Equal Access to Education: An Evaluation of the Roma Teaching Assistant Programme in Serbia

M. Battaglia and L. Lebedinski

Discussion Paper 2014-21

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain

UCL

Equal Access to Education: An Evaluation of the Roma Teaching Assistant Programme in Serbia*

Marianna Battaglia† Lara Lebedinski‡

July 7, 2014

Abstract

This paper investigates the effects of a remedial education programme - the Roma Teaching Assistant Programme - targeting the socially excluded and marginalised Roma ethnic minority in Serbia. By using first-hand collected data, we find evidence that children exposed to the programme went more to school. We do not find an effect on dropouts or marks for all grades. An examination of heterogenous effects suggest that children in the first grade benefitted more from the programme as compared to their older peers through lower dropouts and better grades. Overall, our results suggest that well-targeted remedial education programmes can be successful in boosting outcomes of low performers.

Keywords: primary education, remedial education, Roma, ethnic minority, absences, dropouts

JEL classification codes: I21, J15, D04

*We would like to thank Tito Boeri, Matz Dahlberg, Eliana La Ferrara, Andreas Madestam, Michele Pellizzari, Petra Thielmann and participants at AMID summer school 2010, Bocconi University 2011 Development Reading Group, Bocconi University PhD seminar 2011, Spring Meeting of the Young Economists 2012, Third International Workshop on Applied Economics of Education 2012 and CESifo Economic Studies 2012 conference. We are grateful to the personnel of the Serbian Ministry of Education, especially to Božidar Jovanović and Tünde Kovacs-Cerović, for giving us the authorisation for this project and for answering all our questions. We are indebted to Nikola Duvnjak from OSCE for answering all our questions about the pilot phase of the programme. This paper is produced as part of the project Actors, Markets, and Institutions in Developing Countries: A micro-empirical approach (AMID), a Marie Curie Initial Training Network (ITN) funded by the European Commission under its Seventh Framework Programme Contract Number 214705 PITN-GA-2008-214705. The usual disclaimer applies.

†Departamento de Fundamentos del Análisis Económico, University of Alicante, 03690 San Vicente del Raspeig, Alicante, Spain, Tel: +34 965 90 32 18, Email: mbattaglia@ua.es

‡Corrsponding author, IRES, Université catholique de Louvain, Department of Economics, 3 Place Montesquieu, 1348 Louvain-la-Neuve, Belgium, Tel: +32 104 74 337, Email: lara.lebedinski@uclouvain.be

1 Introduction

Roma are mainly located in South Eastern Europe and with a population of approximately 6 million people they constitute the largest ethnic minority in the continent (Open Society Institute, 2008).[1] They experience severe social exclusion in terms of high poverty and unemployment levels, low educational attainments and no participation in the political and cultural life. Roma people are poorer than other population groups and more likely to fall into poverty and remain poor. They have persistent disadvantages in education, including low school attendance and overrepresentation in special schools and schools for adult education.[2] Roma are often employed in low skilled jobs where they earn below average wages. They often lack access to credit and property ownership and are overdependent on social benefits.

The problems of the Roma minority have become evident with the expansion of the European Union. The visa liberalisation and the adhesion to the Union of countries like Romania and Bulgaria - where the percentage of Roma population is high - have indirectly led to migration flows towards Western Europe. The receiving countries encountered new problems: increased number of asylum requests, appearance of informal settlements, increased number of unemployed and inadequacy of the education system to accommodate new foreign pupils. The extraordinariness of the phenomenon led to hot discussions within the European countries and civil society. It increased the interest of the EU in the EU candidate countries with a high percentage of Roma population, e.g. Serbia. Understanding the impact of programmes targeting Roma is useful not only for the sending countries, but also for the receiving ones.

Schooling has always been considered a needed measure to improve living conditions of Roma people and to foster their integration: higher enrolment rates and better achievement at school are expected to lead to persistent effects in the labour market and in the reduction of poverty in the long-run. The goal of this paper is to evaluate the impact of the main programme targeting Roma inclusion in education in South Eastern Europe in the first year of its introduction: the Roma Teaching Assistant Programme (RTA).[3]

[1] The number of Roma and the subsequent numbers refer to the following countries: Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Kosovo, Latvia, Lithuania, Republic of Macedonia, Moldova, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia and Ukraine.

[2] The net enrolment rate of Roma in primary education varies among the countries and it is still low, in the range of 40% to 60%. Moreover, students may enrol at the beginning of the year, but may not actually attend school: the percentage of completion rates of primary school is in the range of only 30% to 40% in most countries in 2008 (Open Society Institute, 2008). Special schools are schools for children with special educational needs. Schools for adult education were initially introduced with the idea to provide basic literacy knowledge to adult pupils. Nowadays, however, they are mainly attended by pupils who enrolled at school late and by pupils who decided to return to school after dropping out.

[3] The Serbian name of the programme is *Romski Asistenti - Pomoć u Nastavi*. The Ser-

Roma assistants - one per each school - participate in regular lessons where they provide additional help to Roma pupils who have difficulties in following classes. They organise additional lessons, help them with their homework and assignments and once per week they visit their parents. We examine if this remedial education programme reduces dropouts, raises attendance and improves the marks of Roma pupils. This study uses primary collected data during 5 months in the Summer-Autumn 2010. We employ two different econometric strategies and their combination. First, we exploit the gradual implementation and the intensity of the programme in order to base the evaluation of its impact on a comparison of *Early* and *Late Enrollees*. Second, we compare children exposed to the programme to older cohorts less exposed to it. There is evidence that all children exposed to the programme went on average more to school. We also find evidence that marks improved in mathematics and Serbian for first graders. Higher impacts are obtained in schools with a lower number of Roma: the higher is their number, the lower the impact of the programme on the outcomes of interest. This is especially the case for girls, for whom being in a school with a lower number of Roma turns out to be more favourable. Boys respond to the programme with fewer absences in schools with fewer Roma.

Two main mechanisms may help to explain the empirical evidence. First, the intervention alters the inputs in the education production function by providing more instruction time to pupils. The second potential mechanism works through cultural transmission of preferences for education. For an ethnic group which has extremely low educational outcomes in different countries, it is possible that this fact can be, at least partly, explained by preferences for education. In the case of RTA, the provision of a role model, a person with the same Roma background, can affect preferences for education of both children and parents and induce children to reduce absences and exert more effort at school.

There are no studies in economic literature that investigate how to improve life circumstances of Roma, in general, and Roma kids, in particular. This paper contributes to the existing literature by providing an accurate overview of the attainments of Roma pupils in Serbian schools, for which so far there were no data available, and by contrasting their achievement to the average Non Roma pupils. More importantly, it adds evidence on short-term effects of remedial education targeting a stigmatised minority group and suggests replicable examples in contexts where minorities suffer

bian Government - together with Montenegro, Croatia, Macedonia, Hungary, Romania, Bulgaria, the Czech Republic and Slovakia - is participating in the *Decade of Roma Inclusion*, an international initiative running from 2005 to 2015 in Central and South-Eastern Europe. The initiative brings together governments, international and non-governmental organizations to improve the welfare of the Roma population, focusing on healthcare, education, employment and housing. Examples of other programmes which introduce Roma school assistants can be found in Czech Republic, Slovakia, Bulgaria and Croatia.

low attainment rates and social exclusion. For Roma people this is the case in many other European countries.

Related Literature This paper adds to the literature on remedial education programmes. Policies targeting low-performing students are generally difficult to evaluate because children with learning difficulties are not randomly assigned to programmes. Their characteristics affect both the selection into the programme and its success, making it difficult to distinguish between the two effects, especially because the selection mechanism is not typically fully observable. A few studies are able to overcome the identification problem and they find support for the effectiveness of remedial education in the short run.[4] Aside from short term-effects, there are two very important questions for policy makers. Does the effect fade out after the programme? Is the effect of remedial education cumulative? Banerjee et al (2007) find that only a year after the end of a remedial education programme, its effect faded largely out. In contrast, children participating in the programme for two consecutive years, could accumulate knowledge and increased their test scores in both years. This suggests that gaps can be reduced, but policy makers should aim at long-term interventions.

This study speaks also to the literature on programmes aiming at improving schooling outcomes of minority communities and at narrowing differences between racial groups. The black-white achievement gap and its evolution has been intensively investigated in the United States.[5] Fryer (2010) demonstrates that today's racial differences in social and economic outcomes are greatly reduced when one accounts for educational achieve-

[4]See for instance Lavy and Schlosser (2005); Jacob and Lefgren (2004); Hanushek, Kain, and Rivkin (2002); Banerjee, Cole, Duflo, and Linden (2007).

[5]The first and most known programme attempting to close the racial gap is the Perry Preschool programme introduced in 1962: it targeted children from disadvantaged socioeconomic backgrounds and consisted of a 2-5-hour daily preschool programme for children aged three years old and weekly home visits by teachers. Schweinhart, Barnes, and Weikart (1993) find that students in the programme had higher marks between the ages of 5 to 27, 21% less grade retention or special services required and 21% higher graduation rates. Other interventions for disadvantaged families followed such as the Abecedarian Project in the '70s, which provided childcare services for four cohorts of children from infancy through age five, and the Early Training Project, consisting in summertime experiences and weekly home visits during the three summers before entering first grade. Attempts have been also made during the primary school through the introduction of after-school programmes (Lauer, Akiba, Wilkerson, Apthorp, Snow, and Martin-Glenn, 2006), merit pay for principals, teachers, and students (Podgursky and Springer, 2007; Fryer, 2010), professional development for teachers (Boyd, Grossman, Lankford, Loeb, and Wyckoff, 2009), and getting parents to be more involved (Domina, 2005), placing disadvantaged students in better schools through desegregation busing (Angrist and Lang, 2004) or altering the neighborhoods in which they live (Jacob, 2004; Sanbonmatsu, Kling, Duncan, and Brooks-Gunn, 2006). The evidence on the efficacy of these interventions is mixed: certain programmes have left the racial achievement gap essentially unchanged. A good overview of the evolution of the black-white gap is given in Neal (2006).

ment and poverty levels. This points to the fact that there is little empirical evidence for discrimination in the recent data. A similar finding emerges in the study of Kertesi and Kezdi (2011) on Roma educational achievement in Hungary. They find that the gap between Roma and Non Roma is substantially larger than the gap between African Americans and whites in United States, but that accounting for health, parenting, school and class fixed effects, and family background, the achievement gap disappears in reading and decreases by 85% in mathematics.

The third relevant strand of the literature is on programmes aiming at achieving better schooling outcomes of the poor. Recent intervention targeting children with low socio-economic background have tried to affect the demand for education. A prominent type of demand side programmes are conditional cash transfers which raised enrolment and attendance in many developing countries. However, policies that promote school enrolment may not promote learning: early contributions indicate that programmes which are effective at reducing absence from school often do not have an impact on achievements of the average student (Schultz, 2004; Miguel and Kremer, 2004). Analogously, Fryer (2011a), through school-based randomized trials in schools designed to test the impact of incentives on student achievement, shows that incentives can raise achievement among even the poorest minority students in the lowest performing schools only if the incentives are given for certain inputs, such as reading books. Providing incentives for achievement in marks is much less effective. Finally, only providing school books and other school material or subsidised school meals does not seem to improve students achievements in the case of students with weaker academic backgrounds (Glewwe, Kremer, and Moulin, 2009; Vermeersch and Kremer, 2005).

The rest of the paper is organised as follow. Section 2 summarises the Roma Teaching Assistant Programme and describes our data. Sections 3 explains our empirical strategy and presents our results. Section 4 discusses the findings and concludes.

2 Description of the programme and preliminary analysis

2.1 The Roma Teaching Assistant Programme

The Roma Teaching Assistant Programme started as a pilot programme implemented by various NGOs in 2002. In 2007 the OSCE took over its coordination and financing. In 2009 the programme started to have a country wide coverage and it is under the coordination of the Ministry of Education. In the scholastic year 2009/2010 48 primary schools had a Roma assistant: 22 schools started with the programme at different points of time between

2002 and 2007; 26 schools started in 2009. The Ministry expanded the programme to other 77 schools starting from November 2010.

Based on when the programme started in a school, the schools can be divided in two groups: schools entering the programme in September 2009 (*Early Enrollees*) and schools entering the programme in September 2010 (*Late Enrollees*). The 22 schools, which joined the programme between 2002 and 2007, are excluded from our analysis. The selection of these schools was not centralised; they were chosen by NGOs based on the share of Roma pupils. For our analysis we will consider only 26 *Early Enrollees* and 77 *Late Enrollees*.

Both schools and potential Roma assistants had to apply to participate in the programme. Among 78 schools that applied in 2009, a commission representing the government institutions together with OSCE representatives, chose 26 *Early Enrollee* schools based on the percentage of Roma students (between 5% and 40%) and preferably, the availability of preschool programme in the school.[6] The requirements for Roma assistants were knowledge of Romani, secondary school diploma and experience in working with children. 158 candidates applied for 26 assistant positions.[7] In 2010 the programme was renamed to *Education for all* and starting with the scholastic year 2010/2011 Roma teaching assistants were renamed to pedagogical assistants. In year 2010, 252 schools applied for 77 assistants. Similar to the first round, schools were eligible only if they had between 5% and 40% of Roma. The only difference between the two rounds was that in the second round the availability of a preschool programme was not considered. The reason is that in 2010/2011 pedagogical assistants started to work in 50 kindergartens offering compulsory preschool programme. Schools not offering the preschool programme could have then been close to kindergartens offering it. The Roma pupil would have been helped by an assistant from her entry in the school anyhow. One could argue that this small change in requirements could lead to a selection bias in the two rounds, but our data do not support this claim.[8] Selection criteria for now pedagogical assistants remained unchanged and out of the 329 applications for the position, 77 were

[6] 64 out of 78 schools that applied had a percentage of Roma between 5% and 40%. Among these 64, OSCE selected 19 schools (out of 26) with a preschool programme, 5 schools (out of 36) with no preschool programme and 2 schools (out of 2) for which no information is available.

[7] The following criteria were taken into account for the ranking of the assistants' applications: highest level of education completed or enrolled (from 10 to 30 points), experience in working with Roma children (0 to 10 points), experience in working on projects related to education (0 to 10 points), motivation (0 to 10 points), attendance of relevant seminars and/or courses (0 to 10 points), experience as Roma teaching assistant (0 to 10 points), knowledge of Romani (0 to 10 points) and additional points (0 to 10 points).

[8] Unfortunately we do not have information on the availability of a preschool programme for schools applying in 2010/2011. Nonetheless, it is worthy to recall that some schools without the compulsory preschool programme have also been selected in the previous year.

accepted to work at schools and another 50 were accepted for kindergartens.

Schools got to know that they would receive an assistant in early June. They did not inform parents about the presence of the RTAs. The programme did not receive publicity from TV and radio. This leads us to believe that parents were not aware of the existence of the RTA before enrolling their children at school. Data also confirm that *Early Enrollees* did not attract more Roma students than *Late Enrollees* in the first year of the programme.[9] There is not selection of children into schools.

Every school received only one assistant. Schools received from the Ministry a description of her duties, but they were free to decide how to allocate the time of the assistant depending on the needs of the school.[10] Activities at school involve both working during regular classes and after-school work. Work with local communities comprises duties such as collecting information about children who did not enrol or who left school, gathering documents for school enrolment, visiting families,[11] cooperation with Roma NGOs, etc. The assistants were advised to work mainly with lower grades, especially the first. Their objectives were to ensure that children go to school, to prevent them from dropping out and to help them to succeed at school. In 2009 the Ministry of Education organised a series of seminars with the goal of providing the necessary knowledge and skills to Roma teaching assistants.[12]

2.2 Preliminary Analysis

2.2.1 Data and Trends of the Variables

We use primary data collected during 5 months in the summer/autumn 2010. School data was not available in a digital format and we visited schools and collected in person data from administrative records. Our dataset contains information on 23 schools (out of 26) among *Early Enrollees* and 15 schools (out of 77) among *Late Enrollees*.[13] Given that the data collection process

[9]Roma pupils joining *Early Enrollees* schools in the pretreatment year - 2008/2009 - corresponded to 2.4% of all Roma enrolled in these schools. In *Late Enrollees* they were 2.1%. In the first year of the programme - 2009/2010 - these percentages were 1.6% and 1.3%. The number of Roma pupils enrolling at school for the first time reduced proportionally between the two years in both types of schools.

[10]There was a suggested time allocation. The 30 weekly hours of the assistant could be distributed in the following way: work at school (19 hours), work with the local community (8 hours) and writing reports and documentation (3 hours).

[11]In most cases Roma live in segregated settlements so that assistants can go to the settlement and visit several families at once.

[12]In total, the assistants attended 22 working days of seminars and courses in the scholastic year 2009/2010. Regular seminars provided the opportunity to the Ministry to understand the problems of the assistants and guide them through the initial difficulties. In 2010/2011 a set of 9 modules, which all assistants had to attend in their first year of service, was devised.

[13]In total, there were 26 schools which got an assistant in 2009/2010. In 3 schools we were not allowed to collect data. These schools do not differ from the other schools either

was costly, our dataset consists only of a subsample of *Late Enrollees*. We selected 15 *Late Enrollees* schools according to the following criteria: first, they had to be in the same district of an *Early Enrollees* school;[14] second, they had to be in a rural/urban municipality as the nearby *Early Enrollees* school; third, they had to share a similar school size to the nearby *Early Enrollees* school and finally, a similar percentage of Roma pupils.[15]

Schools are mainly in Belgrade/Central Serbia and in the South/South-Eastern part of the country, and they are equally distributed in rural and urban areas.[16] Figure 1 reports the distribution of schools in our sample.

The data set contains information on 4 scholastic years − from 2006/2007 to 2009/2010 − for the lower four grades of primary school for 18,268 Roma and Non Roma children. It contains for each year and for each pupil the final mark in mathematics, final mark in Serbian and number of hours of absences in a year. The data set contains individual characteristics, such as gender, year of birth, month of birth and place of birth.[17] School specific data include school size, number of Roma - in both school and class - and whether the school is in an urban setting.

Roma children Table 1 shows summary statistics of the control variables and main outcomes of interest for Roma children in the pre- and treatment year. In the pre-treatment year the mean characteristics of the schools that were enrolled in the programme later (column (2), table 1) resemble those of the schools that enrolled first (column (1), table 1). Boys and girls are equally distributed in both groups of schools and pupils were mainly born in the same town where they attend school. A fifth of all students in the schools are Roma. The number of Roma per class is roughly 4-5 children and the class size is 22.161 in *Early Enrollees* and 23.966 in *Late Enrollees*. The table shows no statistically significant differences between *Early Enrollees* and *Late Enrollees* nor in the student's and school characteristics nor in the outcomes of interest. This similarity between *Early Enrollees* and *Late*

in the number of pupils or in the percentage of Roma children and they are located in different areas: one in Belgrade, one in Central Serbia and one in the South.

[14]A district is made up by more municipalities. In Serbia there are 24 districts and 160 municipalities.

[15]In few cases the school chosen was not available and we needed to select the second option.

[16]10 schools are located in Belgrade; 8 schools in the central area of the country (5 schools in the municipality of Valjevo and 3 in the municipality of Novi Sad); 12 schools in South-Eastern Serbia (3 schools in the municipality of Jagodina, 2 in Kragujevac, 3 in Kruševac, 3 in Zaječar and 1 in Požarevac); 8 schools in the South of the country (6 schools in the municipality of Leskovac and 2 in the municipality of Niš). We define *urban area* a municipality with more than 35,000 inhabitants.

[17]Roma in Serbia are mainly sedentary: they do not move much within the country. Nonetheless, there is a substantial out-migration, especially towards the European Union, and in the last years in-migration has increased due to the wars in Ex-Yugoslavia. Many Roma refugees in Serbia come from Kosovo.

Enrollees schools is also found in the treatment year, providing support for our claim that *Early Enrollees* and *Late Enrollees* are comparable.[18]

By simply comparing average outcomes in the two types of schools, we see that dropouts almost double in the last year in both types of school and absences increase in 2009/2010 in both *Early Enrollees* and *Late Enrollees*, but they increase by less in *Early Enrollee* schools. We believe that both the increase in dropouts and the increases in absences is related to the liberalisation of the visa regime with the European Union. This regime change induced a considerable number of Roma families to migrate to the EU. Finally, we see a minor improvement in all marks from pre- to treatment year. This effect is larger in *Early Enrollees* than in *Late Enrollees*.

Non Roma children In our analysis we focus on Roma children and we use for all of our estimations only the sample of Roma children.[19] Table 2 shows summary statistics of the control variables and main outcomes of interest for Non Roma children in the pre- and treatment year. No statistically significant differences are found neither in the pre- nor in the treatment year in the pupils' characteristics or outcomes. The only statistically significant difference is in the dropout rate, which however is close to zero and not problematic in the lower primary school grades.

The differences between Roma and Non Roma children are striking. Three important aspects need to be stressed when comparing them. First, dropouts are almost exclusively of Roma children. Moreover, Roma children are absent from school approximately three to four times more than Non Roma children. Lastly, on a grading scale of 1 to 5, the difference of almost two marks between Roma and Non Roma pupils in Serbian and mathematics is substantial.

3 Econometric Strategy

We want to examine the impact of the Roma Teaching Assistant Programme on dropouts, attendance and marks of Roma pupils in the first year of its implementation. In the analysis we use only Roma children. We exploit the gradual implementation of the programme. Our treatment group consists of schools which started to implement the programme in September 2009 (*Early Enrollees*), whereas the control group is a subsample of schools which got the assistants starting from November 2010 (*Late Enrollees*).We argue that the selection of schools to enter the programme earlier/later can be treated "as if random" for several reasons. First, the selection criteria

[18]The only significant difference is found for the place of birth: there are less migrant children in treated schools.

[19]The only exception is section on spillover effects in the appendix.

remained almost the same in both rounds.[20] Second, one could argue that schools applying in the first round were more motivated, but schools which applied in the first year could also apply in the second year. However, some schools which applied in 2009 did not apply anymore in the year after.[21] Thus, if they really were more motivated and of better quality, it is hard to understand why they did not want to be part of the programme anymore in 2010. Third, we do know that observable characteristics do not differ between schools applying in the first year and schools applying in the second year.[22] Fourth, we are in possession of data for 3 years prior to the introduction of the programme and the placebo tests for these years support our claim of no systematic differences between the two groups of schools. Fifth, we do know that the committee for schools selection - composed of the Minister of Education and other representatives of the Ministry, representatives of National Council, OSCE and of the Ministry for Human and Minority Rights - rated schools based on their shown interest and motivation (application) in the same way, in both years.[23]

A second possible estimation strategy is to exploit the fact that older cohorts were less exposed to the programme (control group) to younger cohorts (treated group) in *Early Enrollees*-treated schools.[24]

The main advantage of using as a control group schools which enrolled later in the programme is that its impact would not be confounded with other government policies that took place in the year of its introduction. For instance, in 2009/2010 all first grade pupils got free text books and in the last few years the Ministry strongly suggests to schools to reduce repetition rates especially in the lower grades. The weakness of this control group is that we cannot be completely certain that unobservable characteristics are the same in *Early Enrollee* and *Late Enrollee* schools. In order to purge time-invariant school characteristics, we can use older cohorts in the treatment schools as a control group. Nonetheless, this econometric strategy relies on the strong assumptions that there were no government interventions over the period - which is not exactly our case - and that the outcomes have a regular trend

[20] In both rounds the programme was advertised in newspapers *Politika* and *Prosvetni Pregled*, the last being a newspaper for people working in the education sector; in addition to the advertisement, in 2010/2011 schools' directorates - one directorate may be responsible for more than a municipality - informed schools directly.

[21] 47% of schools which applied in 2009, and did not get selected, did it again in 2010 and two thirds of them got selected in the second year (16 out of 26 schools applying in both years). Among these schools, only 2 schools, corresponding to 12.5%, is present in our subsample.

[22] Schools which applied in 2010 are in the same areas of schools of 2009 and they have almost the same percentage of Roma, on average 13.99% compared to 13.07%. They are not statistically different (p-value=0.458). These are the only information we have on schools which applied and did not get selected.

[23] In both rounds the committee gave priority to schools in the poorest municipalities or with huge Roma settlements (Subotica, Novi Sad, Niš, Kragujevac, Belgrade).

[24] This method is similar in spirit to Duflo (2004).

over the years. By combining the *Early - Late Enrollees* analysis with the cohort specification we better take into account strengths and weaknesses of both approaches.

3.1 First Approach: Comparison of *Early Enrollees* vs. *Late Enrollees*

Our first econometric strategy exploits the fact that some schools received the assistants prior to other schools. We compare *Early Enrollee* schools with *Late Enrollee* schools in the years 2008/2009 - year before the introduction of the programme - and 2009/2010 - year of the introduction of the programme.

3.1.1 Average treatment approach

Our specification (1) is a difference-in-difference model with school fixed effects:

$$Y_{ijt} = \beta_0 + \delta_t + \rho_j + \beta_1 treatment_j * post_t + \beta_2 X'_{ijt} + \varepsilon_{ijt} \qquad (1)$$

The outcome variables Y_{ijt} are dropout, hours of absences and final marks in Serbian and mathematics of individual i, in school j at time t. δ_t is a time fixed effect, ρ_j corresponds to school fixed effects, and $treatment_j * post_t$ is the interaction term between the dummies for treatment status of the school and treatment year. Dropout is a dummy variable taking value one if child i dropped out of school during year t, otherwise it is equal to zero. With school fixed effects we are able to control for time-invariant unobservable school characteristics as well as unobservable geographical characteristics. The control variables X'_{ijt} are school size, school size squared, number of Roma in school, number of Roma in school squared, percentage of Roma per class, class size, class size squared, the gender of the child (=1 if the child is female), age, age squared, and whether the kid is a migrant (=1 if the child was born in the same town where she attends school). The coefficient of interest (β_1) is the difference-in-difference estimator of the interaction term between treatment and time that captures the difference in outcomes between the treatment and control schools.

Results for the different outcomes of interest are reported in table 3. For all outcomes we estimate the regressions without and with controls (columns (1) and (2)). We then split our sample by gender (columns (3) and (4)) to assess whether the impact of the intervention differs by gender.[25]

[25]There is vast evidence on gender gaps in education and their underlying causes. Buchmann, DiPrete, and McDaniel (2008) give a very good review of the literature. For this reason we also control for the gender of the assistant, but it does not turn out to be significant in any specification.

Overall, results show that the programme had a statistically significant impact only on hours of absences: pupils exposed to the programme were on average almost 17 hours less absent from school (0.121 standard deviations).[26] This is especially the case for male, whose reduction in absences is of roughly 26 hours (5 days) or equivalently 0.176 standard deviations. Dropouts and marks in both subjects are unaffected by the programme. At first it could seem surprising that dropouts do not respond to the programme, but our preliminary data analysis has shown that dropouts are not problematic in the lower 4 grades. The fact that the inclusion of control variables does not change the magnitude of our coefficients suggests that our coefficients are very robustly estimated.

There are only three settlements where children from both control and treatment schools live together and it is possible that in these settlements children in *Early Enrollee* and children in *Late Enrollees* schools interact. If one were to believe that there were spillover effects from treated children on children from control schools, this would imply that coefficients in our regressions are underestimated.

Placebo Regressions The difference-in-difference approach relies on the parallel trends assumption. We assume that, in the absence of the programme, treatment and comparison schools would have had a parallel trend in the average outcomes of interest. An obvious way to examine the robustness of our results is to estimate the same regressions (specification (1)) for the years 2006/2007 versus 2007/2008 and for the years 2007/2008 versus 2008/2009. By estimating the same regressions for pretreatment years, we can test if the outcomes in the two groups of schools were regular before the introduction of the programme. Significant difference-in-difference coefficients in placebo regressions would invalidate our estimation strategy and would question the adequacy of our comparison group.

Placebo tests are summarised in table 4 and we do not find significant coefficients. The diff-in-diff estimates for mathematics and Serbian are insignificant and negative for the period immediately preceding the introduction of the programme (2007/2008 versus 2008/2009). Even if one were to believe that there was a trend prior to the programme, then our estimates for all outcomes are underestimating the effect of the programme in the average treatment approach.[27] The placebo regressions support our claim that *Late Enrollees* are a good counterfactual for *Early Enrollees*.

[26] On average, Roma pupils are absent from school 143 hours (28 days) in a year.

[27] Remember that we expect the coefficient of *treatment*post* in our main regression to be positive for marks (Serbian and mathematics).

3.1.2 Intensity of treatment approach

The design of the intervention permits us to investigate whether the effect of the programme varies with the number of Roma per school. Each school has only one assistant: the higher the number of Roma per school, the less intense is the programme. If the assistant has to help a high number of students, she will help less each of them: she will be less present both in regular classes and in activities regarding their homework and assignments.

Our intensity specification is a variation of the previous approach; it still uses within school variation of Roma, but now we exploit also the variation in the number of Roma between schools. We divide schools in two equally sized groups with 19 schools.[28] The dummy *intensive* is equal to one for schools with fewer than 43 Roma.[29] The main difference to the prior model is that we interact the dummy *intensive* with treatment and time. We believe that differentiating the schools in groups helps to better understand the role of the number of Roma on the impact of the programme. The coefficient of interest is now β_6.

The intensity of treatment is modeled:

$$Y_{ijt} = \beta_0 + \delta_t + \beta_1 treatment_j + \beta_2 treatment_j * post_t + \beta_3 intensive\_Roma_{jt} + \tag{2}$$
$$+ \beta_4 intensive\_Roma_{jt} * post_t + \beta_5 intensive\_Roma_{jt} * treatment_j +$$
$$+ \beta_6 intensive\_Roma_{jt} * treatment_j * post_t + \varepsilon_{ijt}$$

Results with and without controls, and for both genders are reported in table 5. The intensity of the programme clearly plays a role in explaining its effects. The lower is the number of Roma in a school, and similarly the more the assistant can help them, the higher is the impact on the outcomes of interest. Absences, for instance, reduce on average by roughly 36 hours or equivalently 0.228 standard deviations in a year in schools with less Roma, compared to *Late Enrollee* schools. The fall in absences is driven by a reduction in absences of boys, almost 60 hours less. The effects disappear in schools with a higher number of Roma. Marks in both mathematics and Serbian increase for pupils in *Early Enrollee* schools with a lower number of Roma, but again these effects do not result in schools with a higher number of Roma. The impacts are especially large for females, for whom being in a school with a lower number of Roma seems to be more favourable: on average, if exposed to the programme in a school with less Roma, their marks

[28] The average school number of Roma between the two years - pre- and treatment year - is used to define the two groups. The threshold is here 43 pupils in a school. Our balance tests for the pretreatment year are reported in the Appendix (table 10).

[29] In A.2 we vary the threshold and we find that the effect of the programme is stronger, the lower the threshold. This is exactly what we would expect. Remember that this approach delivers the effect of the programme on schools for which we define, based on the threshold, that the programme is intensive. A lower threshold means that less Roma are in that school and for this reason the programme is more effective.

in mathematics and Serbian increase by 0.680 (0.581 standard deviations) and 0.499 (0.419 standard deviations). The improvement in test scores is larger in mathematics than in languages and this different response to interventions by subject has been found also in other studies (Fryer, 2011b). Possible explanations for the disparity in treatment effects by subject area are offered in Fryer (2011b). One theory suggests that language skills development occurs at an earlier age than the development of higher cognitive skills. Another prominent theory suggests that language acquisition and reading test scores are influenced by factors outside of the classroom. This is plausible because language skills are used more outside of the classroom than mathematics skills.[30]

Placebo Regressions We estimate placebo regressions of estimation (2) for the years 2006/2007 versus 2007/2008 and for the years 2007/2008 versus 2008/2009. Placebo tests are summarised in table 6. The coefficients for school grades are positive, but not significant for the period preceding the programme (2007/2008 versus 2008/2009). The coefficients we obtain for the treatment year are much larger than coefficients from our placebo regressions. For instance, in mathematics (Serbian) we obtain an increase of 0.417 (0.307) in marks in treatment year, and this number is only -0.011 (0.113) for the pretreatment placebo regression. Similarly, for absences the coefficient is negative, but with -7.490 hours much smaller than the coefficient -36.390 from the main regression. Again, the placebo regressions confirm that *Late Enrollees* are a good counterfactual for *Early Enrollees*.

3.2 Second approach: Cohort regressions and triple difference

We know that assistants worked mostly with the first grade[31] and in the second approach we compare kids in the first grade (young cohorts) with kids in older grades - second, third and fourth - (old cohorts) in the pre- and treatment year in treated schools. Here we assume that the difference in marks between first graders' and second, third and fourth graders' would have been constant over time, in the absence of the programme. We also implicitly assume the absence of policies targeting specific grades. With this econometrics strategy, we are able to purge time-invariant school characteristics.

The specification in this section informs us whether the programme was

[30] For the corresponding references from the education literature see Fryer (2011b).

[31] We know this fact from informal discussions with assistants. We have also confirmed this fact with survey data for Belgrade schools (only a subsample of schools) for which we have collected additional data. In treated schools, the children with whom the assistant actually worked were, on average, 9.73 years old, whereas children with whom the assistant did not work, were 10.44 years old (Battaglia and Lebedinski, 2013).

successful for the children enrolled in the first grade. We first estimate the following regression for *Early Enrollees*:

$$Y_{ijt} = \beta_0 + \beta_1 young_i + \beta_2 post_t + \beta_3 young_i * post_t + \varepsilon_{ijt} \tag{3}$$

where Y_{ijt} are again dropout, hours of absences and final marks in Serbian and mathematics of individual i, in school j and at time t; $young_t$ is equal to 1 when the child is at the first grade; $post_t$ is equal to 1 in the year of the treatment (2009/2010). The coefficient of interest is now β_3 which tells us how the first graders have performed compared to the older grades.

The same regression (3) is then estimated for *Late Enrollees* and the triple difference between treated and control schools and cohorts is captured by γ_3 in the following specification:

$$\begin{aligned} Y_{ijt} \quad &= \beta_0 + \beta_1 young_i + \beta_2 post_t + \beta_3 young_i * post_t + \gamma_1 treatment_j * post_t + \\ &+ \gamma_2 young_i * treatment_j + \gamma_3 young_i * post_t * treatment_j + \varepsilon_{ijt} \end{aligned} \tag{4}$$

Unlike estimation (3), regression (4) does not control for possible selection bias. By estimating cohort regressions on a pooled sample of *Early* and *Late Enrollees*, we can control for government policies targeting specific grades.[32]

The regressions are estimated with controls and we also inspect the impacts by gender. Results are in table 7.

When comparing first graders with older pupils in only *Early Enrollee* schools (column (1)), our coefficients of interest have, with the exception of absences, the correct sign, but are not statistically significant. We are not overly concerned that absences increase, because we observe a similar pattern also in control schools (see column (2) in table 7). We believe that the reason for the overall increase is that dropouts rise in 2009/2010 as a consequence of visa liberalisation because some families have migrated to EU countries.[33] Results of the triple interaction for the full sample are reported in column (3) in table 7. Dropouts are statistically significant. Absences decrease for the young cohort, but they do not reach statistical significance. Pupils exposed to the programme in the first grade get higher marks than first graders in control schools (with respect to their older mates). This effect is partly driven by the worse outcomes of first grade pupils in control schools. We speculate that this happens, because first graders are overall worse than first graders in the preceding year. On average, being in a *Early Enrollee* school increases marks in mathematics and Serbian by 0.284 and 0.296 standard deviations for first graders.

[32] For instance, free schools books for first grade or lower repetition rates in general.

[33] Dropouts are included in our regressions and have a high number of absences.

Placebo Regressions We test the robustness of our results by estimating placebo regressions (regression (3), (4)) for the years 2006/2007 versus 2007/2008 and for the years 2007/2008 versus 2008/2009 (see table 8). None of the relevant coefficients is significant. The size of the coefficient for Serbian for 2007/2008 versus 2008/2009 is insignificant, but not completely negligible. Absences for the two placebo tests are positive, but we obtain a reduction in absences for the treatment year. Similarly, the sign of the coefficient suggests a worsening of marks for Serbian for younger cohorts, exactly the opposite of what we find for the year of treatment.

4 Conclusion

In this paper we estimate the impact of the Roma Teaching Assistant Programme in its first year of implementation on schooling outcomes. We use a difference-in-difference approach by exploiting the gradual introduction of the programme. We argue that the assignment to enter the programme earlier/later can be treated as if random. As an additional check, we use a second econometric strategy. We compare pupils of the first grade from treated schools with older cohorts from the same schools. This econometric strategy controls well for school specific characteristics, but we are not able to control for government interventions over the observed period. We can combine the two approaches (*Early - Late Enrollees* with cohort analysis) and estimate the triple difference between young and old cohorts in treated and control schools.

Results of our analysis show that the programme had a positive effect and started to reduce the gap between Roma and Non Roma students both in school achievements and attendance. There is evidence that absences fell by 0.121 standard deviations in treated schools (17 hours, i.e. 3 to 4 days). The remedial education part of the programme targeted mainly first graders. Our analysis suggests that, for this group, marks have improved by almost 0.284 standard deviations in mathematics and 0.296 standard deviations in Serbian. For the lower primary school grades, dropouts are low in both *Early* and *Late Enrollee* and are not a major problem. Higher impacts are obtained in schools with a lower number of Roma: the higher is their number, the less the assistant can help them, and the lower is the impact of the programme on the outcomes of interest in schools with fewer Roma. Girls seem to benefit more from the programme in terms of better marks, whereas boys exhibit lower absences. With our data it is not possible to explain this differential impact of the programme. We do not know if assistants worked more with girls (boys) or if girls (boys) responded better to the programme in the case of marks (absences).

While first graders in treated schools perform better than their older colleagues, overall the programme does not seem to have a significant im-

pact on pupils' achievement. This is likely the case because assistants work mainly with lower grades and young cohorts are those really exposed to them. Therefore, the general modest effects should not be interpreted as a failure of the programme. Moreover, this study looks only at its impact in the first year. It is possible that assistants and schools need some time to adjust to the new role of the assistant and that the full benefit from them will come at a later stage. This idea is also supported by the literature on the importance of the experience of teachers which emphasises that gains in teaching skills are largest in the initial years of teaching (Rivkin, Hanushek, and Kain, 2005). Still, our results suggest that the programme is more effective in schools with less Roma. We are aware that it is possible that there are systematic differences between schools with a lower and a higher share of Roma. One could argue, for instance, that Roma in schools with a lower share are more willing to adapt and assimilate to the majority population. We cannot be certain that the same effects could be attained in schools with higher percentage of Roma if more assistants were assigned to these schools. It still seems plausible that if the goal of the policy maker is to close the gap in schooling outcomes, then more Roma teaching assistants should be assigned to schools with more Roma students.

We are not able to delve into the mechanism which drives our results and our estimates are derived from reduced form regressions. Still, our results could and probably are driven by two main mechanisms. First, the RTA interventions alters the inputs of the education productions function. The input of teaching is augmented and this can explain better marks of first grade students. A second mechanism possibly at play is the cultural transmission of preferences for education. The presence of a person coming from the same Roma ethnic background could alter the preferences for education and beliefs of Roma pupils, which induces them to attend more school and to exert more effort at school.

References

Angrist JD, Krueger AB (1991) Does Compulsory School Attendance Affect Schooling and Earnings? The Quarterly Journal of Economics 106(4):979–1014

Angrist JD, Lang K (2004) Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. The American Economic Review 94(5):1613–1634

Banerjee AV, Cole S, Duflo E, Linden L (2007) Remedying Education: Evidence from Two Randomized Experiments in India. The Quarterly Journal of Economics 122(3):1235–1264

Battaglia M, Lebedinski L (2013) The Curse of Low Aspirations: Remedial Education and Perceived Returns to Education of Roma People. Mimeo

Baucal A (2009) Development of Mathematical and Language Literacy among Roma Students. Psihologija 39(2):207–227

Boyd D, Grossman P, Lankford H, Loeb S, Wyckoff J (2009) Teacher Preparation and Student Achievement:. Education Evaluation and Policy Analysis 31(4):416–440

Buchmann C, DiPrete T, McDaniel A (2008) Gender Inequalities in Education. Annual Review of Sociology 34:319–337

Currie J, Thomas D (1999) Early Test Scores, Socioeconomic Status and Future Outcomes. NBER Working Paper 6943, National Bureau of Economic Research, Inc.

Domina T (2005) Leveling the Home Advantage: Assessing the Effectiveness of Parental Involvement in Elementary School. Sociology of Education 78(3):233–249

Duflo E (2004) Schooling and Labor Market Consequences of School Construction in Indonesia Evidence from an Unusual Policy Experiment. Journal of Development Economics 74(1):199–250

Fryer RG (2010) Racial Inequality in the 21st Century: The Declining Significance of Discrimination. Handbook of Labor Economics 4:855–971

Fryer RG (2011a) Financial Incentives and Student Achievement: Evidence from Randomized Trials. The Quarterly Journal of Economics 126(4):1755–1798

Fryer RG (2011b) Injecting Successful Charter School Strategies into Traditional Public Schools: Early Results from an Experiment in Houston.

NBER Working Papers 17494, National Bureau of Economic Research, Inc

Glewwe P, Kremer M, Moulin S (2009) Many Children Left Behind? Textbooks and Test Scores in Kenya. American Economic Journal: Applied Economics 1(1):112–135

Hanushek EA, Kain JF, Rivkin SG (2002) Inferring Program Effects for Specialized Populations: Does Special Education Raise Achievement for Students with Disabilities. Review of Economics and Statistics 84(4):584–599

Hoxby CM (2000) The Effects of Class Size on Student Achievement: New Evidence from Population Variation. The Quarterly Journal of Economics 115(4):1239–1285

Jacob BA (2004) Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago. American Economic Review 94(1):233–258

Jacob BA, Lefgren L (2004) Remedial Education and Student Achievement: A Regression Discontinuity Analysis. Review of Economics and Statistics 86(1):226–244

Kertesi G, Kezdi G (2011) The Roma/Non-Roma Test Score Gap in Hungary. American Economic Review: Papers & Proceedings 101:3:519–525

Lauer PA, Akiba M, Wilkerson SB, Apthorp HS, Snow D, Martin-Glenn ML (2006) Out-of-School-Time Programs: A Meta-Analysis of Effects for At-Risk Students. Review of Educational Research 76(2):275–313

Lavy V, Schlosser A (2005) Targeted Remedial Education for Underperforming Teenagers: Costs and Benefits. Journal of Labour Economics 23(4):839–874

Miguel E, Kremer M (2004) Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. Econometrica 72(1):159–217

Murnane RJ, Willett JB, Levy F (1995) The Growing Importance of Cognitive Skills in Wage Determination. The Review of Economics and Statistics 77(2):251–66

Neal D (2006) Why Has Black-White Skill Convergence Stopped?, Handbook of the Economics of Education, vol 1, Elsevier, chap 9, pp 511–576

Neal DA, Johnson WR (1996) The Role of Premarket Factors in Black-White Wage Differences. Journal of Political Economy 104(5):869–95

Open Society Institute (2007) Equal Access to Quality Education For Roma. Technical report, Open Society Foundation

Open Society Institute (2008) International Comparative Data Set. Technical report, Open Society Foundation

Podgursky MJ, Springer MG (2007) Teacher Performance Pay: A Review. Journal of Policy Analysis and Management 26(4):909–949

Rivkin SG, Hanushek EA, Kain JF (2005) Teachers, Schools, and Academic Achievement. Econometrica 73(2):417–458

Sanbonmatsu L, Kling JR, Duncan GJ, Brooks-Gunn J (2006) Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment. The Journal of Human Resources 41(4):649–691

Schultz PT (2004) School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program. Journal of Development Economics 74(1):199–250

Schweinhart LJ, Barnes HV, Weikart DP (1993) Significant Benefits: The High/Scope Perry Preschool Study Through Age 27. High Scope Press, Ypsilanti, MI

Vermeersch C, Kremer M (2005) Schools Meals, Educational Achievement and School Competition: Evidence from a Randomized Evaluation. Policy Research Working Paper Series 3523, The World Bank

5 Tables

Table 1: Averages of control variables and outcomes in pre- and treatment year: Roma

	Pre-treatment year			Treatment year		
	Early Enrollees (1)	Late Enrollees (2)	Difference (3)[(1)-(2)]	Early Enrollees (4)	Late Enrollees (5)	Difference (6)[(4)-(5)]
Characteristics						
Female	0.502	0.471	0.031	0.486	0.469	0.017
			(0.023)			(0.026)
Age	8.748	8.675	0.073	8.677	8.742	-0.065
			(0.089)			(0.110)
Born in the same town	0.867	0.814	0.053	0.877	0.807	0.070*
			(0.038)			(0.035)
Roma per school	0.223	0.193	0.030	0.235	0.194	0.041
			(0.056)			(0.057)
School size	304.937	361.506	-56.569	301.217	362.581	-61.364
			(52.963)			(56.046)
% of Roma per class	0.221	0.183	0.038	0.234	0.185	0.049
			(0.056)			(0.057)
Class size	22.161	23.966	-1.804	22.438	24.213	-1.775
			(1.424)			(1.381)
Outcomes						
Dropout[a]	0.015	0.019	-0.004	0.026	0.035	-0.009
			(0.006)			(0.009)
Absences (hours)	118.103	125.378	-7.275	134.037	155.528	-21.491
			(13.722)			(16.808)
Serbian[b]	2.430	2.547	-0.117	2.496	2.568	-0.072
			(0.123)			(0.144)
Mathematics[b]	2.284	2.370	-0.086	2.365	2.408	-0.043
			(0.125)			(0.156)
Number of schools	23	15		23	15	
Number of Roma pupils	1241	811		1268	847	

Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Dropout is equal to 1 if child dropped out of school during the year; otherwise 0.

[b] Marks range from 1 (worst) to 5 (best). They are categorical.

Table 2: Averages of control variables and outcomes in pre- and treatment year: Non Roma

	Pre-treatment year			Treatment year		
	Early Enrollees (1)	Late Enrollees (2)	Difference (3)[(1)-(2)]	Early Enrollees (4)	Late Enrollees (5)	Difference (6)[(4)-(5)]
Characteristics						
Female	0.487	0.477	0.010	0.469	0.486	-0.017
			(0.014)			(0.012)
Age	8.421	8.400	0.021	8.426	8.416	0.009
			(0.066)			(0.033)
Born in the same town	0.922	0.912	0.010	0.930	0.923	0.007
			(0.011)			(0.011)
Roma per school	0.223	0.193	0.030	0.235	0.194	0.041
			(0.056)			(0.057)
School size	304.937	361.506	-56.569	301.217	362.581	-61.364
			(52.963)			(56.046)
% of Roma per class	0.221	0.183	0.038	0.234	0.185	0.049
			(0.056)			(0.057)
Class size	22.161	23.966	-1.804	22.438	24.213	-1.775
			(1.424)			(1.381)
Outcomes						
Dropout[a]	0.001	0.0006	0.0004	0.001	0.000	0.001**
			(0.0006)			(0.0005)
Absences (hours)	39.159	36.231	2.938	42.549	40.276	2.273
			(2.535)			(2.743)
Serbian[b]	4.396	4.328	0.068	4.434	4.339	0.094
			(0.070)			(0.072)
Mathematics[b]	4.255	4.179	0.076	4.296	4.208	0.088
			(0.080)			(0.081)
Number of schools	23	15		23	15	
Number of Non Roma pupils	4303	3374		4122	3514	

Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Dropout is equal to 1 if child dropped out of school during the year; otherwise 0.

[b] Marks range from 1 (worst) to 5 (best). They are categorical.

Table 3: Average treatment approach

	Effect of programme in treatment year			
	all	all	female	male
	(1)	(2)	(3)	(4)
Dropout[a]				
post	0.017**	0.015**	0.001	0.027**
	(0.007)	(0.006)	(0.010)	(0.012)
treatment*post	-0.006	0.003	0.028*	-0.018
	(0.009)	(0.009)	(0.014)	(0.014)
No. observations	4167	4039	1951	2088
Absences				
post	31.236***	32.853***	22.456***	42.034***
	(7.856)	(9.078)	(10.797)	(10.764)
treatment*post	-17.299**	-16.679*	-4.713	-26.119**
	(7.856)	(9.078)	(10.797)	(10.764)
No. observations	3980	3868	1871	1997
Serbian[b]				
post	0.039	0.046	0.079	0.027
	(0.060)	(0.048)	(0.055)	(0.050)
treatment*post	0.044	0.012	-0.035	0.058
	(0.069)	(0.066)	(0.075)	(0.080)
Mathematics[b]				
post	0.051	0.065	0.096	0.041
	(0.069)	(0.062)	(0.080)	(0.056)
treatment*post	0.046	0.030	0.015	0.053
	(0.081)	(0.077)	(0.091)	(0.085)
No. observations	4085	3961	1916	2045
Controls[c]	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes

This table reports the effect of the programme on dropouts, absences and Serbian and mathematics. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Dropout is equal to 1 if child dropped out of school during the year; otherwise 0.

[b] Marks range from 1 (worst) to 5 (best). They are categorical.

[c] Controls included are school size, school size squared, number of Roma in school, number of Roma in school squared, percentage of Roma per class, class size, class size squared, female (=1), age, age squared, and migrant (=1).

Table 4: Average treatment approach - Placebo

	Placebo tests for pretreatment years					
	2006/2007 and 2007/2008			2007/2008 and 2008/2009		
	all	female	male	all	female	male
	(1)	(2)	(3)	(4)	(5)	(6)
Dropout[a]						
treatment*post	-0.003	0.006	-0.010	0.015	0.014	0.014
	(0.009)	(0.017)	(0.016)	(0.010)	(0.025)	(0.014)
No. observations	3640	1776	1864	3897	1897	2000
Absences						
treatment*post	0.955	-2.002	3.566	9.558	2.921	14.935
	(12.592)	(17.435)	(14.753)	(13.864)	(19.448)	(10.602)
No. observations	3542	1732	1810	3788	1850	1938
Serbian[b]						
treatment*post	0.059	0.141	-0.025	-0.094	-0.103	-0.053
	(0.080)	(0.105)	(0.103)	(0.077)	(0.093)	(0.072)
Mathematics[b]						
treatment*post	0.080	0.093	0.067	-0.057	-0.102	0.007
	(0.066)	(0.075)	(0.080)	(0.077)	(0.098)	(0.080)
No. observations	3585	1750	1835	3846	1876	1970
Controls[c]	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

This table reports the results of the placebo regressions for pretreatment years for the average treatment approach. The outcomes of the regressions are dropouts, absences, Serbian and mathematics. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Dropout is equal to 1 if child dropped out of school during the year; otherwise 0.

[b] Marks range from 1 (worst) to 5 (best). They are categorical.

[c] Controls included are school size, school size squared, number of Roma in school, number of Roma in school squared, percentage of Roma per class, class size, class size squared, female (=1), age, age squared, and migrant (=1).

Table 5: Intensity of treatment

	Effect of programme in treatment year			
	all	all	female	male
	(1)	(2)	(3)	(4)
Dropout[a]				
treatment*post	-0.007	-0.001	0.009	-0.008
	(0.008)	(0.008)	(0.014)	(0.012)
intensive*treatment*post	0.008	0.001	0.044	-0.038
	(0.025)	(0.023)	(0.044)	(0.035)
No. observations	4167	4039	1951	2088
Absences				
treatment*post	-8.707	-4.089	1.886	-9.058
	(7.720)	(8.394)	(8.324)	(11.931)
intensive*treatment*post	-19.312	-36.390	-9.522	-59.759**
	(27.119)	(24.179)	(36.096)	(24.684)
No. observations	3980	3868	1871	1997
Serbian[b]				
treatment*post	-0.027	-0.058	-0.154*	0.041
	(0.078)	(0.080)	(0.084)	(0.094)
intensive*treatment*post	0.330*	0.307**	0.499**	0.108
	(0.169)	(0.150)	(0.186)	(0.151)
Mathematics[b]				
treatment*post	-0.053	-0.078	-0.169	0.012
	(0.081)	(0.080)	(0.093)	(0.088)
intensive*treatment*post	0.454**	0.417**	0.680***	0.161
	(0.152)	(0.143)	(0.179)	(0.138)
No. observations	4085	3961	1916	2045
Controls[c]	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes

This table reports the effect of the programme on dropouts, absences and Serbian and mathematics. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Dropout is equal to 1 if child dropped out of school during the year; otherwise 0.

[b] Marks range from 1 (worst) to 5 (best). They are categorical.

[c] Controls included are school size, school size squared, percentage of Roma per class, class size, class size squared, female (=1), age, age squared, migrant (=1), and *intensive* (=1). *Intensive* is equal to 1 for schools with fewer than 43 Roma.

Table 6: Intensity of treatment - Placebo

	Placebo tests for pretreatment years					
	2006/2007 and 2007/2008			2007/2008 and 2008/2009		
	all	female	male	all	female	male
	(1)	(2)	(3)	(4)	(5)	(6)
Dropout[a]						
intensive*treatment*post	0.010	0.056	-0.037	-0.014	-0.072	0.042
	(0.016)	(0.050)	(0.039)	(0.019)	(0.040)	(0.035)
No. observations	3640	1776	1864	3897	1897	2000
Absences						
intensive*treatment*post	39.511	89.215	-11.344	-7.490	-40.031	25.702
	(42.393)	(60.105)	(34.917)	(26.991)	(34.883)	(25.084)
No. observations	3542	1732	1810	3788	1850	1938
Serbian[b]						
intensive*treatment*post	0.002	-0.145	0.164	0.113	-0.039	0.240
	(0.148)	(0.190)	(0.184)	(0.163)	(0.187)	(0.207)
Mathematics[b]						
intensive*treatment*post	0.026	-0.150	0.199	-0.011	-0.135	0.100
	(0.169)	(0.208)	(0.205)	(0.133)	(0.189)	(0.159)
No. observations	3585	1750	1835	3846	1876	1970
Controls[c]	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

This table reports the results of the placebo regressions for pretreatment years for the intensity of treatment approach. The outcomes of the regressions are dropouts, absences, Serbian and mathematics. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Dropout is equal to 1 if child dropped out of school during the year; otherwise 0.

[b] Marks range from 1 (worst) to 5 (best). They are categorical.

[c] Controls included are school size, school size squared, number of Roma in school, number of Roma in school squared, percentage of Roma per class, class size, class size squared, female (=1), age, age squared, migrant (=1), and *intensive* (=1). *Intensive* is equal to 1 for schools with fewer than 43 Roma.

Table 7: Cohort regression

	Effect of programme in treatment year								
	all			female			male		
	Early Enrollees	Late Enrollees	All	Early Enrollees	Late Enrollees	All	Early Enrollees	Late Enrollees	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dropout[a]									
young*post	0.014	0.079***	0.080***	0.031	0.071***	0.077***	-0.003	0.087***	0.083***
	(0.013)	(0.019)	(0.018)	(0.020)	(0.019)	(0.018)	(0.012)	(0.028)	(0.028)
young*post*treatment			-0.066***			-0.047*			-0.087***
			(0.022)			(0.027)			(0.029)
No. observations	2438	1601	4039	1200	751	1951	1238	850	2088
Absences									
young*post	23.579*	57.493**	54.639**	40.707**	67.618**	67.603**	9.740	52.423**	49.480**
	(11.537)	(21.263)	(22.235)	(16.000)	(28.938)	(26.469)	(14.055)	(22.036)	(24.287)
young*post*treatment			-31.867			-28.524			-40.337
			(24.945)			(30.772)			(27.852)
No. observations	2336	1532	3868	1152	719	1871	1184	813	1997
Serbian[b]									
young*post	0.079	-0.255**	-0.300***	0.052	-0.284	-0.352**	0.101	-0.295*	-0.328**
	(0.102)	(0.104)	(0.101)	(0.113)	(0.173)	(0.170)	(0.161)	(0.154)	(0.139)
young*post*treatment			0.382**			0.416*			0.423*
			(0.149)			(0.215)			(0.212)
Mathematics[b]									
young*post	0.082	-0.241	-0.291*	-0.019	-0.344	-0.426*	0.180	-0.205	-0.239*
	(0.116)	(0.143)	(0.152)	(0.132)	(0.234)	(0.241)	(0.155)	(0.142)	(0.135)
young*post*treatment			0.381*			0.428			0.412*
			(0.194)			(0.279)			(0.207)
No. observations	2395	1567	3962	1180	736	1916	1215	831	2046
Controls[c]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table reports the effect of the programme on dropouts, absences, Serbian and mathematics, using the cohort comparison methodology. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Dropout is equal to 1 if child dropped out of school during the year; otherwise 0.

[b] Marks range from 1 (worst) to 5 (best). They are categorical.

[c] Controls included are school size, school size squared, number of Roma in school, number of Roma in school squared, percentage of Roma per class, class size, class size squared, female (=1), age, age squared, migrant (=1) and *young* (=1). *Young* is equal to 1 when the child is at the first grade.

Table 8: Cohort regression - Placebo

	Placebo tests for pretreatment years					
	2006/2007 and 2007/2008			2007/2008 and 2008/2009		
	Early Enrollees	Late Enrollees	All	Early Enrollees	Late Enrollees	All
	(1)	(2)	(3)	(4)	(5)	(6)
Dropout[a]						
young*post	0.003	0.008	0.011	-0.020	-0.036**	-0.037***
	(0.017)	(0.017)	(0.015)	(0.015)	(0.012)	(0.013)
young*post*treatment			-0.010			0.019
			(0.024)			(0.020)
No. observations	2259	1381	3640	2389	1508	3897
Absences						
young*post	-14.021	-56.385*	-54.237*	-9.425	-13.207	-13.596
	(17.068)	(27.822)	(28.643)	(13.600)	(16.765)	(18.055)
young*post*treatment			34.591			10.100
			(32.454)			(23.240)
No. observations	2203	1339	3542	2331	1457	3788
Serbian[b]						
young*post	0.030	0.143	0.105	-0.167	0.034	0.043
	(0.207)	(0.194)	(0.203)	(0.178)	(0.154)	(0.156)
young*post*treatment			-0.024			-0.244
			(0.278)			(0.231)
Mathematics[b]						
young*post	0.164	0.117	0.086	-0.101	-0.115	-0.113
	(0.237)	(0.180)	(0.185)	(0.145)	(0.197)	(0.196)
young*post*treatment			0.122			-0.022
			(0.286)			(0.239)
No. observations	2232	1354	3586	2364	1482	3846
Controls[c]	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

This table reports the results of the placebo regressions for pretreatment years for cohort comparison methodology. The outcomes of the regressions are dropouts, absences, Serbian and mathematics. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Dropout is equal to 1 if child dropped out of school during the year; otherwise 0.

[b] Marks range from 1 (worst) to 5 (best). They are categorical.

[c] Controls included are school size, school size squared, number of Roma in school, number of Roma in school squared, percentage of Roma per class, class size, class size squared, female (=1), age, age squared, migrant (=1) and *young* (=1). *Young* is equal to 1 when the child is at the first grade.

6 Figures

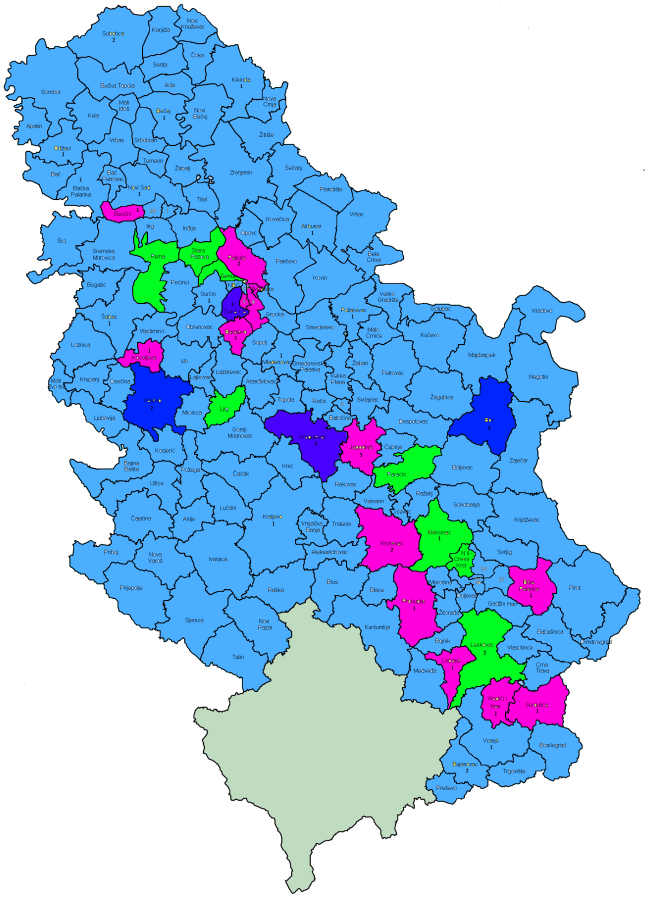

Figure 1: Location of the schools with assistants
This figure reports the distribution of schools in our sample. In pink municipalities there are only *Early Enrollee* school; in green municipalities there are only *Late Enrollee school* and in dark blue municipalities there are both *Early* and *Late Enrollee* schools.

A Appendix

A.1 Institutional Setting

A.1.1 Living conditions of Roma in Serbia

Data on Roma in Serbia are inaccurate and scarce. Surveys often lack information about ethnic identity of the respondents. More importantly, when asked about their ethnicity, some Roma people do not declare themselves as Roma. Most of them consider themselves both Roma and Serbian and the question of nationality allows only one answer.

The official 2002 census counts 108,000 Roma, while estimates put forward a number between 350,000 and 500,000 or approximately 5-7% of the overall population (Open Society Institute, 2007). Most Roma live in segregated settlements and have different demographic characteristics from the rest of the population. According to the World Bank Living Standard Measurement Survey (LSMS) 2003 - which provides a boosted sample of Roma in Serbia - their households are more numerous than the average household, they have more children and their population is younger. The percentage of male Roma who declare to have worked over the last week is similar to the national average (69%). Contrary to men, the participation of females is 34% and considerably lower than the national average (53%). Overall, approximately 60% of Roma have a consumption below the poverty line and weekly consumption of food per household member in Roma households is half the national average.

Turning to education, 60% of Roma younger than 18 years old have not completed primary education. In contrast, only 20% of overall population do not have a primary school diploma. Out of all children of primary school-age, 30% of Roma do not attend school whereas this is the case for only 1% of the overall population of primary school-age. Using data from the National Assessment Study conducted with third grade students, Baucal (2009) finds that after the first 3 years of school Roma pupils lag 2.2 - 2.5 years behind the average student. Also, children from Roma ethnic minority performed worse on standardised tests than Non Roma children with the same socio-economic background.

The main barriers of access to education for Roma are absence of documents, financial constraints, parents' low educational background, child labour, discrimination from teachers and pupils and language barriers (Open Society Institute, 2008). In the recent years Serbian schools started enrolling children with incomplete documents, but there is still a minor number of children not able to enrol due to lack of documents. According to the law, the local government should inform schools and parents that children who reach the school-age in the municipality have to enrol at school. But Roma are often not regularly registered as residents in the municipality and the local government is not able to reach out to them. School books and addi-

tional school material are a significant burden for the budget of poor families and the most poor among Roma children do not even own adequate clothing for winter months and live in overcrowded homes where they do not have adequate conditions to pursue their studies. A majority of Roma parents has low educational attainment and this implies that they often cannot help their children with their school work. In addition, some parents attach little value to schooling and education. These reasons together imply that the perceived benefits of going to school are lower than respective costs. Moreover, in some cases Roma children help their parents in their work. Also, Roma pupils can face discrimination from teachers and other pupils. There is anecdotal evidence that they are often seated in the last row in classrooms, that teachers do not read their homework and that teachers do not encourage them in their studies. Another problematic issue is that a considerable share of them is sent to special schools. Finally, in a survey conducted by UNICEF - Multiple Indicator Cluster Survey, 2006 - only 10% of Roma declare Serbian to be their mother tongue. Children may face difficulties at school due to limited knowledge of Serbian.[34]

A.1.2 Primary Education System in Serbia

In Serbia, school is compulsory until age 15. Children enrol at primary school if they are aged at least 6.5 years at the start of the scholastic year in September. Since 2007 the attendance of at least 6 months of a cost-free preschool programme is compulsory; in 2010 the length of the compulsory preschool has been extended to 9 months.[35]

Primary school consists of 8 years. In the first four grades pupils get one teacher who teaches all compulsory subjects except English, while in the upper four years pupils have one teacher per subject. In the first grade teachers use descriptive marks; from the second grade on, the range of marks is 1 to 5 with 1 being the insufficient and worst mark. The marks are categorical. During each semester, a child gets 4 marks for each subject. Out of the 4 marks, at least 2 marks are derived from written tests. The teacher can decide what to use as assessment for the remaining two marks. He could give additional written tests, give marks based on oral examination, homework or class participation. The final mark at the end of the year is the arithmetic mean of 8 marks. If a pupil gets at least one insufficient mark at the end of the year, her teacher can decide whether to let her pass to the upper grade or to ask her to take the retake exam in August. In the last few

[34] With few exceptions, the rest declare Romani to be their mother tongue.

[35] The obligatory preschool programme has been introduced in order to facilitate the transition to school for children from lower socio-economic backgrounds. In the initial years the capacities of preschool institutions were not sufficient to enrol all preschool children. Hence, some children, mainly from poorer families or in rural areas, could not be enrolled in preschool. However, due to the lack in the enforcement of the law, they were let to enrol in school also without having attended the compulsory preschool programme.

years the Ministry of Education has suggested that schools reduce repetition rates, especially in the lower four grades.

There are no school fees for primary school, but indirect costs such as books and other school material can pose a considerable cost for some parents.[36] The Ministry of Education aims at reducing the cost of education and the first graders in 2009/2010 are the first generation to receive free text books. The plan is that this generation and all younger generations obtain free school books in the future.

A.2 Sensitivity analysis of the threshold of intensity of treatment

In the main body of this paper, we examine how the effect of the programme varies based on the number of Roma in a school. We split the schools in two equally sized groups based on the number of Roma in the school. We define a dummy 'intensive' which takes value 1 if a school is among schools with *less Roma* (at most 43 Roma children), and value 0 if a school is among schools with *more Roma* (more than 43 Roma children). In this section we report how our results change when we vary this threshold.

Table 9 shows the results of the intensity of treatment approach for different thresholds. We rank the schools based on the number of Roma, starting with the lowest, and divide then the schools in 10 deciles. Column (1) of the table reports the results if we set the dummy 'intensive' equal to 1 for schools in the first decile and 0 otherwise. In column (2) we set the dummy 'intensive' equal to 1 for schools in the first and second decile, and 0 otherwise, etc. The reported coefficient (*intensive*treatment*post*) captures the effect of the programme on treated children in schools where the programme is 'intensive'. Essentially, going from the left to the right columns we increase the number of schools and thus the number children for which we consider the programme to be 'intensive'. For instance, in column (1) children in schools with less than 16 Roma are considered to be treated. In column (2) this number rises to 23 Roma per school, etc. We expect that the effect of the programme drops as we go from left to the right columns and this is exactly what this table shows. The results in table 9 are only indicative and should be interepreted with caution especially for the lower deciles. For lower deciles we consider only very few schools to be actually treated (in the case of column (1) only one treated school falls into the first decile).

[36] On average, in Serbia costs associated with school (books and other school material) correspond to almost 2% of yearly household income (LSMS 2003). Based on a survey we conducted in Belgrade, for Roma people these costs account for 6% of their yearly household income.

Table 9: Intensity of treatment: Sensitivity analysis

	1st decile (1)	up to 2nd decile (2)	up to 3rd decile (3)	up to 4th decile (4)	up to 5th decile (5)	up to 6th decile (6)	up to 7th decile (7)	up to 8th decile (8)	up to 9th decile (9)
Dropout[a]									
intensive*treatment*post	0.032	0.031	-0.011	-0.000	0.001	0.010	0.007	-0.007	0.005
	(0.032)	(0.037)	(0.033)	(0.029)	(0.023)	(0.018)	(0.017)	(0.015)	(0.012)
Absences									
intensive*treatment*post	-47.636	-47.970	-75.063**	-49.650*	-36.390	-20.643	-11.530	-12.136	7.308
	(34.443)	(34.788)	(31.415)	(29.079)	(24.179)	(17.668)	(16.936)	(15.706)	(13.351)
Mathematics[b]									
intensive*treatment*post	0.806***	0.650**	0.617***	0.422**	0.417***	0.391***	0.347***	0.280**	0.353***
	(0.126)	(0.258)	(0.214)	(0.178)	(0.143)	(0.121)	(0.117)	(0.105)	(0.080)
Serbian[b]									
intensive*treatment*post	0.204	0.500	0.457**	0.364**	0.307**	0.217*	0.265**	0.175	0.133
	(0.136)	(0.354)	(0.190)	(0.163)	(0.150)	(0.122)	(0.105)	(0.105)	(0.091)
No. observations	3961	3961	3961	3961	3961	3961	3961	3961	3961
Controls[c]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table reports the effect of the programme on dropout, absences, Serbian and mathematics for different thresholds. We order the schools based on the number of Roma in a school. In column (1) we define the programme as 'intensive' in schools of the 1st decile (max. 16 Roma per school), column (2) up to 2nd decile (max. 23 Roma per school), column (3) up to 3rd decile (max. 29 Roma per school), column (4) up to 4th decile (max. 34 Roma per school), column (5) up to 5th decile (max. 43 Roma per school), column (6) up to 6th decile (max. 48 Roma per school), column (7) up to 7th decile (max. 52 Roma per school), column (8) up to 8th decile (max. 66 Roma per school), column (9) up to 9th decile (max. 116 Roma per school). Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Dropout is equal to 1 if child dropped out of school during that year; otherwise 0.

[b] Marks range from 1 (worst) to 5 (best). They are categorical.

[c] Controls included are school size, school size squared, percentage of Roma per class, class size, class size squared, female (=1), age, age squared, migrant (=1), and *intensive* (=1). *Intensive* is equal to 1 in column (1) if school falls into first decile, in column (2) if school falls into first or second decile, etc.

Table 10: Balancing tables for intensity of treament approach for pre- and treatment year

	Pre-treatment year			Treatment year		
	Early Enrollees (1)	Late Enrollees (2)	Difference (1-2) (3)	Early Enrollees (1)	Late Enrollees (2)	Difference (1-2) (3)
Dropout:[a]						
Less Roma	0.015	0.026	-0.011	0.035	0.044	0.009
			(0.016)			(0.026)
More Roma	0.021	0.019	-0.002	0.023	0.033	-0.010
			(0.007)			(0.010)
Absences (hours):						
Less Roma	143.691	146.399	-2.708	141.220	171.947	-30.727
			(33.530)			(39.283)
More Roma	109.800	120.637	-10.837	131.583	151.127	-19.544
			(13.648)			(18.262)
Serbian:[b]						
Less Roma	2.529	2.613	-0.084	2.680	2.461	0.219
			(0.234)			(0.212)
More Roma	2.397	2.531	-0.135	2.436	2.598	-0.162
			(0.147)			(0.167)
Mathematics:[b]						
Less Roma	2.380	2.513	-0.133	2.549	2.281	0.268
			(0.219)			(0.213)
More Roma	2.252	2.337	-0.085	2.305	2.443	-0.138
			(0.151)			(0.182)

This table reports the results of balancing tests for pretreatment and treatment year for the intensity of treatment approach. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Dropout is equal to 1 if child dropped out of school during that year; otherwise 0.

[b] Marks range from 1 (worst) to 5 (best). They are categorical.

A.3 Cost of the programme

This section aims to give an estimate of the cost of the programme and to contrast this programme to alternative interventions available to the government. At this early stage of the programme it is not possible to provide an estimate of the monetary benefits. Still, we discuss two possible channels through which this programme could affect future labour market outcomes and thus, could have a monetary gain. First, the programme could raise completion rates of primary school and consequently, the average educational attainment of treated Roma children. Higher educational attainment implies higher earnings (Angrist and Krueger, 1991). We do not find that dropout rates fall, because in the lower 4 grades of primary school the dropout rates are already low. Primary school lasts 8 years and we believe that dropout rates in the higher grade (from 5th to 8th) are the ones actually affected by the programme. A second channel through which the programme could affect children's labour market outcomes is through better grades. Several studies (Murnane, Willett, and Levy, 1995; Neal and Johnson, 1996; Currie and Thomas, 1999) have demonstrated a positive relationship between pupils' test scores at school and their subsequent earnings.

The main cost of the programme are the (gross) wages of the assistants which amount to 450 Euros per month. In addition to that, in 2009/2010, assistants attended 7 trainings totalling up to 22 days. Our estimation suggests that the training per assistant had a cost of approximately 1,200 Euros.[37] The turnover of the assistants in the first year was very low with only two people quitting in that period. This is especially important because it indicates that training costs will not be bore each year. Not taking into account the training costs, we calculate that the programme costed 100 Euros per Roma student per year.[38] We know that in Belgrade 53% of Roma children were actually treated (Battaglia and Lebedinski, 2013), assuming that the take up rate was the same also in the whole of Serbia, this yields a cost of 200 Euros per treated child.

There are no cost estimates of other programmes in Serbia to which this programme could be directly compared. We still discuss two other measures which we consider alternatives to the RTA programme. One possible

[37] We calculate the cost for accommodation and stay during the training to be 40 Euros per day plus transport costs of 10 Euros for 7 travels. This amounts to 950 Euros. There were 22 days of training and we add additional 14 days for preparation of training and reporting. The cost of trainers is 150 Euros per day and the wage cost is 5,400 Euros. Accommodation, stay and travel costs of trainers are 1,510 Euros (22 days*40 Euros + 7 travels *10 Euros). Therefore, the total costs for trainers including their stay and accommodation is 6,910. Dividing this sum with 26 (number of assistants) gives 265 Euros per assistant. The cost of accommodation per assistant (950 Euros) plus training per assistant (265 Euros) adds up to a total cost of the training of 1,215 Euros.

[38] 450 Euros *12 month [wages]* 23 schools /1268 Roma students = 97,95 Euros.

alternative would be to use professional teachers to give remedial education classes. The gross wage of teachers is in the range of 900 to 1,200 Euros depending on the experience and thus twice the amount paid to the assistants. Assuming that two assistants have at least the same performance as one teacher, the RTA programme is more cost effective. Aside from being less costly, an additional advantage of RTAs is that they act as role models for Roma children and this cannot be easily quantified in monetary terms. A second option available to the policy maker would be class size reduction, but this intervention did not yield satisfactory results in some settings (Hoxby, 2000) and is a very costly intervention. Another disadvantage of the class size reduction is that it does not specifically target disadvantaged children.

All of these facts, taken together, indicate that the programme is with 200 Euros per child per year not overly costly and that it also performs well with respect to its alternatives.

A.4 Spillover effects - Non Roma pupils

We can investigate whether this programme also affects Non Roma pupils. We employ both econometric strategies and their combination and find that neither absences reduced nor marks improved for Non Roma students. The presence of a Roma assistant do not improve Non Roma schooling outcomes. Results are reported in tables 11 and 12.

These results, combined together with the previous ones, provide some evidence that the programme is succeeding in reducing the gap between Roma and Non Roma children, both in school achievements and attendance.

Table 11: Average treatment approach - Non Roma

	Effect of programme in treatment year on Non Roma		
	Absences	Serbian[a]	Mathematics[a]
	(1)	(2)	(3)
post	5.025***	-0.011	0.011
	(1.185)	(0.031)	(0.018)
treatment*post	-1.586	0.054	0.029
	(1.725)	(0.034)	(0.025)
No. observations	14686	14982	14981
Controls[b]	Yes	Yes	Yes
School FE	Yes	Yes	Yes

This table reports the effect of the programme on Non Roma for absences, Serbian and mathematics. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Marks range from 1 (worst) to 5 (best). They are categorical.

[b] Controls included are school size, school size squared, number of Roma in school, number of Roma in school squared, percentage of Roma per class, class size, class size squared, female (=1), age, age squared, and migrant (=1).

Table 12: Cohort regression - Non Roma

	Effect of programme in treatment year on Non Roma								
	Absences			Serbian[a]			Mathematics[a]		
	Early Enrollees	Late Enrollees	All	Early Enrollees	Late Enrollees	All	Early Enrollees	Late Enrollees	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
young*post	0.658	3.178	3.259	0.033	-0.056	-0.066	0.044	-0.062	-0.070
	(2.682)	(3.274)	(3.057)	(0.072)	(0.077)	(0.069)	(0.058)	(0.085)	(0.079)
young*post*treatment			2.787			-0.037			-0.049
			(4.986)			(0.107)			(0.111)
No. observations	8099	6587	14686	8232	6750	14982	8232	6749	14981
Controls[b]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table reports the effect of the programme on Non Roma for absences, Serbian and mathematics. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Marks range from 1 (worst) to 5 (best). They are categorical.

[b] Controls included are school size, school size squared, number of Roma in school, number of Roma in school squared, percentage of Roma per class, class size, class size squared, female (=1), age, age squared, migrant (=1) and *young* (=1). *Young* is equal to 1 when the child is at the first grade.

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2014/3082/21