.48	1734	0.4	0.49	1654	0.37	0.48
NEET	1709	0.53	0.5	1592	0.51	0.5	2545	0.57	0.49	2638	0.55	0.5

Source: own estimation from the 2001 and 2011 Hungarian Censuses. The average probability of starting school at age 7. Born below the cutoff: born 150 days (5 months) before 1 June. Born above the cutoff: born 150 days (5 months) after 1 June.

2.2. The reform

In an attempt to reduce the number of dropouts, the CSL age was increased from 16 to 18 in 1996. As mentioned above, students had to attend school until the end of the academic year in which they turned 16 before the reform. The Public Education Act (1996) increased compulsory school attendance until the end of the academic year in which students turned 18. The reform was introduced with students enrolling in elementary school in the 1998/99 academic year. Thus, these students (and their parents) knew already at the time of enrolment that they had to stay in school two years longer than the previous cohorts. The enforcement of the new regulation was strict (Mártonfi, 2011b), although there might have been some noncompliance, i.e. if one had a child or got married. Should children have missed school for an extended period, parents would have been fined or imprisoned for up to five years (Kazuska, 2012). Mártonfi (2011b) documents that the majority of students did show up in school as intended. Fig. 1 shows the share of young people still in school before the reform, in the 2001 Hungarian Census, and after the reform, in the 2011 Hungarian Census[3]. The share of 17-year-olds in school on average was already high before the reform, at 87 %, and by 2011, this ratio increased to 97 %. Among Roma students, the share of those in school at age 17 increased from 38 % to 92 % between 2001 and 2011. The data suggest that the average length of schooling increased by about 0.48 years[4] (between age 17 and age 19) on average and by about 0.98 years among the Roma.

Although the Act introduced other measures as well, increasing the CSL age was the only element causing sharp changes for those starting elementary school right before the reform, in the 1997/98 academic year, versus right after the reform, 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 four grades (Grade 9-12) and thus not to end before age 18 for anybody). This process began in the 1998/1999 academic year. As a result, when the first treated cohort enrolled in secondary school at age 14 in 2006, the adjustments to the length of secondary school programs had been adapted for half a decade.

All actors of the education system supported the reform at the time of its enactment in 1996. However, it has been viewed controversially since the first treated cohort reached age 16. As it was grandfathered in, the Act pushed all implementation costs to the future government of 2008 when the number of students started to increase in schools (National Institute of Public Education, 2011). Although the number of potentially affected students could have been predicted well in advance, the education policy did not actively support the implementation of the reform in 2008-2011. The schools and their leading bodies began to realize that they lacked the tools to handle the emerging problems

[3] See more information about the data in Section 3.

[4] The sum of increases in the probability of still being in school at ages 17, 18 and 19.

Table A7
Descriptive statistics: sample balance and outcome variables at age 25 (Census data, total sample)

	Census 2001						Census 2016					
	Born below the cutoff			Born above the cutoff			Born below the cutoff			Born above the cutoff		
	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD
	1975						1990					
Year of birth	77897	1975	0	75463	1975	0	4058	1990	0	4097	1990	0
Month of birth	77897	3.03	1.42	75463	7.94	1.39	4058	3.03	1.42	4097	8.01	1.4
Roma	77897	0.02	0.15	75463	0.02	0.15	4058	0.05	0.04	4097	0.05	0.05
Female	77897	0.49	0.5	75463	0.49	0.5	4058	0.48	0.5	4097	0.48	0.5
Age	77897	25.2	0.4	75463	25	0	4058	26	0	4097	25.8	0.4
In school	77897	0.11	0.32	75463	0.12	0.33	4058	0.1	0.3	4097	0.13	0.33
Dropout	77897	0.04	0.2	75463	0.05	0.21	4058	0.15	0.36	4097	0.15	0.36
Degree	77897	0.78	0.41	75463	0.78	0.41	4058	0.85	0.36	4097	0.84	0.37
Works	77897	0.78	0.42	75463	0.77	0.42	4058	0.75	0.44	4097	0.73	0.44
Works of not in school	69160	0.8	0.4	66145	0.8	0.4	3650	0.78	0.41	3584	0.77	0.42
NEET	77897	0.18	0.38	75463	0.18	0.38	4058	0.2	0.4	4097	0.2	0.4
	1976						1991					
Year of birth	74000	1976	0	74097	1976	0	4027	1991	0	4113	1991	0
Month of birth	74000	3.02	1.42	74097	7.97	1.4	4027	3	1.42	4113	8.01	1.4
Roma	74000	0.03	0.16	74097	0.02	0.14	4027	0.05	0.05	4113	0.05	0.05
Female	74000	0.49	0.5	74097	0.49	0.5	4027	0.5	0.5	4113	0.48	0.5
Age	74000	24.2	0.4	74097	24	0	4027	25	0	4113	24.81	0.39
In school	74000	0.14	0.35	74097	0.16	0.37	4027	0.15	0.35	4113	0.18	0.38
Dropout	74000	0.04	0.2	74097	0.05	0.22	4027	0.14	0.35	4113	0.14	0.34
Degree	74000	0.79	0.41	74097	0.79	0.41	4027	0.86	0.35	4113	0.86	0.35
Works	74000	0.75	0.44	74097	0.73	0.44	4027	0.73	0.45	4113	0.71	0.46
Works of not in school	63353	0.78	0.41	62003	0.79	0.41	3435	0.79	0.41	3379	0.79	0.41
NEET	74000	0.18	0.39	74097	0.18	0.38	4027	0.18	0.39	4113	0.18	0.38
	1977						1992					
Year of birth	72158	1977	0	70848	1977	0	3981	1992	0	4142	1992	0
Month of birth	72158	3.05	1.42	70848	7.96	1.41	3981	3	1.42	4142	7.97	1.4
Roma	72158	0.03	0.16	70848	0.02	0.15	3981	0.05	0.04	4142	0.05	0.05
Female	72158	0.49	0.5	70848	0.49	0.5	3981	0.49	0.5	4142	0.48	0.5
Age	72158	23.19	0.4	70848	23	0	3981	24	0	4142	23.81	0.4
In school	72158	0.19	0.39	70848	0.22	0.41	3981	0.23	0.42	4142	0.26	0.44
Dropout	72158	0.04	0.2	70848	0.05	0.22	3981	0.15	0.36	4142	0.14	0.35
Degree	72158	0.8	0.4	70848	0.8	0.4	3981	0.84	0.36	4142	0.85	0.35
Works	72158	0.7	0.46	70848	0.67	0.47	3981	0.65	0.48	4142	0.65	0.48
Works of not in school	58458	0.77	0.42	55594	0.77	0.42	3060	0.75	0.43	3069	0.78	0.42
NEET	72158	0.18	0.39	70848	0.18	0.39	3981	0.19	0.39	4142	0.17	0.37

Source: own estimation from the 2001 and 2016 Hungarian Censuses. The average probability of starting school at age 7. Born below the cutoff: born 150 days (5 months) before 1 June. Born above the cutoff: born 150 days (5 months) after 1 June.

(National Institute of Public Education, 2011). The increased number of students put so much strain on unprepared schools (primarily vocational training schools, that most students at risk of dropping out attended) that most school principals viewed the reform unfavourably, according to a 2009 survey (Mártonfi, 2011b). The most frequently expressed problems included that schools had no methods to engage unmotivated students in learning, they were unable to offer a credible perspective on life to these mostly low socioeconomic status and low-skilled students, and they had no expertise in the development of students from troubled backgrounds (Mártonfi, 2011a). Due to these challenges 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 evaluates the raising of the CSL age only[5].

As already mentioned, the reform was introduced with those who enrolled in elementary school in 1998. According to the elementary school enrolment rule, in this period, compulsory schooling started on 1 September of the year in which one reached age 6 by 31 May. Those born on 1 June or later, start elementary school one year later, at age 7. Thus, those compliant with the enrolment rule and born before 1 June 1991, enrolled in elementary school in 1997 under the old CSL age scheme. Those compliant with the enrolment rule and born on 1 June 1991, or later, enrolled in school in 1998 under the new CSL age scheme. This discontinuity at 1 June 1991 in the date of birth of students is going to be the base of my identification strategy.

Compliance with the enrolment rule was not perfect. The age of enrolment was a joint decision of parents, preschool teachers, and in some cases, pedagogical and psychological counsellors employed by public pedagogical service centres. The decision itself was made during preschool. At the time of the reform, preschool attendance was compulsory from age 5. The decision process about elementary school enrolment started with an official opinion of preschool teachers about whether the child was ready to start school. In the case of any doubts, preschool teachers could ask for a "school readiness examination" from the local pedagogical service centre. According to the aggregate administrative data (the statistics of the Public Education Statistics database of the Public Education Information System), about 80 % of a cohort enrolled in elementary school according to the enrolment rule in this period (*compliers*), while most of the rest enrolled a year later (*late starters*) (Table A1 in Appendix A). Early school start is rare (*early starters*), at about 2 %. More detailed information on compliance with the enrolment rule is provided along with the identification assumptions in Section 4.

3. Data

This paper relies on six data sources: (1) the 2001 Hungarian Census, (2) the 2011 Hungarian Census, (3) the 2016 Hungarian Microcensus, (4) the Panel of Linked Administrative Data (Admin3) database, (5) aggregate administrative data from the Public Education Statistics (PES) of the Public Education Information System, and (6) the National

[5] On the reform that cut the CSL age back from 18 to 16, see Adamecz-Völgyi *et al.* (2021).

Table A8
Descriptive statistics: sample balance and outcome variables at age 25 (Census data, Roma subsample)

	Census 2001						Census 2016					
	Born below the cutoff			Born above the cutoff			Born below the cutoff			Born above the cutoff		
	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD
	1975						1990					
Year of birth	1898	1975	0	1714	1975	0	201	1990	0	206	1990	0
Month of birth	1898	3.05	1.4	1714	8	1.42	201	2.9	1.43	206	8.02	1.47
Roma	1898	1	0	1714	1	0	201	1	0	206	1	0
Female	1898	0.49	0.5	1714	0.47	0.5	201	0.47	0.5	206	0.53	0.5
Age	1898	25.19	0.39	1714	25	0	201	26	0	206	25.77	0.42
In school	1898	0.02	0.12	1714	0.03	0.16	201	0.02	0.14	206	0.02	0.15
Dropout	1898	0.04	0.2	1714	0.04	0.21	201	0.63	0.48	206	0.69	0.46
Degree	1898	0.13	0.34	1714	0.16	0.37	201	0.36	0.48	206	0.29	0.45
Works	1898	0.46	0.5	1714	0.51	0.5	201	0.51	0.5	206	0.45	0.5
Works of not in school	1869	0.46	0.5	1671	0.51	0.5	197	0.51	0.5	201	0.44	0.5
NEET	1898	0.53	0.5	1714	0.48	0.5	201	0.48	0.5	206	0.54	0.5
	1976						1991					
Year of birth	1897	1976	0	1561	1976	0	199	1991	0	195	1991	0
Month of birth	1897	2.96	1.41	1561	8	1.43	199	3.09	1.49	195	8.02	1.37
Roma	1897	1	0	1561	1	0	199	1	0	195	1	0
Female	1897	0.5	0.5	1561	0.51	0.5	199	0.49	0.5	195	0.5	0.5
Age	1897	24.21	0.41	1561	24	0	199	25	0	195	24.84	0.37
In school	1897	0.02	0.13	1561	0.02	0.15	199	0.01	0.07	195	0.04	0.19
Dropout	1897	0.03	0.18	1561	0.04	0.2	199	0.74	0.44	195	0.71	0.45
Degree	1897	0.15	0.36	1561	0.16	0.36	199	0.26	0.44	195	0.28	0.45
Works	1897	0.52	0.5	1561	0.52	0.5	199	0.52	0.5	195	0.48	0.5
Works of not in school	1866	0.52	0.5	1526	0.52	0.5	198	0.53	0.5	188	0.48	0.5
NEET	1897	0.47	0.5	1561	0.47	0.5	199	0.47	0.5	195	0.5	0.5
	1977						1992					
Year of birth	1831	1977	0	1633	1977	0	218	1992	0	220	1992	0
Month of birth	1831	2.99	1.41	1633	7.92	1.39	218	3.09	1.37	220	8	1.37
Roma	1831	1	0	1633	1	0	218	1	0	220	1	0
Female	1831	0.49	0.5	1633	0.49	0.5	218	0.5	0.5	220	0.4	0.49
Age	1831	23.21	0.41	1633	23	0	218	24	0	220	23.82	0.39
In school	1831	0.02	0.13	1633	0.02	0.14	218	0	0.07	220	0.03	0.18
Dropout	1831	0.04	0.19	1633	0.05	0.21	218	0.73	0.44	220	0.67	0.47
Degree	1831	0.16	0.37	1633	0.16	0.37	218	0.27	0.44	220	0.32	0.47
Works	1831	0.51	0.5	1633	0.5	0.5	218	0.4	0.49	220	0.54	0.5
Works of not in school	1798	0.51	0.5	1598	0.5	0.5	217	0.4	0.49	213	0.55	0.5
NEET	1831	0.48	0.5	1633	0.49	0.5	218	0.6	0.49	220	0.43	0.5

Source: own estimation from the 2001 and 2016 Hungarian Censuses. The average probability of starting school at age 7. Born below the cutoff: born 150 days (5 months) before 1 June. Born above the cutoff: born 150 days (5 months) after 1 June.

Table A9
Descriptive statistics: secondary school enrolment and dropping out by school tracks at age 20 (Census data, total sample)

	Census 2001						Census 2011					
	Born below the cutoff			Born above the cutoff			Born below the cutoff			Born above the cutoff		
	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD
	1980						1990					
Goes to secondary school	63834	0.86	0.35	63618	0.86	0.35	49694	0.9	0.3	51936	0.91	0.29
Goes to vocational training school	63834	0.34	0.47	63618	0.33	0.47	44844	0.26	0.44	47034	0.24	0.43
Goes to high school	63834	0.6	0.49	63618	0.61	0.49	44844	0.97	0.18	47034	0.96	0.19
Drops out from secondary school	54586	0.04	0.21	54561	0.05	0.22	49694	0.14	0.34	51936	0.13	0.34
Drops out from vocational training school	21616	0.08	0.28	20870	0.1	0.3	11632	0.11	0.31	11235	0.12	0.32
Drops out from high school	38585	0.02	0.14	38930	0.02	0.14	43348	0.02	0.13	45312	0.02	0.14
	1981						1991					
Goes to secondary school	61317	0.86	0.35	60533	0.86	0.35	51892	0.91	0.28	52670	0.92	0.27
Goes to vocational training school	61317	0.32	0.46	60533	0.29	0.46	47385	0.25	0.43	48377	0.23	0.42
Goes to high school	61317	0.62	0.49	60533	0.63	0.48	47385	0.96	0.2	48377	0.95	0.22
Drops out from secondary school	52528	0.03	0.18	51934	0.03	0.18	51892	0.13	0.33	52670	0.12	0.33
Drops out from vocational training school	19377	0.06	0.23	17756	0.06	0.24	11634	0.12	0.32	11172	0.13	0.34
Drops out from high school	37836	0.02	0.14	38005	0.02	0.13	45373	0.02	0.14	45837	0.02	0.14
	1982						1992					
Goes to secondary school	63834	0.86	0.35	63618	0.86	0.35	49950	0.93	0.26	51795	0.93	0.25
Goes to vocational training school	63834	0.34	0.47	63618	0.33	0.47	46365	0.24	0.43	48180	0.23	0.42
Goes to high school	63834	0.6	0.49	63618	0.61	0.49	46365	0.93	0.26	48180	0.91	0.29
Drops out from secondary school	54586	0.04	0.21	54561	0.05	0.22	49950	0.11	0.31	51795	0.1	0.31
Drops out from vocational training school	21616	0.08	0.28	20870	0.1	0.3	11284	0.12	0.32	11009	0.12	0.32
Drops out from high school	38585	0.02	0.14	38930	0.02	0.14	43042	0.02	0.14	43607	0.02	0.13

Source: own estimation from the 2001 and 2011 Hungarian Censuses. The average probability of starting school at age 7. Born below the cutoff: born 150 days (5 months) before 1 June. Born above the cutoff: born 150 days (5 months) after 1 June.

Table A10
Descriptive statistics: secondary school enrolment and dropping out by school tracks at age 20 (Census data, Roma subsample)

	Census 2001						Census 2011					
	Born below the cutoff			Born above the cutoff			Born below the cutoff			Born above the cutoff		
	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD
	1980						1990					
Goes to secondary school	1796	0.23	0.42	1737	0.23	0.42	2483	0.47	0.5	2539	0.48	0.5
Goes to vocational training school	1796	0.19	0.39	1737	0.19	0.39	1161	0.67	0.47	1214	0.63	0.48
Goes to high school	1796	0.07	0.25	1737	0.06	0.24	1161	0.81	0.39	1214	0.8	0.4
Drops out from secondary school	422	0.21	0.41	400	0.27	0.44	2483	0.67	0.47	2539	0.67	0.47
Drops out from vocational training school	333	0.23	0.42	323	0.28	0.45	779	0.34	0.47	770	0.36	0.48
Drops out from high school	121	0.11	0.31	106	0.11	0.32	943	0.17	0.37	971	0.17	0.38
	1981						1991					
Goes to secondary school	1789	0.26	0.44	1577	0.25	0.43	2471	0.48	0.5	2542	0.53	0.5
Goes to vocational training school	1789	0.19	0.39	1577	0.19	0.39	1192	0.64	0.48	1345	0.64	0.48
Goes to high school	1789	0.08	0.27	1577	0.07	0.26	1192	0.79	0.4	1345	0.76	0.43
Drops out from secondary school	464	0.19	0.39	393	0.18	0.38	2471	0.67	0.47	2542	0.63	0.48
Drops out from vocational training school	344	0.18	0.38	293	0.19	0.39	766	0.38	0.48	860	0.37	0.48
Drops out from high school	146	0.16	0.37	114	0.09	0.28	946	0.2	0.4	1025	0.18	0.39
	1982						1992					
Goes to secondary school	1709	0.28	0.45	1592	0.28	0.45	2545	0.59	0.49	2638	0.58	0.49
Goes to vocational training school	1709	0.2	0.4	1592	0.19	0.4	1497	0.64	0.48	1536	0.61	0.49
Goes to high school	1709	0.09	0.29	1592	0.08	0.28	1497	0.74	0.44	1536	0.75	0.43
Drops out from secondary school	475	0.13	0.34	442	0.12	0.33	2545	0.58	0.49	2638	0.55	0.5
Drops out from vocational training school	334	0.14	0.35	310	0.14	0.35	957	0.34	0.48	941	0.28	0.45
Drops out from high school	153	0.06	0.24	135	0.07	0.26	1104	0.16	0.37	1152	0.15	0.36

Source: own estimation from the 2001 and 2011 Hungarian Censuses. The average probability of starting school at age 7. Born below the cutoff: born 150 days (5 months) before 1 June. Born above the cutoff: born 150 days (5 months) after 1 June.

Table A11
Descriptive statistics: sample balance and outcome variables in October 2011 by year of birth (Admin3 data, total sample)

	Born below the cutoff			Born above the cutoff			Born below the cutoff			Born above the cutoff		
	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD
	1980						1990					
Year of birth	30868	1980	0.00	30340	1980	0.00	24286	1990	0.00	25102	1990	0.00
Month of birth	30868	3	1.41	30340	8	1.41	24286	3	1.42	25102	8	1.40
Age	30868	31	0.00	30340	31	0.00	24286	21	0.00	25102	21	0.00
Female	30868	0.49	0.50	30340	0.50	0.50	24286	0.49	0.50	25102	0.48	0.50
Employed	30868	0.56	0.50	30340	0.57	0.50	24286	0.30	0.46	25102	0.27	0.44
Hours worked	15507	38.09	6.03	15516	38.11	6.03	5788	36.10	8.53	4801	35.67	8.84
Log hourly wage	15507	6.81	0.62	15516	6.81	0.61	5788	6.42	0.40	4801	6.38	0.40
Low-skilled jobs	15507	0.09	0.29	15516	0.09	0.29	5788	0.14	0.35	4801	0.16	0.37
Disappeared	30868	0.06	0.23	30340	0.05	0.22	24286	0.02	0.15	25102	0.02	0.15
	1981						1991					
Year of birth	29472	1981	0.00	29049	1981	0.00	25284	1991	0.00	25453	1991	0.00
Month of birth	29472	3	1.41	29049	8	1.41	25284	3	1.42	25453	8	1.41
Age	29472	30	0.00	29049	30	0.00	25284	20	0.00	25453	20	0.00
Female	29472	0.50	0.50	29049	0.49	0.50	25284	0.48	0.50	25453	0.48	0.50
Employed	29472	0.56	0.50	29049	0.57	0.50	25284	0.24	0.43	25453	0.20	0.40
Hours worked	15009	38.09	6.01	14989	38.12	5.88	3878	35.61	9.04	2729	35.23	9.39
Log hourly wage	15009	6.80	0.61	14989	6.80	0.61	3878	6.37	0.39	2729	6.36	0.41
Low-skilled jobs	15009	0.09	0.28	14989	0.09	0.29	3878	0.14	0.35	2729	0.13	0.33
Disappeared	29472	0.05	0.22	29049	0.05	0.22	25284	0.02	0.14	25453	0.02	0.14
	1982						1992					
Year of birth	26556	1982	0.00	26887	1982	0.00	23818	1992	0.00	24603	1992	0.00
Month of birth	26556	3	1.42	26887	8	1.40	23818	3	1.42	24603	8	1.40
Age	26556	29	0.00	26887	29	0.00	23818	19	0.00	24603	19	0.00
Female	26556	0.49	0.50	26887	0.48	0.50	23818	0.48	0.50	24603	0.48	0.50
Employed	26556	0.57	0.50	26887	0.57	0.49	23818	0.19	0.39	24603	0.16	0.37
Hours worked	13727	38.13	5.87	14053	38.09	6.11	1781	35.23	9.29	1037	35.00	9.43
Log hourly wage	13727	6.80	0.60	14053	6.78	0.59	1781	6.31	0.44	1037	6.27	0.45
Low-skilled jobs	13727	0.09	0.29	14053	0.09	0.29	1781	0.11	0.31	1037	0.09	0.28
Disappeared	26556	0.05	0.22	26887	0.05	0.21	23818	0.02	0.14	24603	0.02	0.13

Source: own estimation from the Admin3 database. Born below the cutoff: born 5 months before 1 June. Born above the cutoff: born 5 months after 1 June.

Assessment of Basic Competencies (NABC) database.

The 2001 and the 2011 Censuses cover the total population with about 120,000-100,000 observations per year of birth. The data were collected in the spring of 2001 (autumn of 2011), when the main cohort of interest was about 10 (20) years old. The 2016 Microcensus was conducted in 2016 when the main cohort of interest was 25 years old and covers a 10 % representative sample of the population. The 2001 and 2016 data contain information on the birth year and month of individuals, while the 2011 data also contains the day of birth. All three Censuses have self-reported information on ethnicity (and no other data source captures ethnicity in Hungary).

The 2001 Census is the only single data source that allows to examine

Table A12
Descriptive statistics: sample balance and outcome variables in October 2016 by year of birth (Admin3 data, total sample)

	Born below the cutoff			Born above the cutoff			Born below the cutoff			Born above the cutoff		
	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD
	1975						1990					
Year of birth	41120	1975	0.00	40560	1975	0.00	24238	1990	0.00	25055	1990	0.00
Month of birth	41120	3	1.42	40560	8	1.40	24238	3	1.42	25055	8	1.40
Age	41120	41	0.00	40560	41	0.00	24238	26	0.00	25055	26	0.00
Female	41120	0.50	0.50	40560	0.50	0.50	24238	0.49	0.50	25055	0.48	0.50
Employed	41120	0.69	0.46	40560	0.69	0.46	24238	0.61	0.49	25055	0.61	0.49
Hours worked	24644	37.83	6.55	24300	37.73	6.59	13117	37.76	6.72	13360	37.66	6.85
Log hourly wage	24644	7.13	0.63	24300	7.14	0.63	13117	7.03	0.51	13360	7.01	0.51
Low-skilled jobs	24644	0.11	0.32	24300	0.11	0.32	13117	0.12	0.33	13360	0.12	0.33
Disappeared	41120	0.11	0.31	40560	0.11	0.32	24238	0.10	0.30	25055	0.09	0.29
	1976						1991					
Year of birth	38621	1976	0.00	38771	1976	0.00	25218	1991	0.00	25415	1991	0.00
Month of birth	38621	3	1.42	38771	8	1.39	25218	3	1.42	25415	8	1.40
Age	38621	40	0.00	38771	40	0.00	25218	25	0.00	25415	25	0.00
Female	38621	0.50	0.50	38771	0.49	0.50	25218	0.48	0.50	25415	0.48	0.50
Employed	38621	0.69	0.46	38771	0.69	0.46	25218	0.60	0.49	25415	0.58	0.49
Hours worked	23042	37.75	6.59	23018	37.71	6.68	13170	37.54	7.00	12849	37.41	7.19
Log hourly wage	23042	7.15	0.63	23018	7.15	0.63	13170	6.99	0.49	12849	6.97	0.47
Low-skilled jobs	23042	0.11	0.31	23018	0.11	0.31	13170	0.12	0.33	12849	0.13	0.34
Disappeared	38621	0.11	0.31	38771	0.12	0.32	25218	0.09	0.29	25415	0.09	0.29
	1977						1992					
Year of birth	37280	1977	0.00	36313	1977	0.00	23760	1992	0.00	24562	1992	0.00
Month of birth	37280	3	1.40	36313	8	1.41	23760	3	1.42	24562	8	1.40
Age	37280	39	0.00	36313	39	0.00	23760	24	0.00	24562	24	0.00
Female	37280	0.49	0.50	36313	0.50	0.50	23760	0.48	0.50	24562	0.48	0.50
Employed	37280	0.68	0.46	36313	0.68	0.46	23760	0.56	0.50	24562	0.53	0.50
Hours worked	22004	37.76	6.49	21436	37.80	6.54	11538	37.21	7.36	11213	37.00	7.61
Log hourly wage	22004	7.15	0.64	21436	7.15	0.63	11538	6.94	0.47	11213	6.91	0.46
Low-skilled jobs	22004	0.11	0.31	21436	0.11	0.31	11538	0.14	0.34	11213	0.14	0.35
Disappeared	37280	0.11	0.31	36313	0.11	0.32	23760	0.09	0.28	24562	0.08	0.28

Source: own estimation from the Admin3 database. Born below the cutoff: born 5 months before 1 June. Born above the cutoff: born 5 months after 1 June.

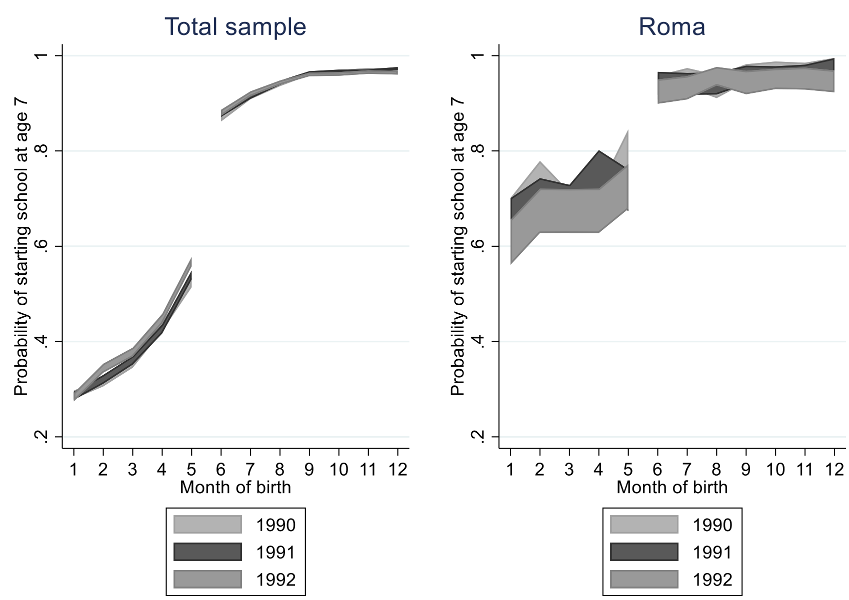

Fig. B1. The probability of starting school at age 7 among those born in 1990-1992
Source: own estimation from the 2001 Hungarian Census. The average probability of starting school at age 7 by month of birth, plotted with the 95% confidence intervals of the means. Those born in Jan-May: late-starters or grade repeaters. Those born in June-Dec: compliers to the elementary school enrolment rule. No. of observations: 370,344 and 14,147.

compliance with the enrolment rule for the cohorts of interest.[6] While it does not capture directly which academic year children started school, it registers which grade of elementary school students were attending at the time of the data collection. Thus, I infer the year of school enrolment from the grades they were attending in 2001. In terms of outcome variables, all three Censuses capture the following self-reported information: attending school by school track, highest degree, and whether one

[6] Individual-level administrative data on school enrolment are only available from 2009 in Hungary, when the first treated cohort was already 18 years old.

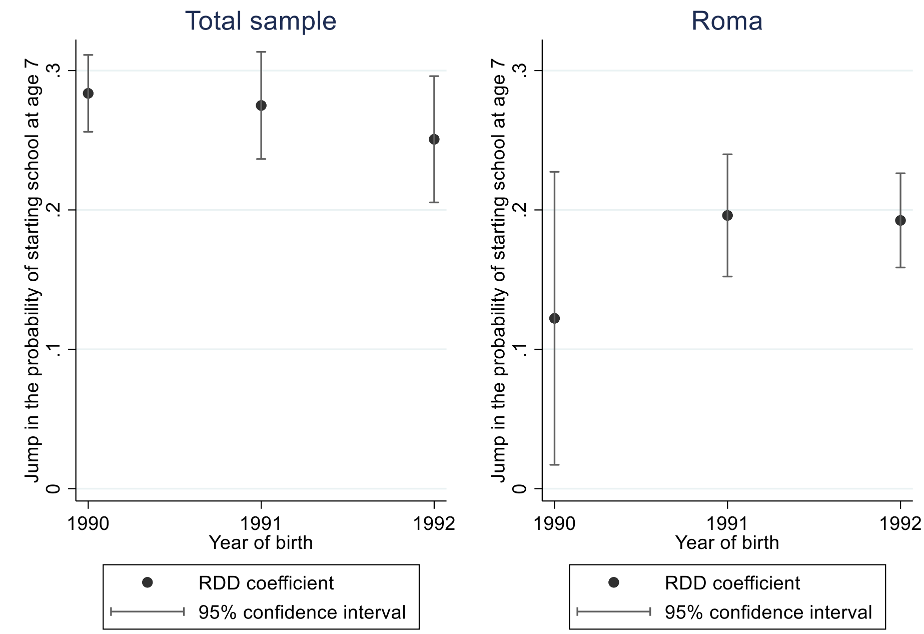

Fig. B2. The size of the jump (RDD coefficients) in the probability of starting school at age 7 among those born in 1990-1992
Source: own estimation from the 2001 Hungarian Census. RDD coefficients estimated using 150-day bandwidths. No. of observations: 276,770 and 10,459.

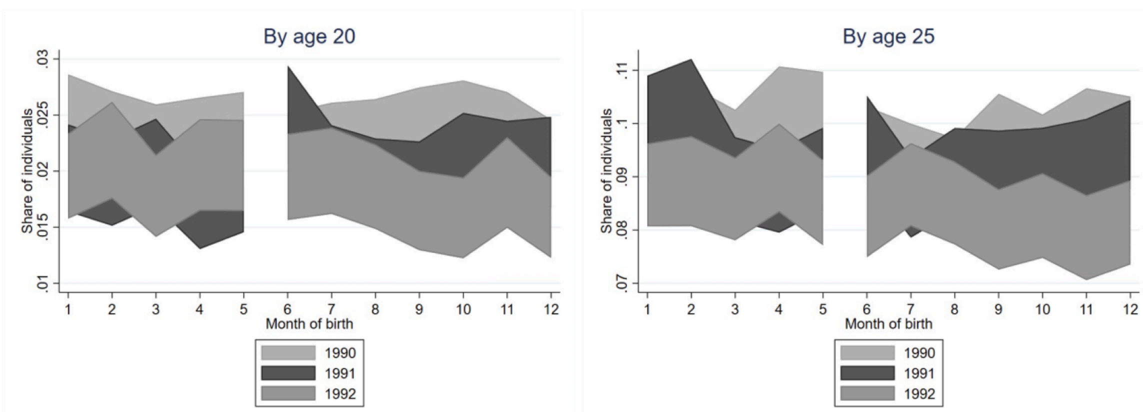

Fig. B3. The probability of disappearance from administrative data around the cutoff for those born in 1990-1992
Source: own estimation from the Admin3 database. The mean probability of disappearance from the data by month of birth, along with their 95% confidence intervals. Disappearence: not being registered as employed, unemployed, in school, no visit to family doctor for at least 12 consecutive months and never showing up again until October 2011 (by age 20) or 2016 (age 25); used as a proxy for migration. No. of observations: age 20: 176,879; age 25: 176,519.

has a job. The 2011 Census also captures the number of successfully completed grades. The Censuses do not capture information on grade retention.[7] All outcome variables from the Censuses are defined in block A of Table A2 in Appendix A, while the descriptive statistics of those born around the cutoff are provided in Table A4–Table A10 in Appendix A.

The Admin3 is an administrative panel database that is provided by the Databank of the Centre for Economic and Regional Studies (KRTK). The anonymized dataset links individual monthly data on employment, wage and occupation from the National Tax and Customs Administration, healthcare data (inpatient and outpatient care events) from the National Insurance Fund Administration, and school enrolment data from the Educational Authority (Sebők, 2019). It covers a random 50 % of the population born before 1 Jan 2003 (people with a Social Security Number in 2003). Labour market data are available from 2003 to 2017 while healthcare and education data are available only from 2009 to 2017. Unfortunately, the education data of Admin3 do not cover the period when the cohort of interests were in school, neither do they have information on ethnicity. Thus, I use these data to investigate the effects of the reform on employment, hours worked, log hourly wages and the probability of working in a low-skilled occupation (i.e., an occupation that does not require a qualification) at age 20 and 25, at the same times when the Census data were collected, on the total sample. Furthermore, I also test whether the reform affected the probability of “not being visible” in administrative data, i.e., when someone did not die, did not work, did not get registered as unemployed, did not go to school and did not see a family doctor (GP) by 2011 and 2016 at least once, as a proxy for leaving the country (migration). The outcome variables from Admin3 are defined in block B of Table A2 in Appendix A, while the descriptive statistics of those born around the cutoff are provided in Table A11 and Table A12 in Appendix A.

The PES of the Public Education Information System is the

[7] Descriptive statistics on grade retention are provided in Table A3 in Appendix A.

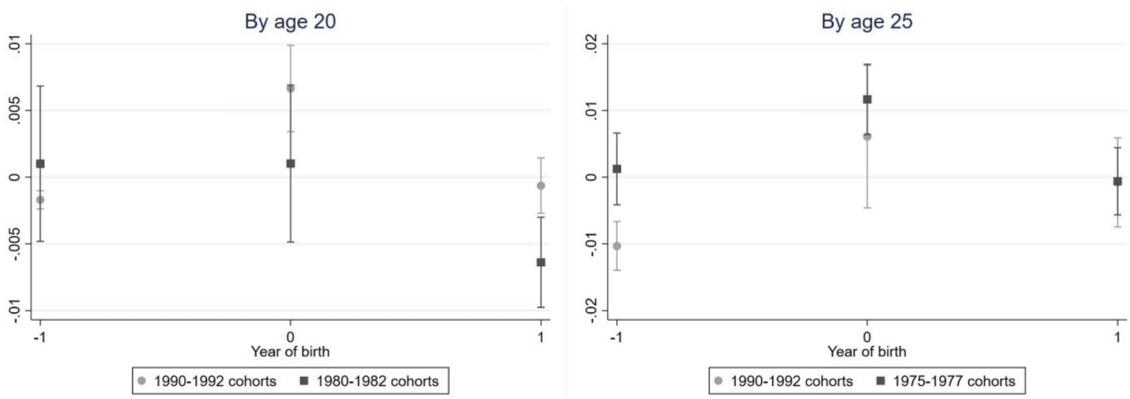

Fig. B4. The estimated RDD coefficients around the reform and comparison cutoffs (outcome: disappearance from administrative data)
Source: own estimation from the Admin3 database. Disappearence: not being registered as employed, unemployed, in school, no visit to family doctor for at least 12 consecutive months and never showing up again until October 2011 (by age 20) or 2016 (age 25); used as a proxy for migration. All coefficients are estimated in separate regressions. RDD coefficient estimates (β_{RDD}) according to Equation (1). Bandwidth: 5 months. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations: age 20: 288,768; age 25: 342,974.

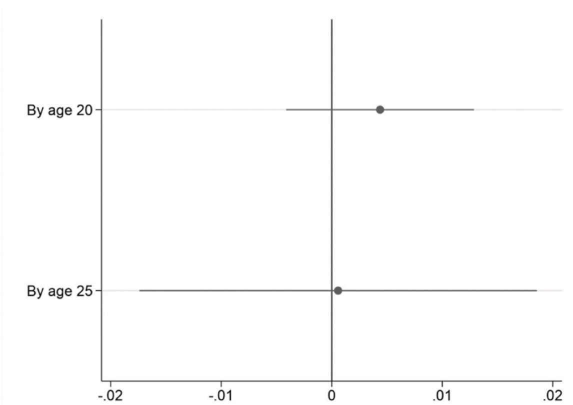

Fig. B5. The effect of the reform on the probability of disappearing from administrative data (DRDD estimates using a 5-month bandwidth)
Source: own estimation from Admin3 database. Disappearence: not being registered as employed, unemployed, in school, no visit to family doctor for at least 12 consecutive months and never showing up again until October 2011 (by age 20) or 2016 (age 25); used as a proxy for migration. All coefficients are estimated in separate regressions. DRDD coefficients estimates (β_{DRDD}) according to Equation (2). Bandwidth: 5 months. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations: age 20: 288,768; age 25: 342,974.

administrative school census. It collects information on schools, school programs, and students. For the cohorts and periods of interest for this paper, the PES provides aggregate data across school cohorts and academic years. I use the PES data as the second data source to show that compliance with the school enrolment rule was similar before and after the reform.

Lastly, the NABC registers the results of centrally organized low-stake math and reading tests taken in Grades 6, 8, and 10. The 10-grade waves of 2007-2010 cover the cohorts of interest. The NABC also provides information on grade retention, students' age, as well as school-level information on the distribution of students in schools. Thus, the NABC allows comparing the student body in secondary schools before and after the reform. The sampling design of the NABC changed in the observation period. The 2007 wave included 30 students from each programme of each school: 43,775 students altogether; however, the date of birth information is complete for only 36,605 students. From 2008 onwards, the sample includes all 10-graders (102,705-112,409 observations per year). Due to the change in the sampling design and the large share of missing information in the 2007 wave, the NABC is only used to provide supportive descriptive evidence, and I do not draw causal inference from the data.

4. Identification strategy and empirical methods

4.1. Identification strategy

As mentioned above, this paper uses a difference-in-regression-discontinuities design identification strategy (Grembi, Nannicini & Troiano, 2016; Hong, Dragan & Glied, 2019) that exploits the elementary school enrolment rule. Children born before 1 June should have, in principle, enrolled in elementary school at age 6, while those born on 1 June or later should have enrolled only the next year, aged around 7.[8] Thus, children were more likely to enrol in elementary school after the reform, in 1998, if they were born on 1 June 1991 or later, than those born before this date. However, being born before or after the enrolment cutoff also means starting elementary school at age 6 vs 7, which might impact schooling and labour market outcomes without reform as well. Furthermore, those who start school at age 7 spend one year less in school until the end of compulsory schooling.

Theoretically, starting school at age 7 (as opposed to age 6) could affect children's outcomes both positively and negatively. Black, Devereux 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. Besides, social development may depend on a child's age relative to their peers. On the other hand, starting school at an older age may be harmful if children could learn more in school than in preschool (or at home). Furthermore, parental investment in helping children with their schoolwork may depend on school starting age as well – parents may provide less help to their children if they start school when they are older. Black, Devereux and Salvanes (2011) examine the effect of school starting age on education outcomes, and they find very small positive effects of starting school

[8] Note that as one's enrolment age differs based on their month of birth, the expression *„starting school at age 6 vs 7"* is imprecise. Those compliant with the enrolment rule, if born before the cutoff, start school between ages 6.3 (born in May) - 6.7 (born in Jan), while those born above the cutoff, start school at ages between 6.8 (born in December) – 7.3 (born in June). At the cutoff, between those born in May vs. June, expected enrolment age goes up from 6.3 to 7.3. Those born below the cutoff, if delayed starting school by one year, will age between 7.3 (born in May) - 7.7 (born in Jan). For simplicity, I'll still refer to the enrolment age as "6" vs. "7".

when younger. In the Hungarian literature, Altwicker-Hámori and Köllő (2012) 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 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 the time of taking the test, as these two are perfectly collinear.

To identify the effect of the reform, and not the reform and starting school at age 7 together, I estimate the differences in the outcome variables around cutoffs among those not affected by the reform (comparison cohorts), and among those affected by the reform, born in 1991 (reform cohort), and interpret the difference of these RDD estimates as the ITT effect of the reform.

I look at the effects of the reform at age 20 (in 2011) and age 25 (in 2016) and use five comparison cohorts for both. Comparison cohorts are summarized in Table 1. As I measure the outcome variables in 2011 and in 2016, those born one year before or after the reform cohort are also one year older or younger and their outcome variables are not necessarily comparable. Thus, I use three further comparison cohorts, born 10 (15) years earlier, who were at the same age as those born +/- 1 year before/after the reform cohort when their outcomes were measured. The difference-in-regression-discontinuities estimates are constructed as the difference in the RDD coefficients between the treated cohort (1991) and the 10(15)-year-before cohort (1981 or 1976), compared to the average differences between the 1990 and 1992 cohorts and the 10(15)-year-before cohorts (1980/1982 and 1975/1977). As it will be detailed below, this construct relies on the assumption that in the lack of the reform, the estimated RDD jumps around the cutoffs would have

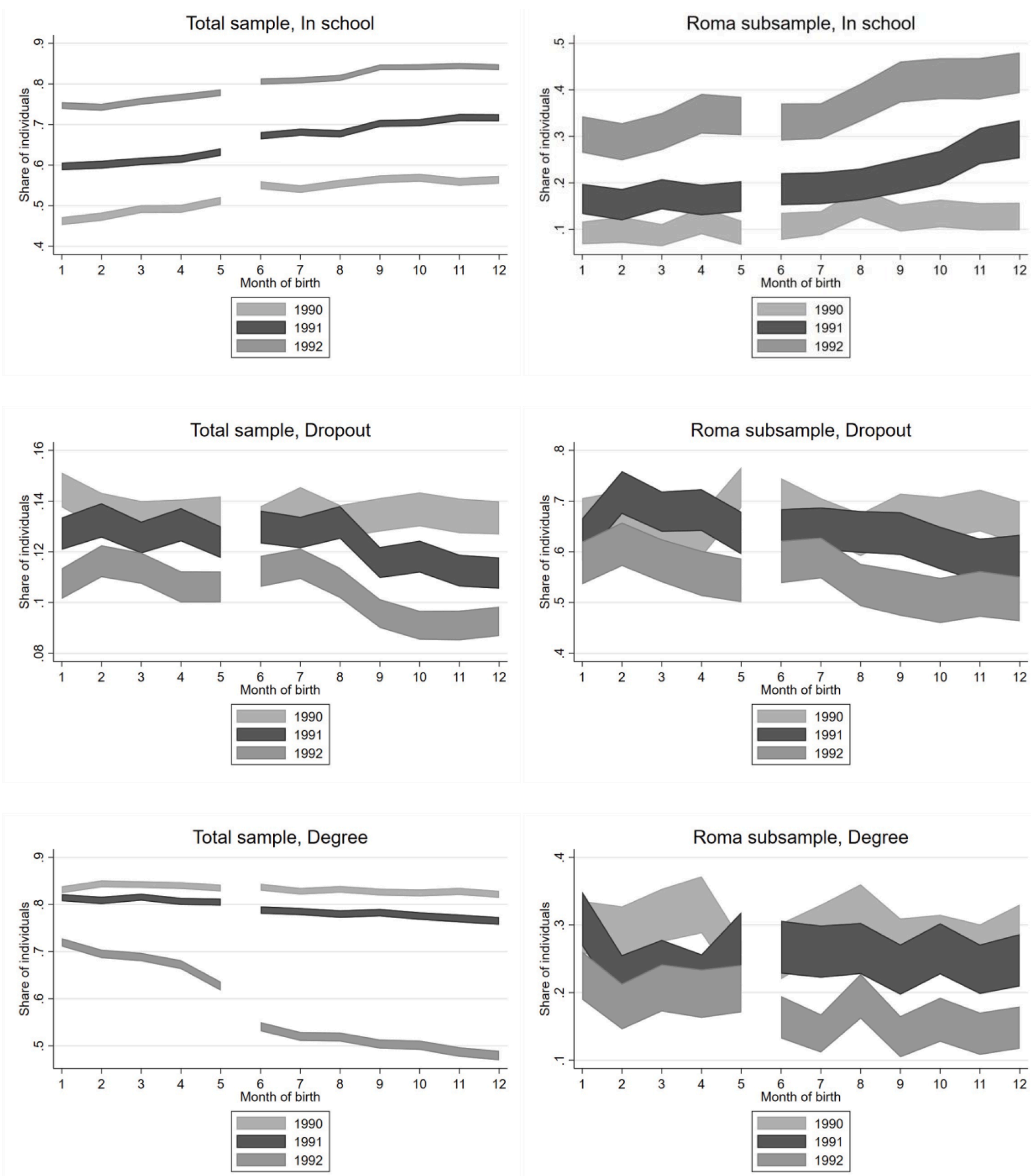

Fig. C1. The outcome variables around the cutoff for those born in 1990-1992, measured at age 20
Source: own estimation from the 2011 Census. The means of the outcome variables by month of birth, along with their 95% confidence intervals. No. of observations: Total sample: 366,608; Roma subsample: 18,176.

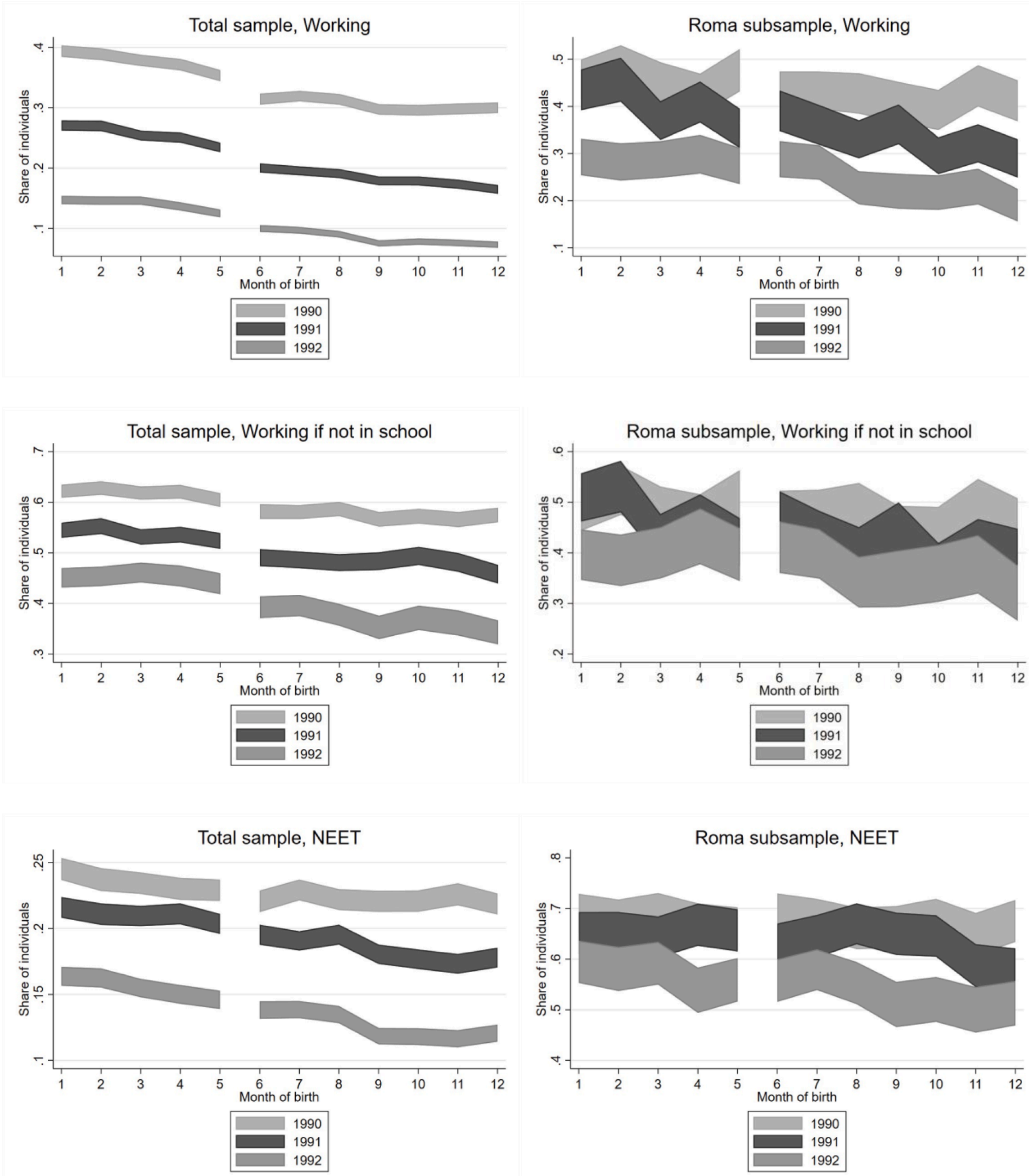

Fig. C1. (*continued*).

changed the same between the 1980/1982 and the 1990/1992 cohorts. Thus, any differential change in the RDD coefficients between the 1981 and 1991 cohorts (or the 1976 and 1991 cohorts) compared to the two other pairs of comparison cohorts must be due to the reform. This assumption will be formally tested in Section 5.

4.2. Identification assumptions

Based on Grembi, Nannicini and Troiano (2016), I rely on three identification assumptions. First, I assume that no manipulations happened regarding school enrolment due to the reform, i.e., no defiers exist. This assumption would fail if parents who dislike (like) schooling wanted to manipulate the system and sent their children born above (below) the cutoff to school one year earlier (later) to avoid (gain) two more compulsory years in school. I assume non-compliers start school late or early because of their general preferences on what is the ideal time of school enrolment, and not because they want to defy the reform. I provide evidence to support this assumption in subsection B1 of Appendix B by showing that the jump in the probability of starting school early vs. late around the cutoff stayed stable for the cohorts born in 1990–1992.

The second identification assumption, that was already mentioned in the previous section, is that the effects of being born above vs below the cutoff are either constant over time (in non-reform years) or are on the same trend. This assumption is similar to the parallel trends assumption of difference in differences strategies and is tested in the next section by plotting the same jumps (RDD coefficients) in the outcome variables around the cutoffs in five comparison birth years.

The third identification assumption requires the effect of the reform to be independent of date-of-birth effects. Practically, this means

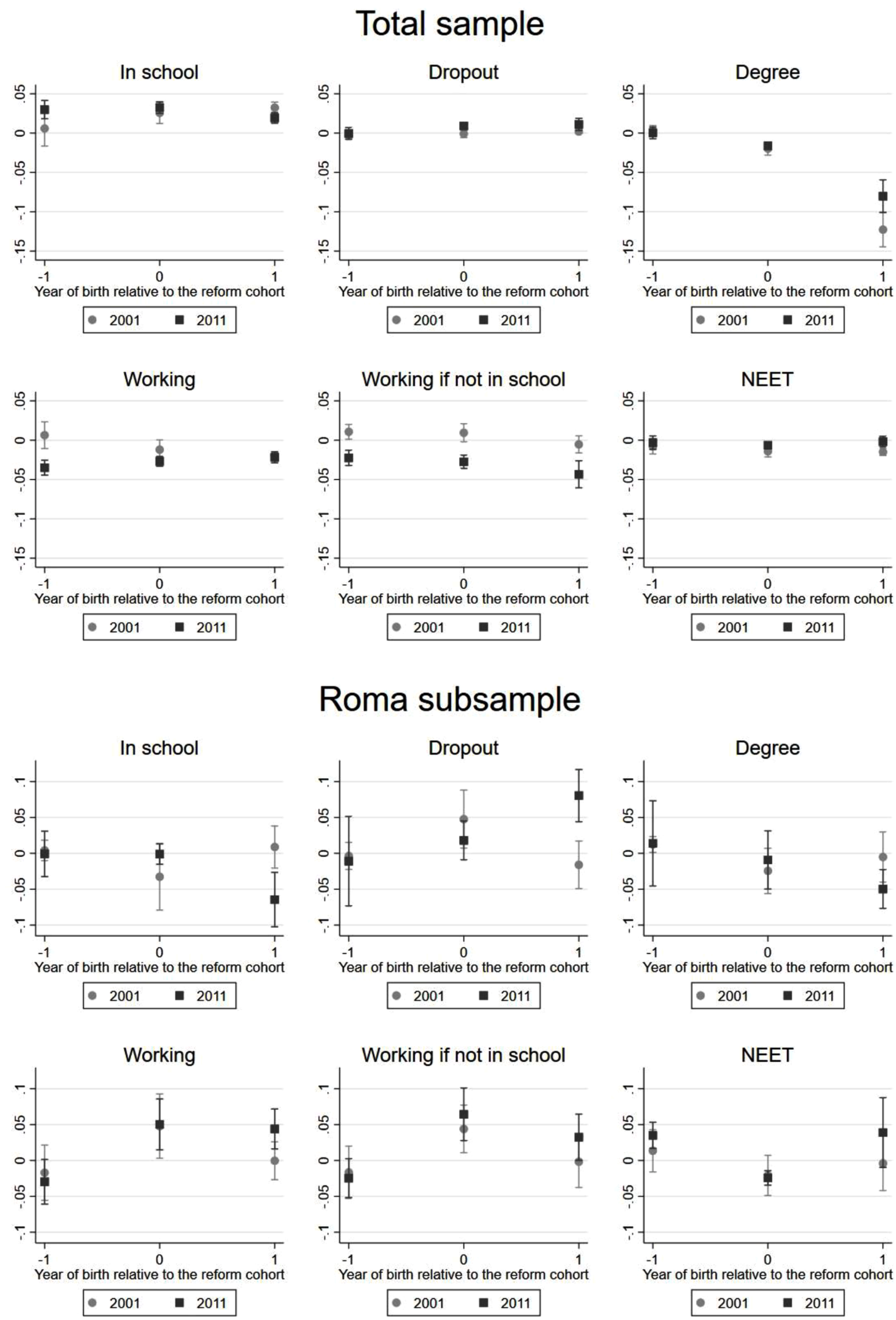

Fig. C2. The estimated RDD coefficients around the reform and comparison cutoffs (outcomes measured at age 20)
Source: own estimation from the 2001 and 2011 Censuses. RDD coefficients estimates (β_{RDD}) according to Equation (1). Bandwidth: 150 days in 2011 and 5 months in 2001. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations: 603,688 and 22,771.

assuming that I would get the same effects even if the cutoff were between different months (say, between October and November, as opposed to May and June). This assumption is inherently impossible to test, so all results are interpreted as local effects around the May-June cutoff.

Lastly, while this is not strictly an identification assumption, I also assume that the reform did not affect the probability of leaving the country (migration). If the reform increased the probability of migration and higher-ability young people became more likely to migrate due to the reform, I would underestimate the effects of the reform (because the data would cover only those who stayed in Hungary). Subsection B2 of Appendix B shows that the reform did not affect the probability that someone "disappeared" from the system, i.e., the probability that they did not die, did not work, not registered as unemployed, did not go to

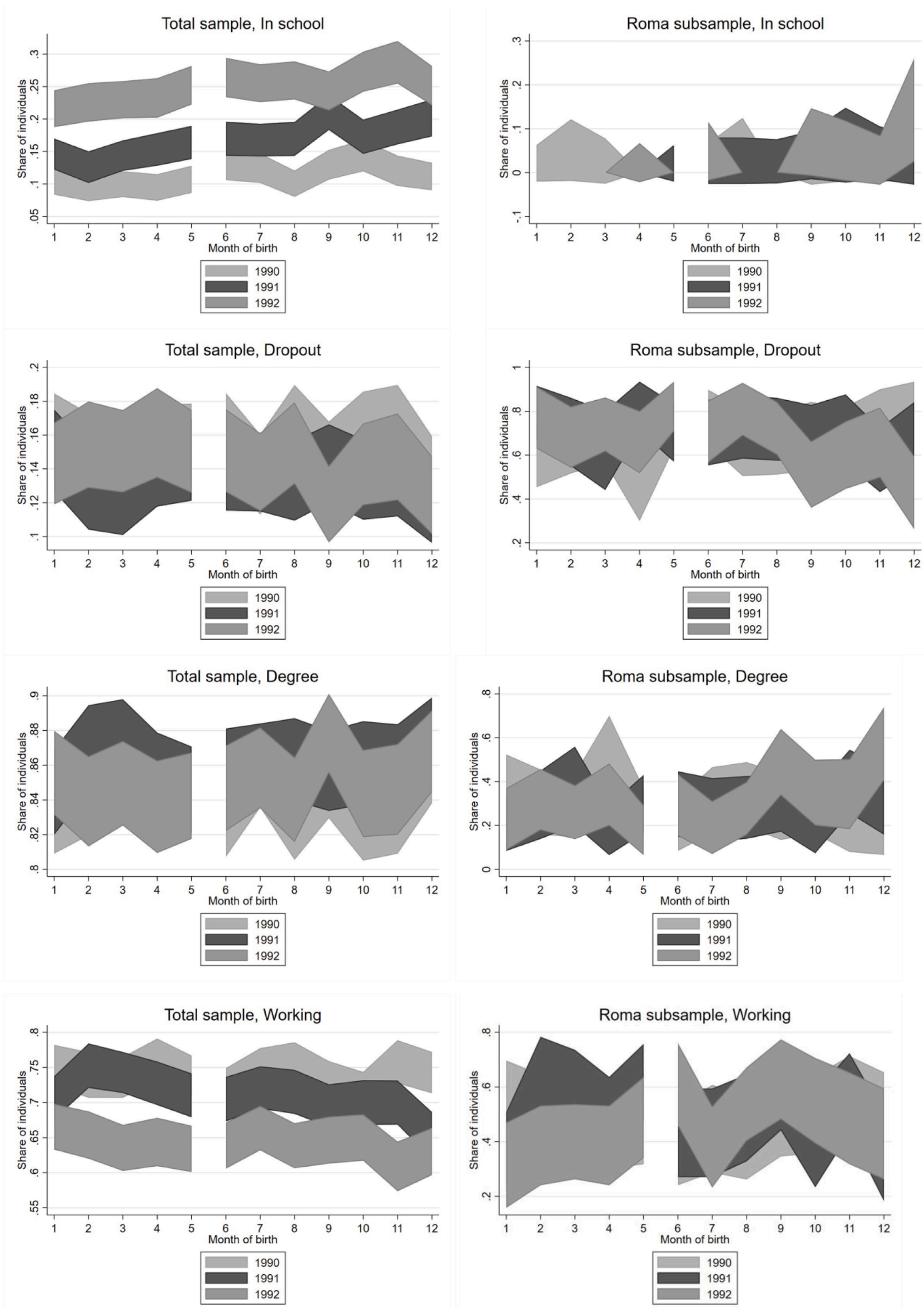

Fig. C3. The outcome variables around the cutoff for those born in 1990-1992, measured at age 25
Source: own estimation from the 2016 Microcensus. The means of the outcome variables by month of birth, along with their 95% confidence intervals. No. of observations: Total sample: 29,129; Roma subsample: 1,462. Note that in the second graph in the first row, "Roma subsample, In School", no one was in school among those born in some birth months, and the exact zeros are not plotted.

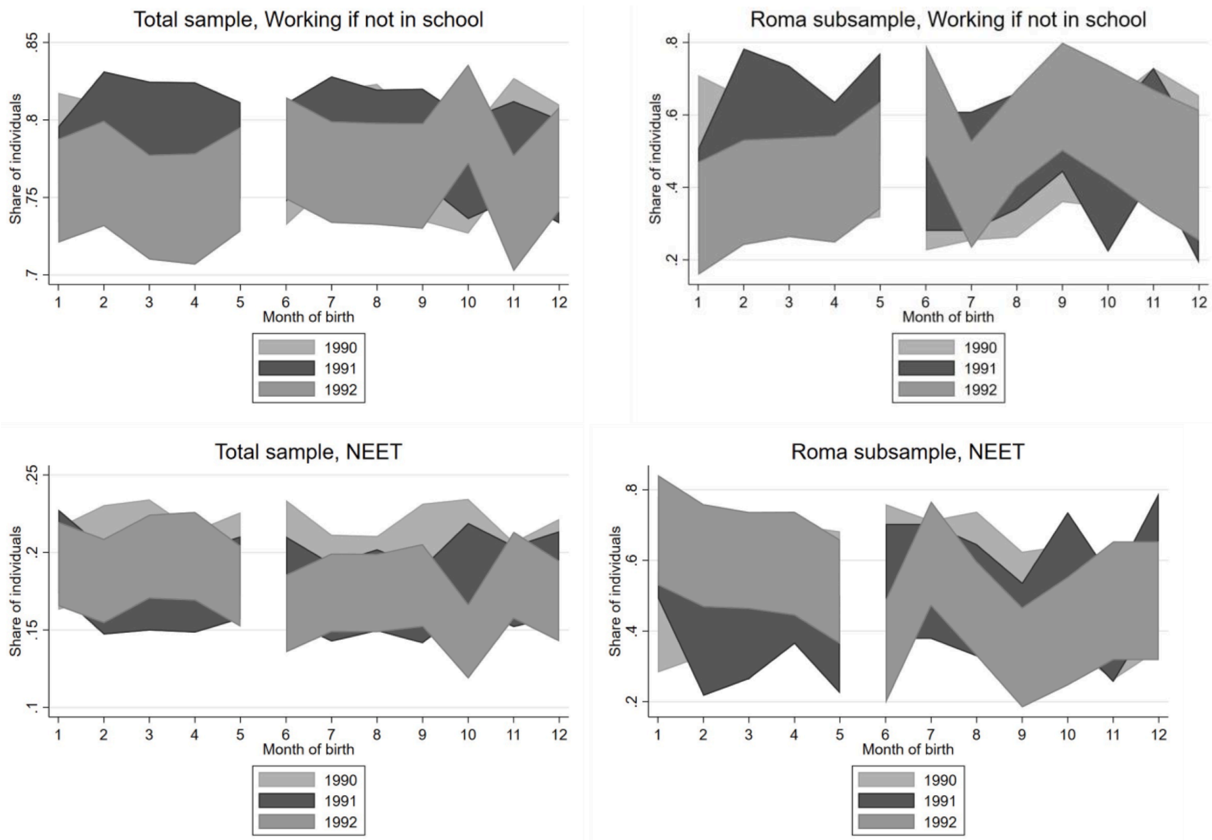

Fig. C3. (*continued*).

school, and did not see their family doctor (GP) between 2009 and 2017, at least once until age 25. Even though being "invisible" in administrative data does not necessarily mean that someone lives abroad, it is very unlikely that someone would not show up in the system at least once for 7 years if they still lived in Hungary. The results of this exercise suggest that reform had no effects on the probability of migration.

4.3. Empirical methods and robustness checks

First, I estimate the differences in schooling outcomes around the 1 June cutoff in the comparison and the reform cohorts in a fuzzy regression discontinuity design (RDD) framework. As discussed before, one can estimate

- the ITT effects of starting school at age 7 (as opposed to age 6) in the comparison cohorts ($RDD_{comparison}$), while
- the ITT effects of starting school at age 7 and the reform together (RDD_{reform}) in the reform cohort.

I estimate the RDD effects by fitting triangular kernel-weighted local linear regressions on both sides of the cutoffs (Hahn, Todd & Van der Klaauw, 2001; Imbens & Lemieux, 2008). The triangular weighting ensures that the closer is an observation to the cutoff, the higher weight it gets. 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 cutoffs (Cheng, Fan & Marron, 1997). For simplicity, I start by setting 150-day (or 5-month[9]) bandwidths for all cohorts and all outcome variables. Then, I apply alternative bandwidths as robustness checks (Fig. C3 in Appendix C).[10]

The RDD models are estimated separately for all cohorts and are of the form

$$y_i = \alpha + \beta_{RDD} * june_i + f(running_i, june_i) + \varepsilon_i, \qquad (1)$$

where

y_i is one of the six main outcome variables of individual i: being in school, dropping out, earning a secondary degree, working, working if not being in school, neither being in school nor in employment (NEET);

$june_i$ is a binary variable which is one if individual i was born on or after 1 June and 0 otherwise;

β_{RDD} is the RDD parameter of interest;

$running_i$ is the running variable that captures how far one's day (in 2011) or month (in 2001 and 2016) of birth is below or above the cutoff (0 if one was born on 1 June);

$f(running_i, june_i)$ is a kernel-weighted local linear function of $running_i$ which is different on the two sides of the cutoff (practically: $\beta_1 * running_i + \beta_2 * june_i * running_i$); and

ε_i is a usual error term.[11]

Theoretically, the fact that in 2001 and 2011, the running variable is

[9] When RDD models are estimated separately, the running variable is in days in the 2011 Census and months in the 2011 and 2016 Censuses (as they do not capture the day of birth). In the pooled DRDD models, the running variable is in months.

[10] A previous working paper version of this article used the bandwidth optimization routine of Calonico *et al.* (2017) to set the bandwidths. However, as the optimal bandwidths are different for all cohorts and outcomes, its use is not intuitive in a DRDD framework. The optimal bandwidths using the method were between 90-150 days around the cutoff, and this interval is covered by applying the 3-5-month bandwidths (but always the same for each cohort) in Figure C3 in Appendix C.

[11] Practically, the following model is estimated in Stata: $y = i.june\#\#c.running\ if\ running \geq - \quad bw\ \&\ running < bw\ [pw = \quad triangular\ weights],\ vce(cluster\ month - and - year - of - birth)$.

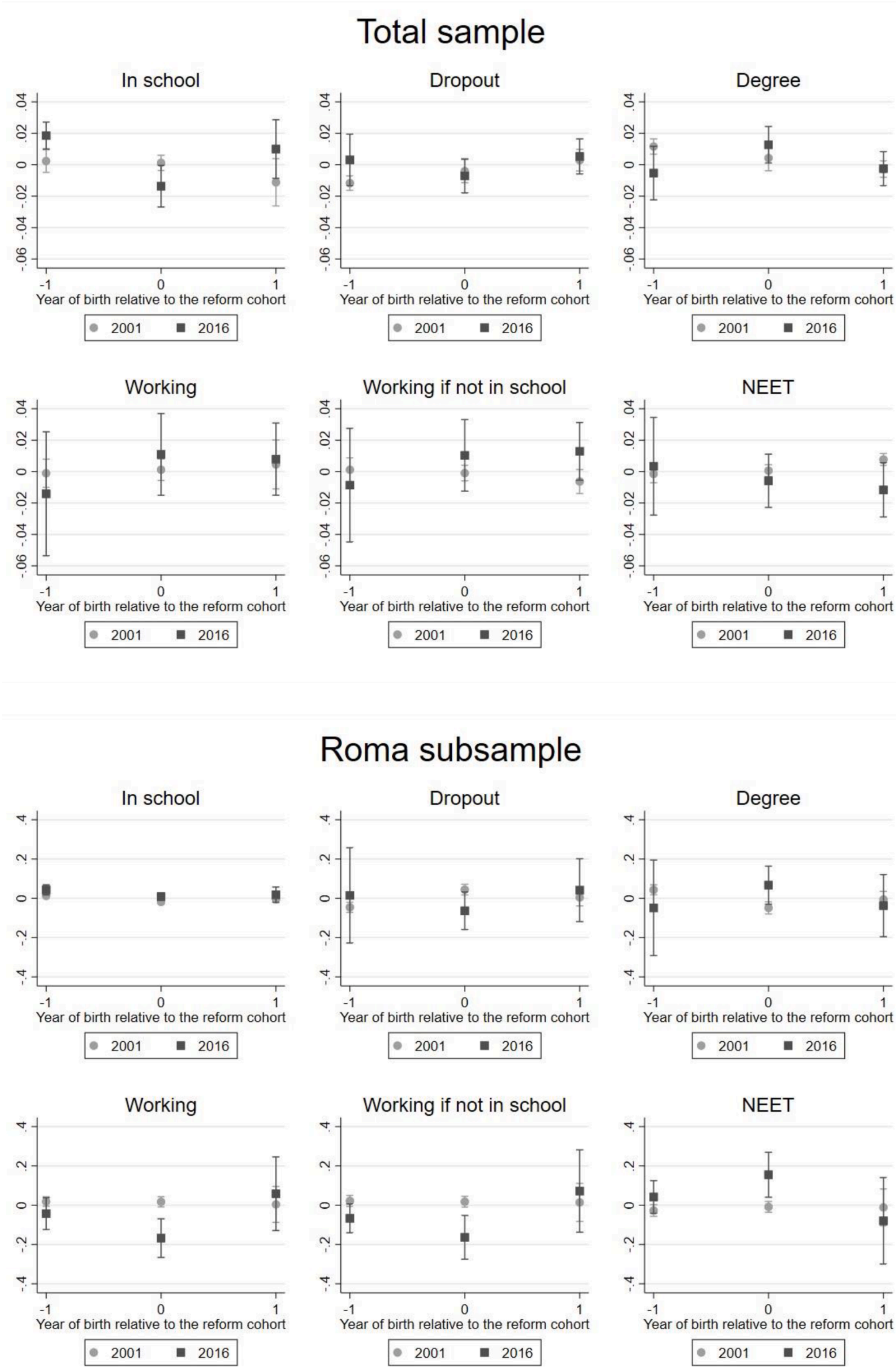

Fig. C4. The estimated RDD coefficients around the reform and comparison cutoffs (outcomes measured at age 25)
Source: own estimation from the 2001 and 2011 Censuses. RDD coefficients estimates (β_{RDD}) according to Equation (1). Bandwidth: 5 months. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations: 422,083 and 10,515 at age 25.

monthly, and not daily, might lead to model specification errors, because it is not possible to go close enough to the cutoff in terms of day of births (Hong, Dragan & Glied, 2019). The traditional way to correct for such specification errors is to cluster the standard errors by the running variable (Lee & Card, 2008). While there is a growing literature examining when and how clustering the standard errors is necessary and useful in general (Abadie et al., 2023) and in RDD setups in particular (Kolesár & Rothe, 2018), this paper follows a conservative approach and employs standard errors clustered by month-and-year-of-birth, as most papers in this literature (Hong, Dragan & Glied, 2019). As it will be

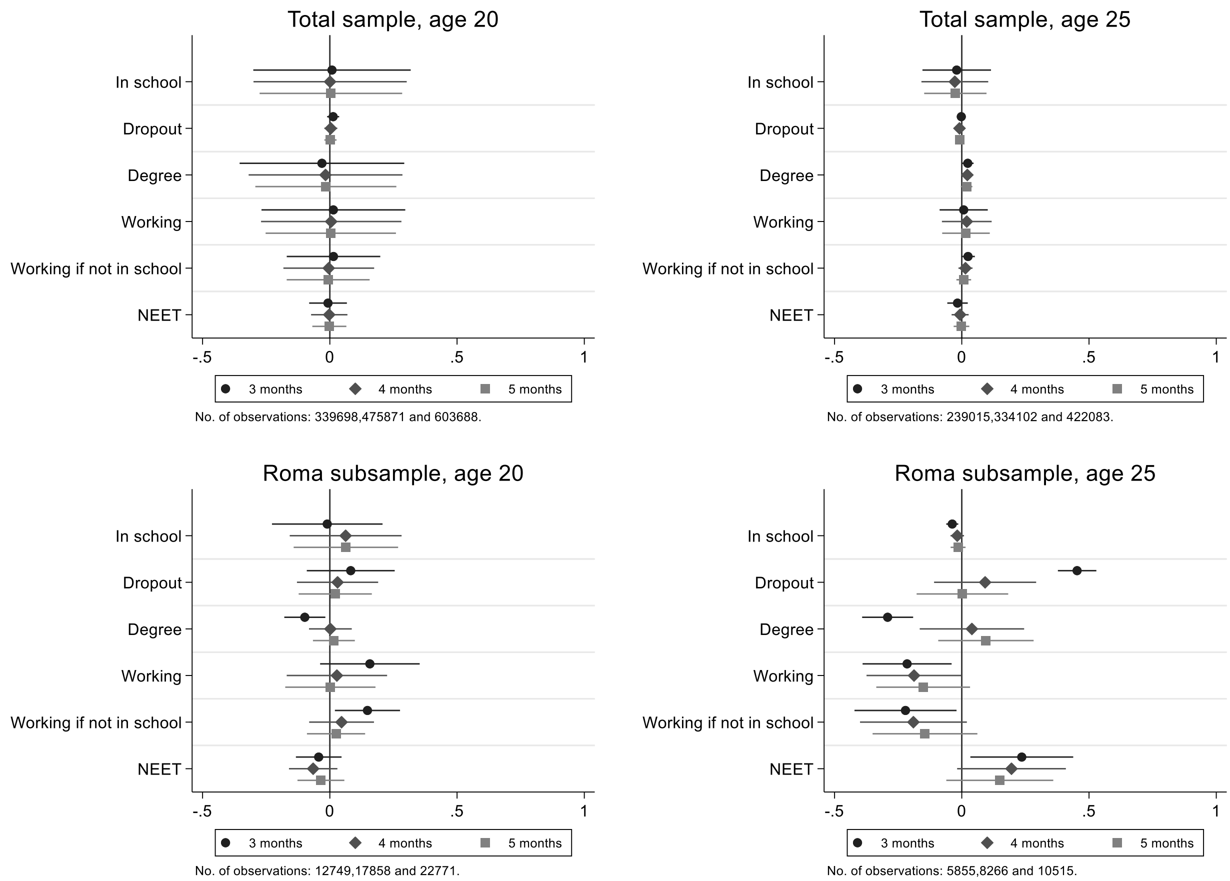

Fig. C5. The effects of the reform on schooling and employment outcomes – DRDD estimates using 3-5 months bandwidths
Source: own estimation from the 2001, 2011 and 2016 Censuses. DRDD coefficients estimates (β_{DRDD}) according to Equation (2). Bandwidth: 3, 4 and 5 months. Confidence intervals are constructed based on robust standard errors clustered by month of birth.

shown, most estimated effects will be very close to zero, so in this case, clustering the standard errors or not would not have any meaningful effects on the conclusions of the paper.

$y_i = \alpha + \beta_{DRDD} * \ june_i * middle\ year\ of\ birth_i * Census_i +$

$$y_i = \alpha + \beta_{DRDD} * \ june_i * middle\ year\ of\ birth_i * Census_i + \beta_1 * \ middle\ year\ of\ birth_i * Census_i \ + \beta_2 * Census_i \ * june_i \ + \ \beta_3 * \ middle\ year\ of\ birth_i * Census_i \ + \ \beta_4 * \ middle\ year\ of\ birth_i + \ \beta_5 * Census_i + \beta_6 * june_i + f(running_i, june_i, year\ of\ birth_i) \ + \ \varepsilon_i, \tag{2}$$

The estimated DRDD models are extensions of the RDD models. The sample of the treated and comparison cohorts are pooled together (using six cohorts at a time), and the local linear regression functions are still different below and above the cutoff for each cohort. The DRDD coefficient is identified by adding a Census dummy variable (that captures average differences occurring in 10(15) years), a middle year of birth dummy variable (that captures how those at age 20 differ from those at age 21 and age 19 on average) and their interaction terms with $june_i$ to the model[12]:

where

$middle\ year\ of\ birth_i$ is a binary variable which is one if individual i was born in 1991, 1981 or 1976 and zero otherwise;

$Census_i$ is a binary variable if individual i is captured by the 2011 (or 2016) Census and 0 if captured by the 2001 Census,

β_{DRDD} is the DRDD parameter of interest;

$year\ of\ birth_i$ is year of birth; and $f(running_i, june_i, year\ of\ birth_i)$ is a kernel-weighted local linear regression function of $running_i$, which is different on the two sides of the cutoff in all cohorts separately.[13]

It was discussed above that Roma ethnic minority students were more likely to drop out than the average before the reform. As they

[12] Note that analytically, this is the same as the simple double subtraction indicated in the last rows of Table 3 (except that the double subtraction would weight all cohorts equally but in reality, younger cohorts are smaller due to decreasing fertility over time). The pooled regression is needed to estimate standard errors.

[13] Practically, the following model is estimated in Stata: $y = i.june\#\#$i. middle year##i.Census + $c.running\#\#i.year\ of\ birth\#i.june\ if\ running \geq -bw\ \&\ running < bw\ [pw = triangular\ weights],\ \ vce(cluster\ month - and - year - of - birth)$.

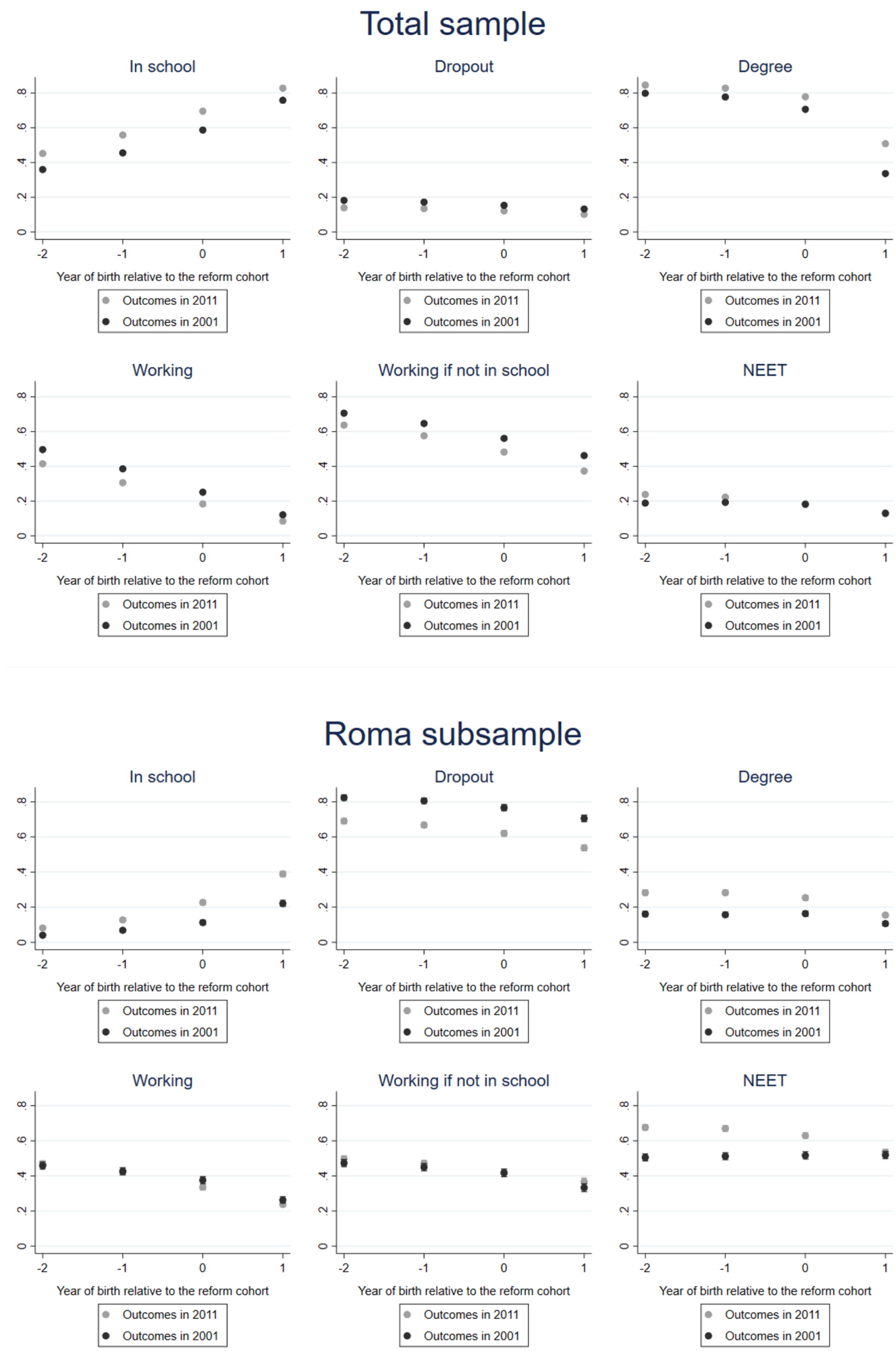

Fig. C6. Parallel trends: the means of the outcome variables by year of birth at age 20 (those born in June-Dec in 1989-1992 and 1979-1982)
Source: own estimation from the 2001 and 2011 Censuses. Sample of those born in June-Dec in 1989-1992 and 1979-1982. Year 0 on the x axis refers to those born in 1991 and 1981. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations: 621,228 and 23,190.

might be seen as a potential implicit target group of the reform, all results are shown on the total sample as well as on the subsample of Roma young people for the outcome variables from the Censuses. As the Admin3 database does not have information on ethnicity, it does not allow to look at the effects for the Roma subsample.

5. Results and robustness checks

Fig. 2 shows the probability of still being in school after completing 5-12 grades in 2011 among those born 150 days below and above the cutoff in 1990-1992. On the total sample, the difference across the two

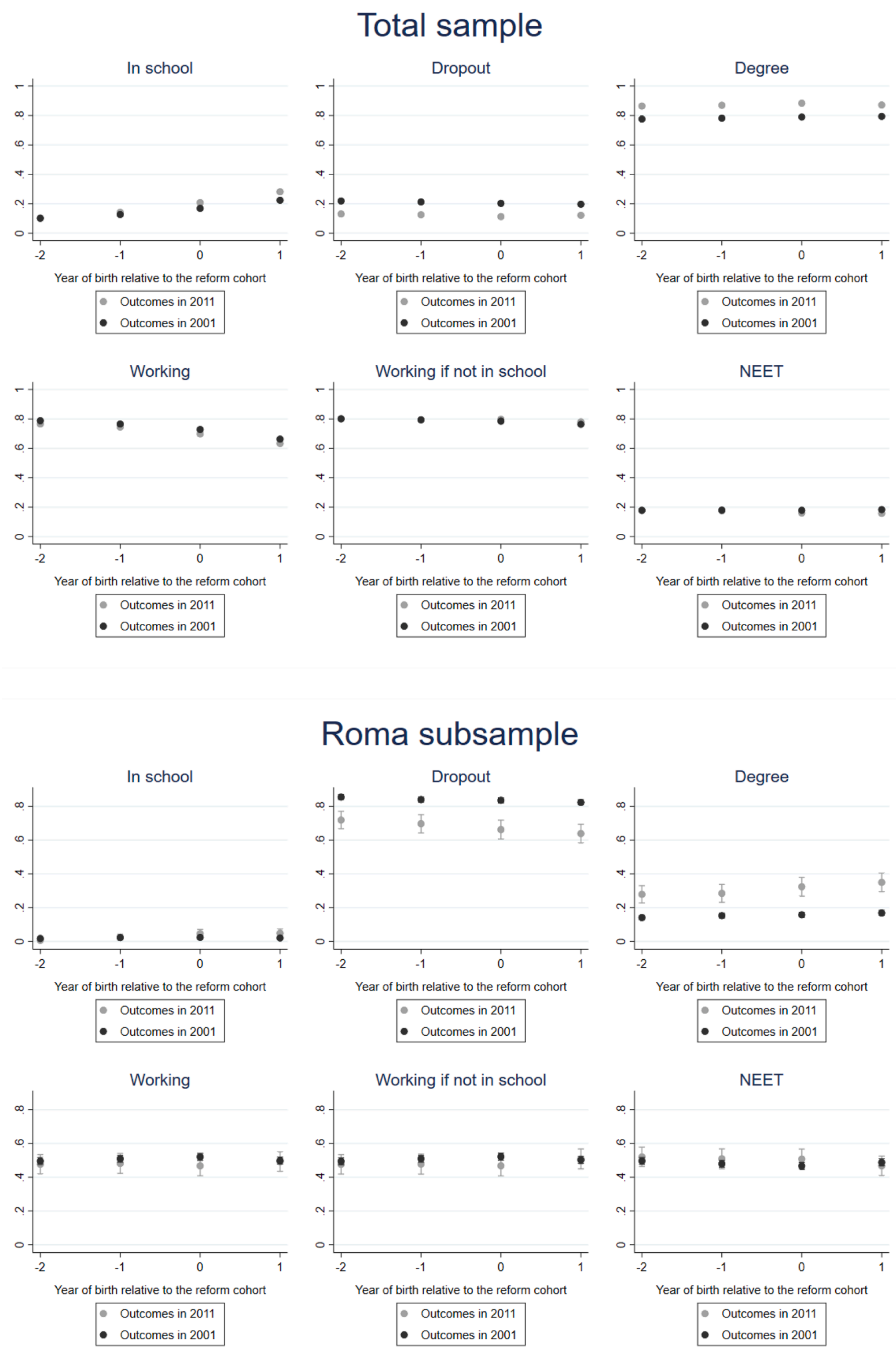

Fig. C7. Parallel trends: the means of the outcome variables by year of birth at age 25 (those born in June-Dec in 1989-1992 and 1974-1977)
Source: own estimation from the 2001 and 2016 Censuses. Sample of those born in June-Dec in 1989-1992 and 1974-1977. Year 0 on the x axis refers to those born in 1991 and 1976. Confidence intervals are constructed based on robust standard errors clustered by year and month of birth. No. of observations: 426,813 and 10,527.

groups is stable across the three cohorts. The data don't suggest that those born above the cutoff in the reform cohort (1991) completed more grades in school than those born below the cutoff, compared to the cohorts born one year before and after. Among the Roma students, however, those born above the cutoff in 1991 were relatively more likely to be still in school after completing 8, 9 and 10 grades than those born below the cutoff, and this difference is not present in the comparison cohorts. We do not see any effects at or above Grade 11, though. These results suggest that the reform increased the number of successfully

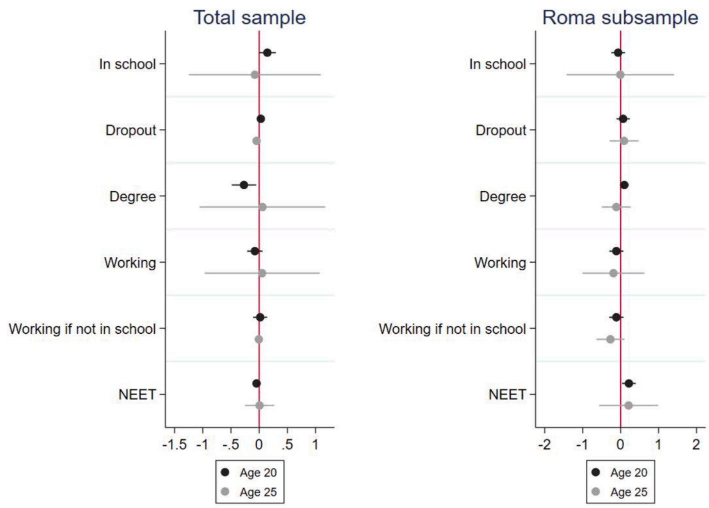

Fig. C8. The effects of the reform on schooling and employment outcomes– robustness checks using a diff-in-diffs identification strategy: DiD coefficients (Census data)
Source: own estimation from the 2001, 2011 and 2016 Censuses. All coefficients are estimated in separate regressions. Confidence intervals are constructed based on robust standard errors clustered by year and month of birth. Age 20: sample of those born in June-Dec in 1989-1992 (*treated*) and 1979-1982 (*control*). No. of observations: 621,228 and 23,190. Age 25: sample of those born in June-Dec in 1989-1992 (treated) and 1974-1977 (control). *Pre-periods*: the first two birth years in each category; *post-periods*: the second two birth years of each category. The plotted coefficients are estimated on the interaction term of *treated* and *post*. No. of observations: 426,813 and 10,527.

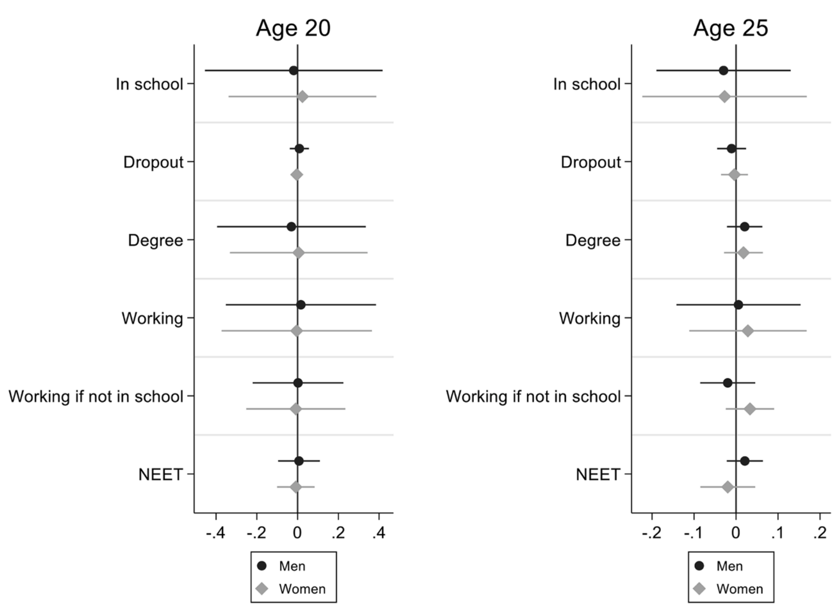

Fig. C9. The effects of the reform on schooling and employment outcomes (DRDD coefficients) by gender (Total sample, Census data)
Source: own estimation from the 2001, 2011 and 2016 Censuses. All coefficients are estimated in separate regressions. DRDD coefficients estimates (β_{DRDD}) according to Equation (2). Bandwidth: 5 months. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations: age 20: men: 309,906, women: 293,782; age 25: men: 215,535, women: 206,548,

completed grades among Roma students, but it did not affect the probability of completing the 12th grade and earning a secondary degree.[14]

Fig. C1 and Fig. C3 in Appendix C plot the averages of the outcome variables by month of birth while Fig. C2 and Fig. C4 plot the estimated RDD coefficients at age 20 and 25 using the Census data. The latter show the credibility of the second identification assumption: these RDD coefficients are either not statistically different from each other in the comparison cohorts or those measured in 2011 (2016) are parallel to those measured in 2001 (on individuals of the same age). Second, they also suggest that the difference of the 1991 and 1981 RDD coefficients is not different from the average difference between the other two pairs of estimates. In other words, they suggest that the reform had no effect on these outcomes. This conclusion is supported by the formal DRDD estimates (Fig. 3). On the total sample, almost all coefficients are very close to zero, and these effects are similar for men and women (Fig. C9 in Appendix C). On the subsample of Roma young people, estimates are not precisely zero, but none are significantly different from zero (Fig. 3).

[14] Note that the *number of successfully completed grades* does not include unsuccessful school years that ended with grade retention. Thus, this measure does not allow us to make an inference about how the reform affected the lenght of schooling.

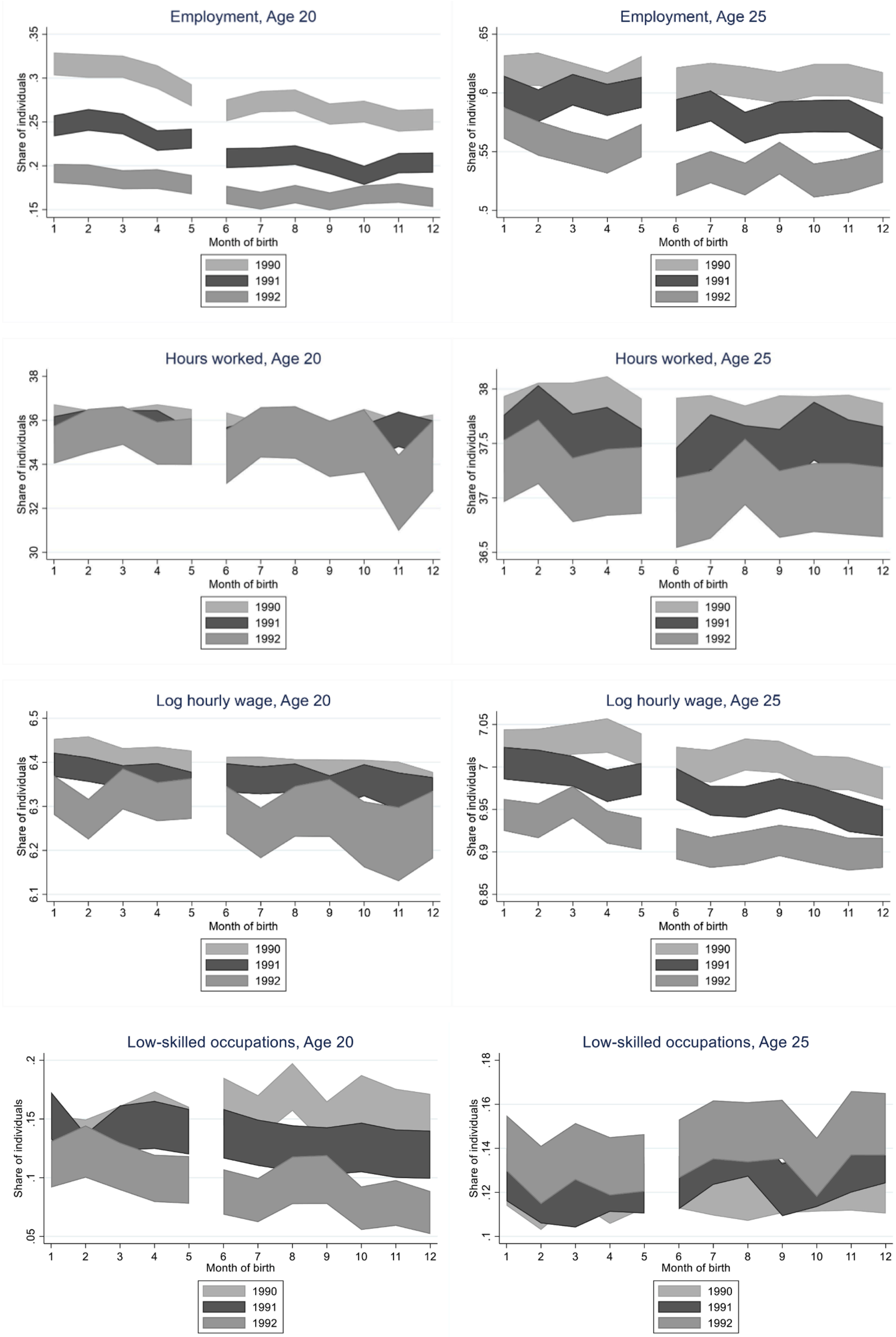

Fig. D1. Outcome variables around the cutoff for those born in 1990-1992 (Admin3 data)
Source: own estimation from the Admin3 database. Outcomes are measured in October 2011 (age 20) and in October 2016 (age 25). No. of observations: Age 20 employment: 176,879; age 20 other outcomes: 39,685; age 25 employment: 176,519; age 25 other outcomes: 102,418.

Robustness checks to these results are provided in Appendix C. Fig. C5 plots the same DRDD coefficients using 3-5 months bandwidths. Figs. C6–C7 present an alternative identification strategy using a difference in differences (DiD) strategy. As shown in Fig. B1 in Appendix B, those born in June-December are very likely to enrol in elementary school at age 7. Thus, I use the subsample of those born in June-December in 1989-1992 (two years before and two years after the introduction of the reform) as the *treated group*, as well as those born exactly 10 and 15 years earlier as before as the *control group*. Then, I set up a DiD strategy using the first two cohorts of the two groups as the *before* and the second two cohorts as the *after* period. Then, the interaction term of *treated group* and *after* identifies the causal effect of the

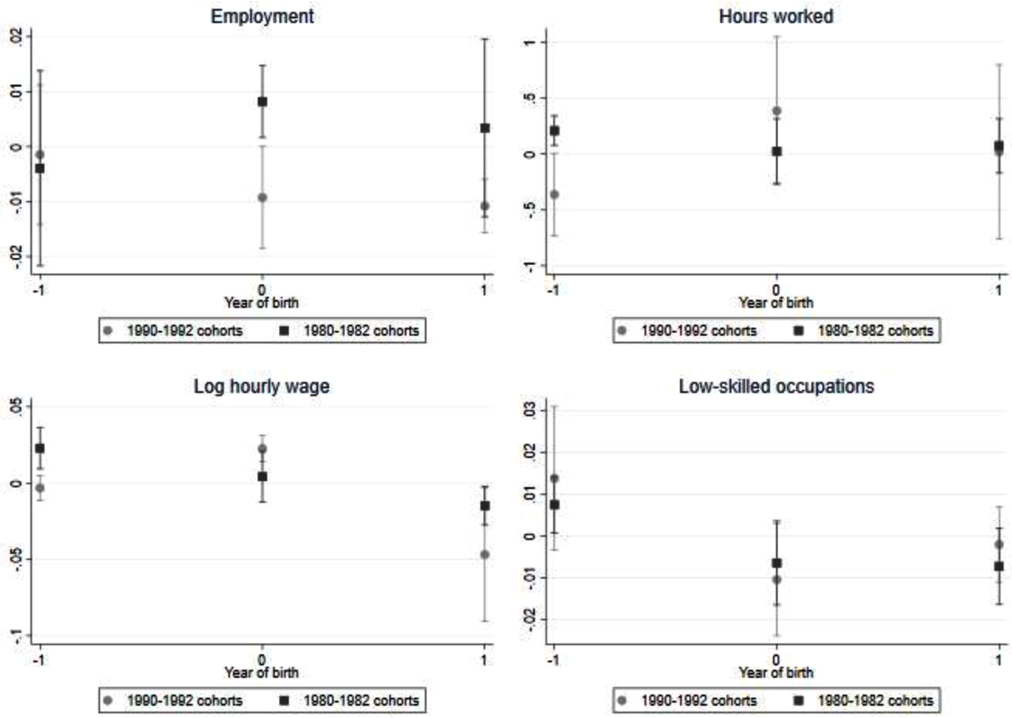

Fig. D2. The estimated RDD coefficients around the reform and comparison cutoffs, age 20 (Admin3 data)
Source: own estimation from the Admin3 database. Outcomes are measured in October 2011. All coefficients are estimated in separate regressions. RDD coefficient estimates (β_{RDD}) according to Equation (1). Bandwidth: 5 months. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations: employment: 288,768; other outcomes: 97,197.

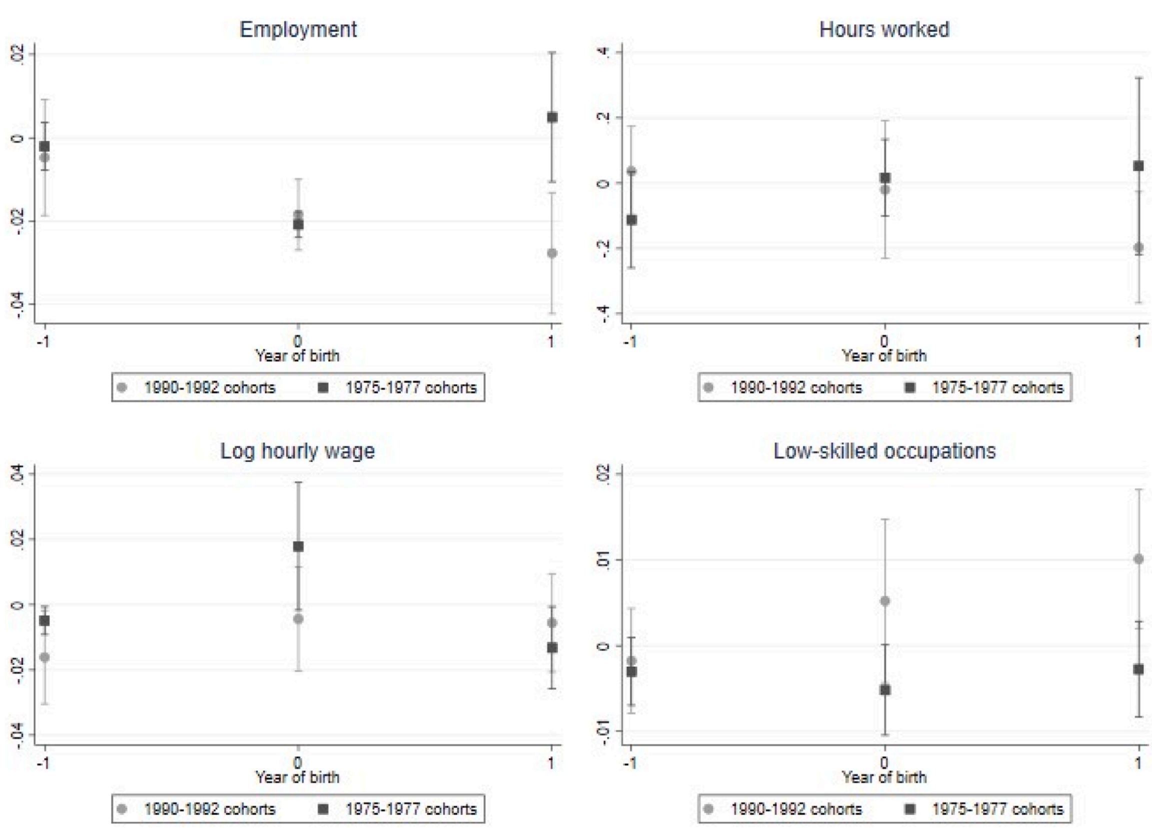

Fig. D3. The estimated RDD coefficients around the reform and comparison cutoffs, age 25 (Admin3 data)
Source: own estimation from the Admin3 database. Outcomes are measured in October 2016. All coefficients are estimated in separate regressions. RDD coefficient estimates (β_{RDD}) according to Equation (1). Bandwidth: 5 months. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations: employment: 342,974; other outcomes: 192,187. Note that the number of observations at age 25 is larger than the number of observations at age 20 because the 1975-1977 cohorts were substantially larger than the 1980-1982 cohorts.

Fig. D4. The effects of the reform: the estimated DRDD coefficients with alternative bandwidths (Admin3 data)
Source: own estimation from the Admin3 database. Outcomes are measured in October 2011 (age 20) and October 2016 (age 25). All coefficients are estimated in separate regressions. DRDD coefficients estimates (β_{DRDD}) according to Equation (2). Bandwidth: 5 months. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations: Age 20: employment: 163,175/228,415/288,768 with 3/4/5-month bandwidths; other outcomes: 55,116/77,137/97,197 with 3/4/5-month bandwidths. Age 25: employment: 193,039/271,099/342,974 with 3/4/5-month bandwidths; other outcomes: 108,030/151,915/192,187 with 3/4/5-month bandwidths. Note that the number of observations at age 25 is larger than the number of observations at age 20 because the 1975-1977 cohorts were substantially larger than the 1980-1982 cohorts.

reform in a DiD setup. Fig. C6 presents the parallel trends for the outcomes at age 20 and Fig. C7 for the outcomes at age 25. Fig. C8 plots the estimated DiD coefficients that are similar to the DRDD estimates. Note that while the DRDD estimates identify local ITT effects around the 1 June cutoff, the DiD estimates identify average ITT effects among those born between June and December; thus, the two are not perfectly comparable. Still, the DiD estimates also being (very close to) zero supports the conclusions drawn from the DRDD estimates.

Next, I investigate the effects of the reform on employment, hours worked, log hourly wages and the probability of working in low-skilled occupations (those that do not require a qualification) at ages 20 and 25 using the Admin3 database. Fig. D1 in Appendix D plots these outcomes for those born in Jan-May versus June-Dec in 1990-1992, while Fig. D2 and Fig. D3 in Appendix D show the estimated RDD coefficients around the cutoff. Fig. 4 presents the DRDD coefficients. These results confirm the earlier findings that the reform did not affect the probability of employment, now using administrative data instead of self-reported information. Furthermore, they also show that the reform had no significant effect on the number of hours worked, log hourly wages, and the probability of working in low-skilled occupations. These effects are robust to alternative bandwidths (Fig. D4 in Appendix D), similarly close to zero for both genders (Fig. D5 in Appendix D) and are also similar in a DiD approach (Fig. D6 and Fig. D7 in Appendix D).

6. Channels and mechanisms

This section briefly investigates the potential reasons why the reform did not have any meaningful effects. First, as already mentioned and suggested by Fig. A3 in Appendix A, it is likely that marginal students were simply too old to be kept in school by the reform until earning a degree.

Second, most marginal students attended vocational training schools. There is plenty of evidence in the literature that being older than one's peers and having low educational achievement, i.e. test scores or GPA, are among the key predictors of dropping out of secondary school (Cratty, 2012). Before the reform, 42 % of those in vocational training schools aged at least 17 already in Grade 10, while in high schools, this ratio was only 22 % (Fig. E1 in Appendix E). The average standardized math test scores of vocational training schools students were about 0.8 standard deviations below the average, but they were even lower among those aged 17 or above, at about -1.0 standard deviation (Fig. E1 in Appendix E). It is likely that the marginal students

Fig. D5. The effects of the reform: the estimated DRDD coefficients by gender (Admin3 data)
Source: own estimation from the Admin3 database. Outcomes are measured in October 2011 (age 20) and October 2016 (age 25). All coefficients are estimated in separate regressions. DRDD coefficients estimates (β_{DRDD}) according to Equation (2). Bandwidth: 5 months. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations: Age 20: employment: 148,042/140,726 for men/women; other outcomes: 55,900/41,297 for men/women. Age 25: employment: 174,484/168,490 for men/women; other outcomes: 101,174/91,013 for men/women. Note that the number of observations at age 25 is larger than the number of observations at age 20 because the 1975-1977 cohorts were substantially larger than the 1980-1982 cohorts.

Fig. D6. Parallel trends: the means of the outcome variables by year of birth at age 20 and 25 (those born in June-Dec in 1989-1992 (treated cohorts) and 1983-1986 (control cohorts); Admin3 data)
Source: own estimation from the Admin3 database. Outcomes are measured around age 20 for both the treated and the control groups to make the sample as similar to the Censuses as possible. At age 20, outcomes for the *treated* group (those born in 1989-1992) are measured in October 2011 as before while the outcomes for the *control* group (those born in 1983-1986) are measured in October 2005. At age 25, outcomes for the *treated* group (those born in 1989-1992) are measured in October 2016 as before while the outcomes for the *control* group (those born in 1983-1986) are measured in October 2010. No. of observations: age 20: 277,902 for employment and 46,730 for the outcome variables of the employed; age 25: 277,269 for employment and 140,289 for the outcome variables of the employed.

are among them.

Descriptive evidence shows that the number of students increased heavily in vocational training schools. The first treated cohort reached Grade 10 in the 2007/2008 academic year. In the following four years, the number of students in vocational training schools increased by 8.2 % (Table E 1 in Appendix E). However, vocational schools' available resources did not follow the increasing number of students (Mártonfi, 2011b, 2011a). Table E 3 in Appendix E shows the average per capita expenditures of vocational schools. Although they did receive some extra funding in 2007 and 2008, their expenditures per student decreased by 21.5 % in the first two grades (Grade 9-10) and by 12 % in Grade 11+ between 2006 and 2011. Thus, while their workload increased, their per capita expenditures decreased.

Besides the higher number of students, their distribution also changed in vocational training schools. The increase in the total number of students (13,023) was almost the same as the increase in the number of disadvantaged students (12,098); thus, the increase came almost exclusively from this student group (Table E 2 in Appendix E). According to the school-level data of the NABC, the share of Roma students more than doubled between the Spring of 2007 and 2011, and the share of students with at least one unemployed parent grew by more than 50 % (Table E 2 in Appendix E).

Although vocational school students were on average older than high school students already before the reform, the age distribution of students shifted more and more to the right after the reform (Fig. in Appendix E). Fig. E1 also shows that the share of grade repeaters in Grade 10 increased from 15.8 % to 20.0 % in this period, which is a 40 % increase. It seems that some students forced to stay in school did not

Fig. D7. The effects of the reform: the estimated DiD coefficients (Admin3 data)
Source: own estimation from the Admin3 database. Outcomes are measured around age 20 for both the treated and the control groups to make the sample as similar to the Censuses as possible. At age 20, outcomes for the *treated* group (those born in 1989-1992) are measured in October 2011 as before while the outcomes for the *control* group (those born in 1983-1986) are measured in October 2005. At age 25, outcomes for the *treated* group (those born in 1989-1992) are measured in October 2016 as before while the outcomes for the *control* group (those born in 1983-1986) are measured in October 2010. All coefficients are estimated in separate regressions. No. of observations: age 20: 277,902 for employment and 46,730 for the outcome variables of the employed; age 25: 277,269 for employment and 140,289 for the outcome variables of the employed. *Pre-periods*: the first two birth years in each category; *post-periods*: the second two birth years of each category. The plotted coefficients are estimated on the interaction term of *treated* and *post*.

complete the grade they were in but rather repeated grades.

There are two potential ways how the reform could have affected the composition of these schools. If more low-ability students would have enrolled in secondary school due to the reform, who would not have enrolled otherwise, this might have mechanically led to a distribution change. However, this wasn't the case. Fig. E2 and Fig. E3 in Appendix E look at the effects of the reform on the probability of secondary school enrolment by school tracks. Fig. E2 plots RDD coefficients around the cutoffs of the reform cohort and five comparison cohorts in a similar fashion to Section 5, and Fig. E3 plots the estimated DRDD coefficients. These results show that the reform did not affect the probability of enrolment, either on average or in vocational training schools. Thus, keeping low-ability students in school for longer caused this composition effect, as those who would have dropped out at age 16 in the lack of the reform stayed in school instead.

Interestingly, it seems that the reform might have increased the probability of dropping out in vocational training schools, especially among Roma students. Besides enrolment, Fig. E2 and Fig. E3 in Appendix E also look at the effect of the reform on the probability of dropping out by secondary school type. Fig. E2 shows that the jump in the probability of dropping out of a vocational school around the cutoff is significantly larger in the reform cohort than in the comparison cohorts; however, the estimated DRDD coefficients are not significant (Fig. E3). As the subsample of Roma students who enrolled in vocational schools is relatively small around the cutoff, standard errors are large, and although the estimated coefficient is meaningful in terms of magnitude (at 10 percentage points), it stays below significance.

7. Discussion

This paper looks at the effects of increasing the CSL age from 16 to 18 on schooling and labour market outcomes in Hungary. It finds that the reform neither decreased the probability of dropping out nor increased the probability of earning a degree. These results are in line with Landis and Reschly (2010), Cabus and De Witte (2011) and Grenet (2013), who showed that increasing the CSL age would not necessarily increase the probability of secondary school completion. While Landis and Reschly (2010), Cabus and De Witte (2011) and Grenet (2013) only look at the average effects of such reforms, this paper also examines the impact of the reform on the most vulnerable students, in particular, those from Roma ethnic minority background.

Lengths of schooling increased by about half a year on average and by a year among Roma ethnic minority students after the reform. However, marginal students were older than their peers due to grade retentions. Despite the two additional years of schooling, these students were not able to complete Grade 12 and achieve a degree by age 18. Among Roma ethnic minority students, who were about five to six times as likely to drop out of school as the average before the reform, the reform increased the probability of completing more grades but did not affect the probability of earning a degree.

The data suggest that vocational training schools that most marginal students attended played a key role. The share of over-aged, low-attainment, disadvantaged students increased massively in these schools after the reform, and vocational training schools were not able to handle the increased workload. The per capita expenditures of these schools even decreased after the reform, and the increasing number of probably unmotivated and low-ability students might have crowded out schools' resources. While students did stay in school for longer, they were not able to complete the requirements of their grades and kept repeating grades. The probability of dropping out might have even increased in

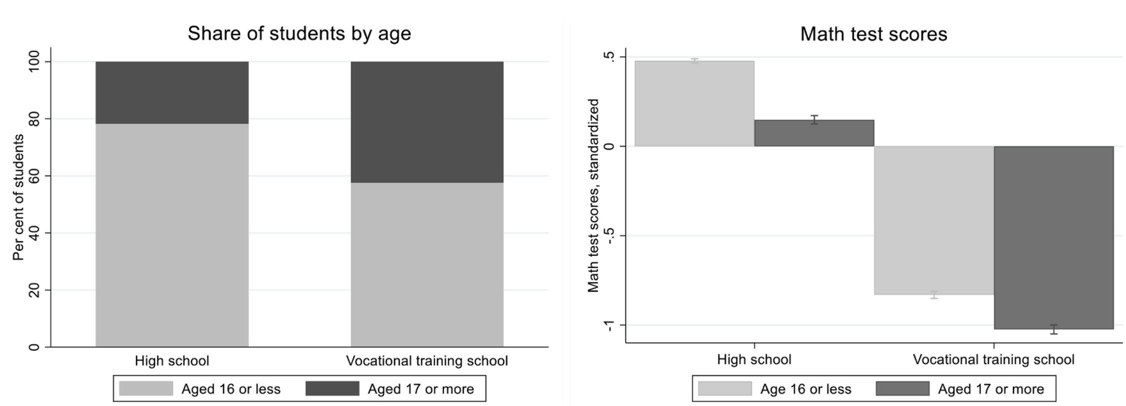

Fig. E1. The age and average standardized math test scores of students in Grade 10 before the reform (2007)
Source: own estimation from the Assessment of Basic Competencies (NABC) database, Grade 10 data in 2007. Weighted by sampling weights. No. of observations: high schools: 27,719, vocational training schools: 8,554.

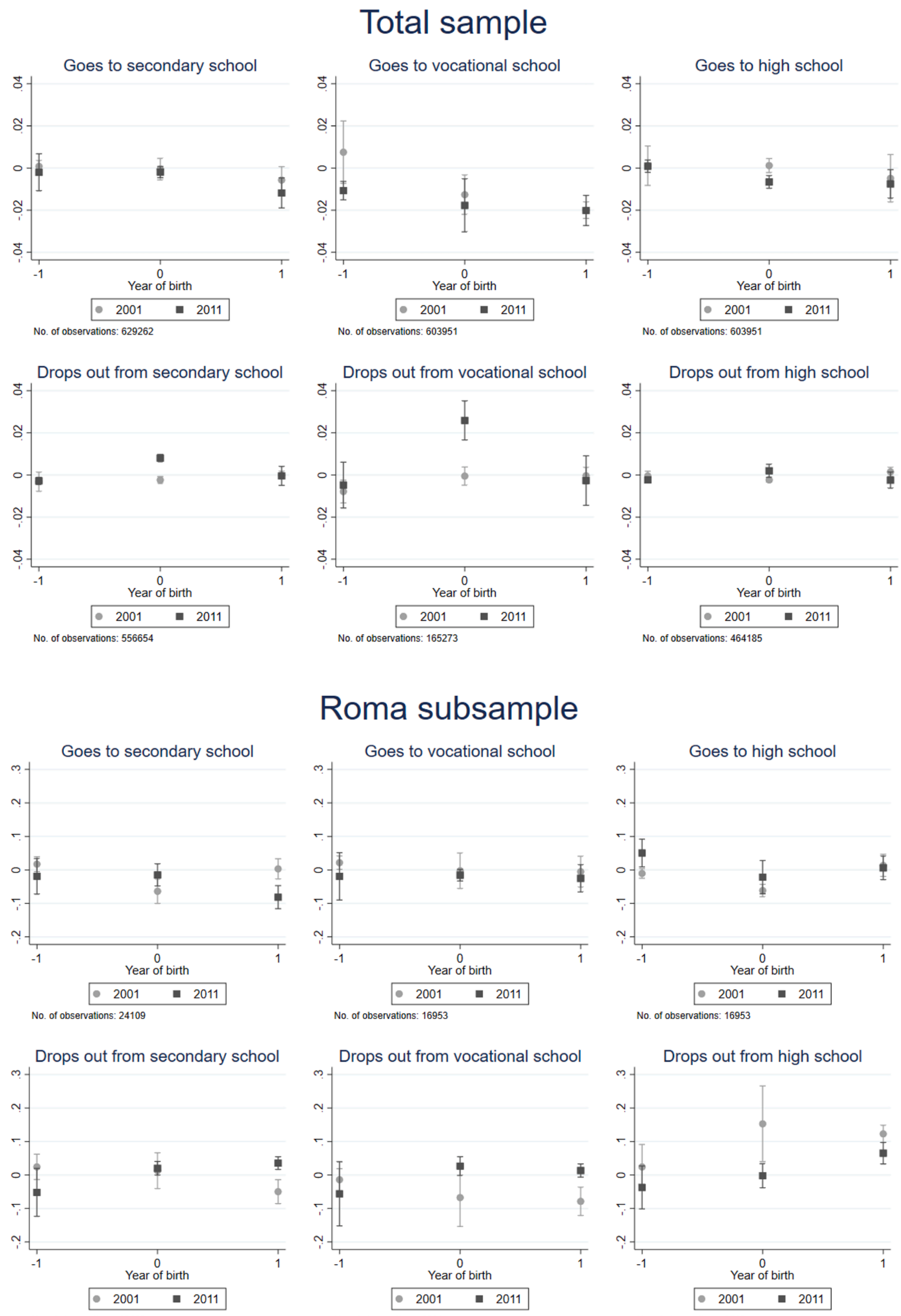

Fig. E2. The effect of the reform on secondary school enrolment and dropping out by school type: RDD coefficients
Source: own estimation from the 2001 and 2011 Censuses. RDD coefficients estimates (β_{RDD}) according to Equation (1). Bandwidth: 150 days in 2011 and 5 months in 2001. Confidence intervals are constructed based on robust standard errors clustered by month of birth.

vocational training schools due to the reform, especially among Roma students; although this latter effect is large but not statistically significant. The deteriorating quality of vocational training schools after the reform might indicate that CSL age reforms are not necessarily reliable instrumental variables for education. Especially, if those likely to be affected by such reforms (in the sense that they would leave school earlier in the lack of the reform) are also the ones affected by potential negative institutional consequences, like crowding out or peer effects.

If we believe that attending school develops human capital, staying in school longer could be beneficial even if it does not increase the probability of earning a secondary degree. This paper finds that the reform had no effect on the probability of employment, hours worked,

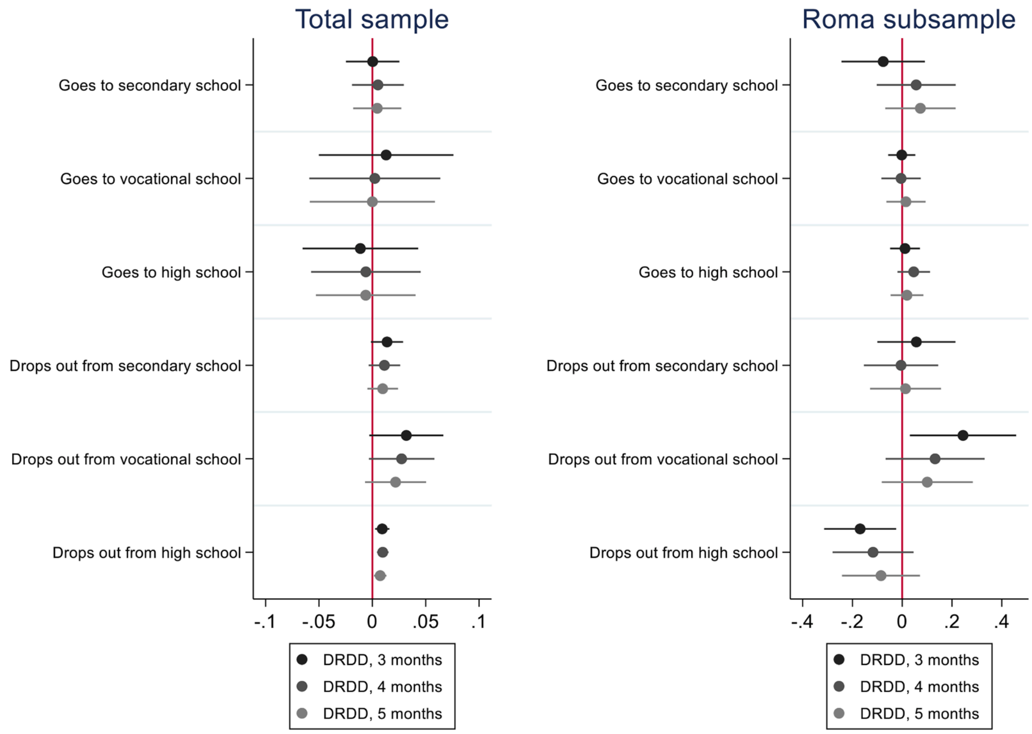

Fig. E3. The effect of the reform on secondary school enrolment and dropping out by school type: DRDD coefficients.
Source: own estimation from the 2001, 2011 and 2016 Censuses. DRDD coefficients estimates (β_{DRDD}) according to Equation (2). Bandwidth: 3, 4 and 5 months. Confidence intervals are constructed based on robust standard errors clustered by month of birth. No. of observations are the same as on Figure E 2.

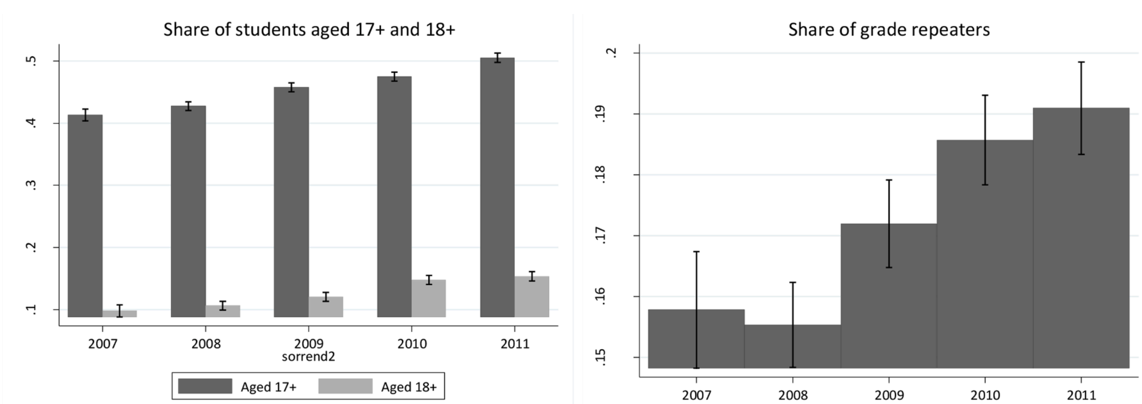

Fig. E4. Students in Grade 10 in vocational training schools (Spring of 2007-2011).
Source: own estimation from the NABC database. Means along with their 95% confidence intervals. Weighted by sampling weights. No. of observations: 82,194. Share of grade repeaters: share of those who repeated grades during their vocational training school studies.

Table E1
The number of students in vocational training schools, 2006-2011

	Academic year					Change between the 2006/2007 and the 2010/2011 academic year	
	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	Index, 2006=100	No. of students
All students	124,466	129,066	128,848	135,268	137,489	110.5	13,023
Full-time students	119,637	123,192	123,865	128,674	129,421	108.2	9,784
Grade repeaters	6,445	6,659	8,322	10,457	11,825	183.5	5,380
Disadvantaged students	28,586	27,916	32,431	37,947	40,684	142.3	12,098
Students facing multiple disadvantages	5,442	8,552	11,114	13,470	12,679	233.0	7,237

Source: own collection from the Statistical Yearbooks of Education, 2007-2012.
[15]Retrieved from https://www.oktatas.hu/kozneveles/kozerdekuadatok.

Table E2
Student composition in vocational training schools

	Academic year					Change between 2006/2007 and 2010/2011
	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	per cent
Roma students	0.178	0.265	0.354	0.373	0.408	229.2
Students receiving child protection subsidy	0.267	0.351	0.407	0.413	0.436	163.3
Students living in a financially deprived family	0.41	0.499	0.462	0.506	0.522	127.3
Students with at least one unemployed parent	0.285	0.371	0.408	0.444	0.435	152.6

Source: own calculation from the school-level data of the National Assessment of Basic Competencies (NABC) database. Schools offering elementary education or academic high school tracks along with a vocational training school track are excluded.

Table E3
Expenditures of vocational training schools, 2006-2011

	2006	2007	2008	2009	2010	2011
No. of students	124,466	129,066	128,848	135,268	137,489	139,823
Vocational education, Grade 9-10						
Expenditures at current prices (million HUF)	15,424	20,595	21,746	19,813	19,149	17,690
Expenditure per student at real prices (1,000 HUF)	123.9	153.6	150.4	123.0	112.3	97.2
Index of per student expenditures, 2006=100	100.0	123.9	121.4	99.3	90.6	78.5
Vocational training (Grade 11+ and advanced vocational training)						
Expenditures at current prices (million HUF)	43,128	53,813	55,875	49,435	55,499	52,105
Expenditure per student at real prices (1,000 HUF)	346.5	401.3	386.5	307.0	325.4	286.4
Index of per student expenditures, 2006=100	100.0	115.8	96.3	79.4	106.0	88.0
Price index, 2006=100	100.0	103.9	112.2	119.1	124.1	130.1

Source: own collection from the Statistical Yearbooks of Education, 2006-2012. Source of price index: National Statistical Office.

log hourly wages and the probability of working in low-skilled jobs at ages 20 and 25. These results show no signs of human capital effects on average, at least up until age 25. However, it is also possible that the signalling effect (Spence, 1973) of not having a secondary degree is stronger than the (potential) human capital development of additional schoolyears, hence the null results. Grenet (2013) argues that in Europe, where examination-based certificates are widespread to prove the completion of secondary education, earning a secondary degree is an essential outcome of education for labour market success. On the contrary, in the US, there is evidence that merely longer schooling had positive effects on wages (Angrist & Krueger, 1991). The incapacitation effect of education that keeps students in school and "*out of trouble*" (Anderson, 2014) might still undoubtedly be a positive outcome of the reform. Adamecz-Völgyi & Scharle (2020) showed that the same reform decreased the probability of teenage motherhood among Roma women through the incapacitation effect of schooling. They find that the reform decreased the probability of getting pregnant during the schoolyear, between September and June, but not during the summer and Christmas breaks, suggesting no human capital effects. Similarly, Adamecz-Völgyi *et al.* (2021) provided evidence about the incapacitation effects of a later reform that cut the CSL age back from age 18 to 16 in Hungary in 2011. They find that the reform increased the probability of being neither in school nor in employment between ages 16 and 18 but had no effects on labour market outcomes at age 19. While they are not able to look at the effects of this later reform at ages 20 and 25 (the ages investigated by the present paper) due to data availability issues, they find no signs of human capital effects at age 19. These results taken together suggest that in the Hungarian education system, two-year longer schooling without the explicit requirement of earning a degree generates important incapacitation effects, but its human capital effects are limited.

These results are not without caveats. First, school dropouts might go back to school later in life. Second, the main empirical strategy of this paper measures the effects of the reform comparing the first treated cohort to the last untreated cohort. Theoretically, it is not impossible that schools could have become better over time in handling the difficulties that they had to face, and subsequent cohorts would have experienced different outcomes. Lastly, long-term effects might also be possible to emerge in the future.

This paper offers useful insights for educational policy. While the existing literature investigates how such reforms work in high-income countries like the US, the UK, Germany or France, our knowledge is limited about what happens when standard educational interventions are implemented in medium-income countries with different (potentially less efficient) educational systems (especially those that allow grade retention). My results make it clear that raising the CSL age is not enough to increase secondary school completion if not accompanied by supply-side elements (schooling expansion). For a successful implementation of this reform, schools should have been explicitly supported by educational policy. This finding is in line with the experience of development programs using demand-side interventions only. For example, 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 & Hoddinott, 2010). Besides providing appropriate resources to schools, reforms should prolong compulsory schooling until earning a secondary degree rather than until a certain age (or aim for a combination of both). Especially in education systems that allow grade retention, the CSL age legislation should always be combined with a requirement of earning a secondary degree.

Declarations of Competing Interest

The authors declare there is no conflict of interests.

Data availability

The authors do not have permission to share data.

Acknowledgements

The present document has been produced using the 2001 and 2011 Hungarian Census and 2016 Hungarian Microcensus data files of the Hungarian Central Statistical Office, and the Panel of Linked Administrative Data (Admin3) database that is a property of the National Health Insurance Fund Administration, the Central Administration of National Pension Insurance, the National Tax and Customs Administration, the

National Employment Service, and the Educational Authority of Hungary. The administrative data was processed by the Databank of the Centre for Economic and Regional Studies. The calculations and the conclusions within the document are the intellectual product of Anna Adamecz as the author. The author gratefully acknowledges financial support from the Hungarian National Scientific Research Program (OTKA), Grant No. 128850. I am grateful to Gábor Kézdi for his advice on how to proceed with this research. I thank Ágnes Szabó-Morvai, John Jerrim, Nikki Shure, Hessel Oosterbeek, Robert Lieli, Marta Lachowska, Ádám Szeidl, Zoltán Hermann, Dániel Horn, and Júlia Varga for their comments to earlier versions of this paper, as well as to the editor and two anonymous referees whose suggestions improved the paper greatly.

Appendix A

Fig. A1; Fig. A2; Fig. A3; Table A1; Table A2; Table A3; Table A4; Table A5; Table A6; Table A7, Table A8, Table A9, Table A10, Table A11, Table A12

Appendix B: Empirical support for the identification assumptions

B1: No manipulation of school enrolment

This subsection provides evidence to support the no-defiers assumption I made in Section 4. I check whether there is a sign of defiance the following way. As mentioned in the Data section, the 2001 Census registers which grade of elementary school those born in 1990/1992 were attending in the spring of 2001. Knowing their grade in 2001 allows me to infer when they must have started school and thus estimate the magnitude of the jump in the probability of starting school at age 7 around the cutoff. However, grade retention is possible and not captured by the Census. In aggregate administrative data, the share of grade repeaters is around 4% in Grade 1 and at or below 2% in Grades 2-4, and these the period (Table A3 in Appendix A). This allows me to assume that the reform did not affect grade retention patterns in the first four grades of elementary school, at ages between 6/7 and 9/10.

Looking at which grade these cohorts attended in 2001, we know with certainty that if someone in the reform cohort was in the 4th grade in the spring of 2001, they must have started school in 1997, or earlier, before the reform. Those in the 3rd or a lower grade either started school in 1998, or later, after the reform, or they started school earlier and repeated grades. Similar logic applies to the comparison groups. Since some of those in Grade 3 in the reform cohort (or in comparable grades in the comparison cohorts) are grade repeaters who started school before the reform (or earlier in the comparison cohorts), I am not able to estimate the exact magnitude of the jump in the probability of being treated. Therefore, I do not use the actual size of these jumps and estimate ITT effects only.

Fig. B1 shows the inferred probability of starting school at age 7 among those born in 1990-1992. Those born below the cutoff are either late starters (as they should have started school at age 6 according to the enrolment rule), or grade repeaters. Those born above the cutoff are compliers to the enrolment rule. Among those born in the reform cohort, in 1991, the probability of starting school at age 7 jumps from 54.0% [53.0;54.9] to 87.8% [87.1; 88.4] on average and from 71.8% [67.4; 76.2] to 94.2% [91.8; 96.6] among the Roma. Below the cutoff, the share of late starters (and potential grade repeaters) is increasing over time as parents became more and more likely to delay school enrolment. The age 7 enrolment rate of the reform cohort below the cutoff fits into this trend.

Looking at the estimated sizes of the jump in the probability of starting school at age 7 (RDD coefficients, see the methods in the next subsection), they are not significantly different from each other among those born in 1990-1992 (Fig. B2). Thus, the inferred age 7 enrolment rates of those in the reform cohort do not show signs of defiance.

B2: No effects on migration

This subsection tests the effect of the reform on migration, proxied by whether one disappeared from the Admin3 administrative database. As detailed in the Data section, I construct a binary variable that equals one if (1) one did not appear in the administrative data at least once for at least 12 consecutive months and (2) never appeared again until October 2011 and 2016.

Fig. B3 shows the average probability of disappearance from the administrative data by month of birth among those born in 1990-1992. Fig. B4 plots the estimated RDD coefficients for the treated and comparison cohorts according to Equation (1) while Fig. B5 presents the estimated DRDD coefficients according to Equation (2). These results suggest that the reform did not affect the probability of migration.

Appendix C: Detailed results from the Census data

Fig. C1, Fig. C2, Fig. C3, Fig. C4, Fig. C5, Fig. C6, Fig. C7, Fig. C8, Fig. C9

Appendix D: Detailed results from the Admin3 database

Fig. D1, Fig. D2, Fig. D3, Fig. D4, Fig. D5, Fig. D6, Fig. D7

Appendix E: Channels and mechanisms

Fig. E1, Fig. E2, Fig. E3, Fig. E4, Table E1, Table E2, Table E3

References

Abadie, A., et al. (2023). When should you adjust standard errors for clustering?*. *The Quarterly Journal of Economics, 138*(1), 1–35. https://doi.org/10.1093/qje/qjac038. Available at:.

Adamecz-Völgyi, A., et al. (2021). *The labor market and fertility impacts of decreasing the compulsory schooling age* (p. 34). Institute of Economics Centre for Economic and Regional Studies. KRTK-KTI WORKING PAPERS 2021/40.

Adamecz-Völgyi, A., & Scharle, Á. (2020). Books or babies? The incapacitation effect of schooling on minority women. *Journal of Population Economics, 33*(4), 1219–1261. https://doi.org/10.1007/s00148-020-00771-9. Available at:.

Adato, M., & Hoddinott, J. (2010). *Conditional cash transfers in Latin America*. IFPRI Books. International Food Policy Research Institute (IFPRI). Available at: https://econpapers.repec.org/bookchap/fprifprib/9780801894985.htm (Accessed 13 July 2020).

Altwicker-Hámori, S., & Köllő, J. (2012). Whose children gain from starting school later? Evidence from Hungary. *Educational Research and Evaluation: An International Journal on Theory and Practice, 18*(5), 459–488.

Anderson, D. M. (2014). In school and out of trouble? The minimum dropout age and juvenile crime. *The Review of Economics and Statistics, 96*(2), 318–331.

Anderson, D. M., Hansen, B., & Walker, M. B. (2013). The minimum dropout age and student victimization. *Economics of Education Review, 35*, 66–74. https://doi.org/10.1016/j.econedurev.2013.03.005. Available at.

Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics, 106*(4), 979–1014. https://doi.org/10.2307/2937954. Available at:.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2011). Too young to leave the nest? The effects of school starting age. *The Review of Economics and Statistics, 93*(2), 455–467.

Cabus, S. J., & De Witte, K. (2011). Does school time matter?—On the impact of compulsory education age on school dropout. *Economics of Education Review, 30*(6), 1384–1398.

Calonico, S., et al. (2017). rdrobust: Software for regression-discontinuity designs. *Stata Journal, 17*(2), 372–404.

Cheng, M.-Y., Fan, J., & Marron, J. S. (1997). On automatic boundary corrections. *The Annals of Statistics, 25*(4), 1691–1708.

Cratty, D. (2012). Potential for significant reductions in dropout rates: Analysis of an entire 3rd grade state cohort. *Economics of Education Review, 31*(5), 644–662. https://doi.org/10.1016/j.econedurev.2012.04.001. Available at.

Cygan-Rehm, K., & Maeder, M. (2013). The effect of education on fertility: Evidence from a compulsory schooling reform. *Labour Economics, 25*, 35–48.

Devereux, P. J., & Hart, R. A. (2010). Forced to be rich? Returns to compulsory schooling in Britain*. *The Economic Journal, 120*(549), 1345–1364. https://doi.org/10.1111/j.1468-0297.2010.02365.x. Available at.

Gábos, A., et al. (2006). Anyagi depriváció Magyarországon, 2009–2015. *Társadalmi Riport, 14*(1), 24.

Green, C., & Paniagua, M. N. (2012). Does raising the school leaving age reduce teacher effort? Evidence from a policy experiment. *Economic Inquiry, 50*(4), 1018–1030. https://doi.org/10.1111/j.1465-7295.2011.00386.x. Available at:.

Grembi, V., Nannicini, T., & Troiano, U. (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics, 8*(3), 1–30. https://doi.org/10.1257/app.20150076. Available at:.

Grenet, J. (2013). Is extending compulsory schooling alone enough to raise earnings? Evidence from French and British Compulsory Schooling Laws*. *The Scandinavian Journal of Economics, 115*(1), 176–210. https://doi.org/10.1111/j.1467-9442.2012.01739.x. Available at:.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica, 69*(1), 201–209. https://doi.org/10.1111/1468-0262.00183. Available at:.

Hajdú, T., Kertesi, G., & Kézdi, G. (2014). Roma fiatalok a középiskolában. Beszámoló a TÁRKI Életpálya-felmérésének 2006 és 2012 közötti hullámaiból. In T. Kolosi, & I. G. Tóth (Eds.), *Társadalmi riport, 2014* (pp. 265–302). Budapest: TÁRKI. Available at: http://real.mtak.hu/20255/ (Accessed 8 October 2019).

Hermann, Z. (2013). *Are you on the right track? The effect of educational tracks on student achievement in upper-secondary education in Hungary* (p. 57). Institute of Economics, Centre for Economic and Regional Studies. Budapest Working Papers on the Labour Market 6.

Hong, K., Dragan, K., & Glied, S. (2019). Seeing and hearing: The impacts of New York City's universal pre-kindergarten program on the health of low-income children. *Journal of Health Economics, 64*, 93–107. https://doi.org/10.1016/j.jhealeco.2019.01.004. Available at:.

Imbens, G., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics, 142*(2), 615–635. https://doi.org/10.1016/j.jeconom.2007.05.001. Available at.

Kazuska, M. (2012). A tankötelezettség múltja, jelene és jövője. *Miskolci Jogi Szemle, VII* (1), 128–142.

Kemény, I., & Janky, B. (2005). Roma population of Hungary 1971-2003. *East European Monographs*, (702), 70.

Kertesi, G. (2005). *A társadalom peremén - Romák a munkaerőpiacon és az iskolában.* Budapest: Osiris Kiadó.

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

Kolesár, M., & Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review, 108*(8), 2277–2304. https://doi.org/10.1257/aer.20160945. Available at:.

Ladányi, J., & Szelényi, I. (2002). Cigányok és szegények Magyarországon, Romániában és Bulgáriában. *Szociológiai Szemle, 2002*(4), 72–94.

Landis, R. N., & Reschly, A. L. (2010). An examination of compulsory school attendance ages and high school dropout and completion. *Educational Policy*. https://doi.org/10.1177/0895904810374851 [Preprint]. Available at:.

Lee, D. S., & Card, D (2008). Regression discontinuity inference with specification error. *Journal of Econometrics, 142*(2), 655–674.

Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *The Review of Economic Studies, 72*(1), 189–221.

Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review, 94*(1), 155–189. https://doi.org/10.1257/000282804322970751. Available at:.

Mártonfi, G. (Ed.). (2011a). *A 18 éves korra emelt tankötelezettség teljesülése és (mellék) hatásai*. Budapest: Oktatáskutató és Fejlesztő Intézet.

Mártonfi, G. (Ed.). (2011b). *Hány éves korig tartson a tankötelezettség? Válaszkísérlet egy rossz kérdésre: Szakpolitikai javaslat*. Budapest: Oktatáskutató és Fejlesztő Intézet.

Meghir, C., & Palme, M. (2005). Educational reform, ability, and family background. *The American Economic Review, 95*(1), 414–424.

Milligan, K., Moretti, E., & Oreopoulos, P. (2004). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics, 88*(9–10), 1667–1695.

National Institute of Public Education. (2011). *Jelentés a magyar közoktatásról 2010. ("Report on the Hungarian Public Education 2010; in Hungarian.")*. Budapest: Oktatáskutató és Fejlesztő Intézet.

OECD. (2014). *PISA 2012 results in focus: What 15-year-olds know and what they can do with what they know* (pp. 1–41). OECD.

OECD. (2015). *Education policy outlook: Hungary* (pp. 1–27).

Oosterbeek, H., & Webbink, D (2007). Wage effects of an extra year of basic vocational education. *Economics of Education Review, 26*(4), 408–419.

Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics, 91*(11), 2213–2229. https://doi.org/10.1016/j.jpubeco.2007.02.002. Available at.

Pischke, J.-S., & von Wachter, T. (2008). Zero returns to compulsory schooling in Germany: Evidence and interpretation. *The Review of Economics and Statistics, 90*(3), 592–598.

Sebők, A. (2019). *The panel of linked administrative data of CERS Databank: A KRTK Adatbank Kapcsolt Allamigazgatási Paneladatbázisa*. Institute of Economics, Centre for Economic and Regional Studies, 2.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355–374. https://doi.org/10.2307/1882010. Available at.