Equal access to education: An Evaluation of the Roma Teaching Assistant Programme in Serbia*

Marianna Battaglia
Bocconi University

Lara Lebedinski
Bocconi University

August 29, 2012

Abstract

Roma constitute a large ethnic minority suffering severe social exclusion, especially in terms of high poverty levels and low educational attainments. This paper investigates the impact of the Roma Teaching Assistant Programme in Serbia in its first year of introduction on the following schooling outcomes: marks, absences and probability to drop out. By using first hand collected data, we employ two different identification 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 not exposed to it. We find that, on average, marks have improved and dropouts have reduced for those children exposed to the programme in their first grade. There is also evidence that overall children exposed to the programme went on average more to school. Higher and more systematic impacts are obtained in schools with a lower number of Roma. We confirm the robustness of our results with placebo tests for the years prior to the introduction of the programme.

Keywords: programme evaluation, primary education, Roma

JEL classification codes: I21, J13, D04

*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.We would like to thank Tito Boeri, 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 and Third International Workshop on Applied Economics of Education 2012. 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. The usual disclaimer applies. Contact email: marianna.battaglia@phd.unibocconi.it; lara.lebedinski@phd.unibocconi.it

1 Introduction

Roma are mainly located in South Eastern Europe and with a population of approximately 6 million of people they constitute the largest ethnic minority in the continent (Open Society Institute, 2008).[1] In all countries they suffer severe social exclusion which can be observed in high poverty and unemployment levels, low educational attainments and no participation in the political and cultural life. Roma 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] significantly lower family permanent incomes, also due to greater household size and lower incidence of home ownership, and lower wages, given the overrepresentation among low skilled jobs. They often lack access to credit and property ownership and are overdependent on social benefits.

The problems of the Roma minority have moved at the forefront of the media's attention especially with the expansion of the European Union. The visa liberalisation and the adhesion to the Union of countries like Romania and Bulgaria, in which the percentage of Roma population is high, have indirectly led to significant migration flows towards the Western countries. Appearance of informal settlements, increased number of unemployed and inadequacy of the education system in receiving new foreign pupils are some of the problems which arose in the receiving areas. The extraordinariness of the phenomenon has led to hot discussions within the European countries and civil society and increased the interest of the European Union in those countries which will likely enter the Union in the future and where a high percentage of Roma population resides, e.g. Serbia. Understanding the impact of programmes targeting Roma is useful not only for these countries but also for the receiving ones, where a minority group becomes also a migrant group.

The aim of this paper is to evaluate the impact of the Roma Teaching Assistant Programme in the first year of its complete introduction. More precisely, we want to examine whether this remedial education programme is effective in reducing drop outs, raising attendance and

[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] 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. Nowdays, however, they are mainly attended by pupils who are late at enroling and by pupils who decided to return to school after dropping out.

improving the marks of Roma pupils, by means of a target increase in instruction time, help in homework and assignments and direct link between assistant and parents.

Schooling has always been considered a needed measure to improve living conditions of Roma people and attempts in this direction have been devised by various countries for many years. Half of the Roma population is indeed younger than 18 years old and focusing on children and young people is broadly recognised as a crucial step towards Roma inclusion: higher enrolment rates and better achievement at school are expected to lead to persistent effects in the labor market and in the reduction of poverty in the long-run. Nonetheless, nowadays the net enrolment rate of Roma in primary education varies among the countries and it is still with 40% to 60% really low. Moreover, students may enrol at the beginning of the year, but may not actually attend school: the percentage of completion rates of primary school are in the range of only 30% to 40% for most countries (Open Society Institute, 2008). However, to the best of our knowledge, there are not systematic studies in economic literature that try to investigate how to improve life circumstances of Roma in general and Roma kids in particular.

This paper is a first solid attempt in this direction and it 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 not data available, and by contrasting their achievement to the average Non Roma pupils.[3] More broadly, it adds evidence on short-term effects of remedial education programmes on minority groups and suggests replicable examples in contexts where minorities suffer low attainment rates and social exclusion. For Roma people this is the case in many other European countries.

The Roma Teaching Assistant Programme[4] is the main programme targeting Roma inclusion in education in South Eastern Europe.[5] It began in Serbia in 2002 as a pilot programme carried out by different NGOs and from 2007 to 2009 was led by the Organisation for Security and Cooperation in Europe (OSCE). Since 2009 the Ministry of Education has been responsible for the coordination of the programme, which for the first time had a broad country coverage. Roma assistants - one per each school - participate in regular lessons where they provide additional

[3]All data are primary data, collected by the authors in the Summer-Autumn 2010.

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

[5]The Serbian 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.

help to Roma pupils who have difficulties in following classes, especially from lower grades. Moreover, they organise additional lessons, help them with their homework and assignments and once per week they visit their parents. In September 2010 the name of Roma assistants has been changed to pedagogical assistants and their target group is no longer only Roma but all children from marginalised groups.[6] Nonetheless, the Ministry of Education expects that mainly Roma children will benefit from this programme.

By using first hand collected data, we employ two different identification 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 not exposed to it. Our results suggest that, on average, marks have improved by almost 0.28 standard deviations in mathematics, 0.35 standard deviation in Serbian and dropouts have reduced by 6.6 percentage points for those children exposed to the programme in their first grade. There is also evidence that all children exposed to the programme went on average at least 17 hours more to school. Higher and more systematic 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 seems to be especially the case for female, for whom being in a school with a lower number of Roma turns out to be more favourable.

The paper is related to three strands of the literature: on remedial education programmes targeting underachieving students, on programmes aiming at improving schooling outcomes of minority communities and at narrowing differences between racial groups and on programmes aiming at achieving better schooling outcomes of the poor.

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 difficult to distinguish between the two effects, especially because the selection mechanism is not typically fully observable. Few studies are able to overcome this problem and find proper counterfactuals: this literature is still scanty and country specific. Among others, Lavy and Schlosser (2005) succeed in evaluating the effects of a remedial education programme for underperforming high school students in Israel. The intervention prepared students for the matriculation exams and aimed at increasing the

[6]The programme is now financed by the European Union and it is named *Education for All - Increasing the Availability and Quality of Education for Children from Marginalised Groups.*

school mean matriculation rates. It took place gradually allowing the authors to use as a control group those schools enroled in the programme later. They found that the programme raised the school mean matriculation rate by 3.3 percentage points. In the United States Hanushek et al. (2002) investigate the effects of targeted programmes on learning-disabled or emotionally disturbed students. They follow those who move in and out of these special programmes and identify programme effectiveness from changes over time in individual performance by comparing academic performance before and after placement into special education. Results suggest that the average special education programme significantly boosts mathematics achievement of special-education students by roughly 0.1 standard deviations. Jacob and Lefgren (2004) study the effect of summer school and grade retention programmes in the United States by using a regression discontinuity design. They find that their net effect is to improve academic performance in reading and mathematics among low-achieving students and that these positive effects remain substantial at least two years following the completion of the experience. Another way to overcome the potential selection bias and understand the impact of these programmes is offered by randomised evaluations. An interesting and successful randomized experiment, for instance, has been conduced in schools in urban India by Banerjee et al. (2007). As part of the programme, underachieving third grade students would have met for two hours each day with an instructor during school hours.[7] mark of children in schools with remedial education improved in both the first and second year of the programme (Banerjee et al., 2007).

The second strand of literature this paper is related to is the literature on programmes aiming at improving schooling outcomes of minority communities and at narrowing differences between racial groups. The black-white mark gap has been intensively investigated in the United states. In the past five decades there have been many attempts to close the racial gap even before kids enter school. The first and most known programme 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.[8] 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 expe-

[7]The remedial classes consisted of 15-20 students and they focused on the core competencies such as literacy and numeracy skills.

[8]Schweinhart et al. (1993) find that students in the programme had higher mark between the ages of 5 to 27, 21% less grade retention or special services required and 21% higher graduation rates.

riences 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 et al., 2006), of merit pay for principals, teachers, and students(Podgursky and Springer, 2007; Roland G. Fryer, 2010), of professional development for teachers (Boyd et al., 2008), and by getting parents to be more involved (Domina, 2005), by placing disadvantaged students in better schools through desegregation busing (Angrist and Lang, 2004) or alter the neighborhoods in which they live (Jacob, 2004; Sanbonmatsu et al., 2006). The evidence on the efficacy of these interventions is mixed: certain programmes have left the racial achievement gap essentially unchanged. However, according to Roland G. Fryer (2010), racial differences in social and economic outcomes are greatly reduced when one accounts for educational achievement and poverty levels. This points to the fact that there is little empirical evidence for discrimination in the recent data. The same has been concluded by Kertesi and Kezdi (2011) in their study on Roma 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 mark gap disappears in reading and decreases by 85% in mathematics.

The third relevant strand of literature is on programmes aiming at achieving better schooling outcomes of the poor. This literature suggests that conditional cash transfers, modelled after the Mexican programme PROGRESA, are successful in improving 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 mark of the average student(Schultz, 2004; Miguel and Kremer, 2004). Moreover,Das et al. (2011) show that, although unanticipated school grants lead to significant improvements in student mark, anticipated grants have no impact on them. Analogously, Roland G. Fryer (2010), 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, increasing in attendance and students pass. Providing incentives for achievement in mark are 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 et al., 2009; Vermeersch and Kremer, 2004).

The rest of the paper is organised as follow. Section 2 gives a general overview of the Serbian education system and summarises the main characteristics of Roma in Serbia. Section 3 gives a description of The Roma Teaching Assistant Programme. Sections 4 and 5 describe our data, the empirical strategy and present our results. Section 6 discusses our findings and concludes.

2 The Education System and Roma in Serbia

2.1 Primary Education System in Serbia

In Serbia, school is compulsory until the age of 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.[9]

Primary school consists of 8 years. In the first four grades pupils have one teacher who teaches all compulsory subjects except English, while in the upper four years of primary school pupils have one teacher per subject. In the first grade children get descriptive marks; from the second grade on, the range of marks is 1 to 5 with 1 being the insufficient and worst mark. If a pupil has at least one insufficient in the lower four grades 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 years the Ministry of Education has suggested to schools to 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.[10] The Ministry of Education aims at reducing the cost of education and the first graders in 2009/2010 are the first generation which received free text books. The plan is that this generation and all younger generations obtain free school books in the future.

[9]The obligatory preschool programme has been introduced in order to facilitate the transition to school for children from lower socioeconomic 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 enroled 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.

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

2.2 Context

Data on Roma in Serbia are inaccurate and scarce. Official census data from 2002 suggest that in Serbia there are 108,000 Roma, although estimates put forward a number of somewhere between 350,000 to 500,000 or approximately 6% of the overall population(Open Society Institute, 2007). Most Roma live in segregated settlements and have considerably 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 - the average household size of Roma population is of 4.5 household members and thus larger than the national average of 3.2. The average number of children younger than 18 years is 2.4 per Roma households, while the population average is only 0.9. 25% of Roma are younger than 10 and approximately 50% of the Roma population is younger than 23. Consequently, the average age of Roma is 25, whereas the average age in the country is 42. The percentage of male Roma who declare to have worked over the last week is similar to the national average (70%). Nonetheless, the participation of females is only around 30% and therefore considerably lower than the national average (50%). 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 children 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. Moreover, children from Roma ethnic minority performed worse on standardised tests than Non Roma children with the same socioeconomic 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.

In the recent years schools started enroling 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 is responsible of informing 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 some of them. School books and additional school material are a significant burden for the budget of poor families and the most poor among Roma children do not even have 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 extremely low compared to the respective costs. Moreover, in some cases Roma children help their parents in their work, e.g. they would go with their parents to collect rubbish or they would help them selling goods on the market, or have to take care of their younger brothers and sisters while the parents are working. 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 they a considerable share of them is sent to special schools for . Finally, in a survey conducted by UNICEF - Multiple Indicator Cluster Survey, 2006 - only 10% of Roma declare Serbian to be their mother tongue. As a consequence, children may face difficulties at school due to limited knowledge of Serbian.

3 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. Since 2009 the programme started to have a country coverage and it is now under the coordination of the Ministry of Education. In the scholastic year 2009/2010 there were 48 primary schools which 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 which have started with the programme in September 2009 (*Early Enrollees*) and schools which were assigned a Roma assistant starting with the scholastic year 2010/2011 (*Late Enrollees*). The 22 schools which joined the programme between 2002 and 2007 are excluded from

our analysis: the selection on these schools was not centralised since they were chosen by NGOs due to their considerable percentage of Roma pupils. For the purpose of our analysis the schools involved in the programme are therefore 26 *Early Enrollees* and 77 *Late Enrollees*.

Both schools and potential Roma assistants had to apply in order to participate in the programme. Among 78 schools which applied in 2009 a commission representing the government institutions together with OSCE representatives, chose 26 *Early Enrollees* schools based on the following two criteria: first, the percentage of Roma students between 5% and 40% and then, preferably, the availability of preschool programme in the school.[11] The requirements for Roma assistants were knowledge of Romani, secondary school diploma and experience in working with children. 158 applications were received for 26 assistant positions.[12]

In 2010 the programme has been renamed to *Education for all* and starting with the scholastic year 2010/2011 Roma Teaching Assistants have been renamed to pedagogical assistants. The same selection criteria for the percentage of Roma students as in 2009/2010 applied for further 77 schools out of 252 which entered the programme in 2010/2011 (*Late Enrollees*). The only difference was that in 2010 also schools not offering the compulsory preschool programme could apply for an assistant. The reason is that in 2010/2011 pedagogical assistants were also introduced in 50 kindergartens which offer the compulsory preschool programme. Schools which were not offering the preschool programme could have then been close to kindergartens offering it. The Roma pupil would have been followed by an assistant from her entry in the school anyhow.[13] Selection criteria for now pedagogical assistants remained unchanged and out of the 329 applications for the position, 77 were accepted to work at schools and another 50 were accepted for kindergartens.

We believe that parents were not aware of the existence of the programme before enroling their children at school. Data also confirm that *Early Enrollees* were not attracting more Roma students than *Late Enrollees* in the first year of the programme.[14] Therefore, we are confident

[11] 64 out of 78 schools which applied had a percentage of Roma beetwen 5% and 40%. Among these 64, OSCE selected 19 schools (out of 26) with a preschool programme, 5 schools (out of 37) with no preschool programme and a school for which no information is available.

[12] For the evaluation of the applications of assistants the following criteria were taken into account: highest level of education completed or enroled (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).

[13] 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.

[14] 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 -

that our analysis is not affected by possible selection of children into schools.

In 2009 the Ministry of Education has organised a series of seminars to provide the necessary knowledge and skills to Roma teaching assistants. In total, the assistants have attended 19 working days of seminars and courses in the scholastic years 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 a set of 9 modules, which all assistants have to attend in their first year of service, has been devised.

Every school receives only one assistant. Schools which are assigned a Roma assistant receive a description of her duties, but they are free to decide how to allocate the time of the assistant depending on the need of the school. However, there is a time allocation suggested by the Ministry. According to this this 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). Activities at school involve both working during regular classes as well as 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,[15] cooperation with Roma NGOs, etc. The assistants were advised to work mainly with lower grades, especially the first. Their objectives are indeed: making sure that children go to school, preventing them from dropping out and helping them to succeed at school.

4 Preliminary Analysis

4.1 Data and Trends of the Variables

We use first hand collected data. They come from administrative records of 23 schools among *Early Enrollees* and 15 schools among *Late Enrollees*.[16] We select the 15 *Late Enrollees* schools out of 77 which got an assistant in 2010 according to the following criteria: firstly, they have to be in the same district of a *Early Enrollees* school;[17] secondly, they have to be in a rural/urban

2009/2010 - these percentages were respectively 1.6% and 1.3%. Thus, the number of Roma pupils enroling at school for the first time reduced between the two years and it did it proportionally in both types of schools.

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

[16] 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 neither in the number of pupils nor in the percentage of Roma children and they are located in different areas: one in Belgrade, one in the central area of the country and one in the South.

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

municipality as the nearby *Early Enrollees* school; thirdly, they have to share a similar school size to the nearby *Early Enrollees* school and finally a similar percentage of Roma pupils.[18]

[insert TABLE 1 here]

Schools are mainly in Belgrade/Central Serbia and in the South/South-Eastern part of the country, and they are fairly located in both rural and urban areas.[19] Figure 1 reports the distribution of schools from which we have collected data. In *pink* municipalities there are only *Early Enrollees* schools; in *green* municipalities there are only *Late Enrollees* schools; and in *dark blue* municipalities there are both *Early* and *Late Enrollees*.

[insert FIGURE 1 here]

The data set contains information on 4 scholastic years, that is from 2006/2007 until 2009/2010, for the lower four grades of primary school for 18,268 children, both Roma and Non Roma. It contains for each year and for each pupil the final mark in Mathematics, final mark in Serbian, end of year average and number of hours of absences in a year. For the scholastic years 2008/2009 and 2009/2010 we have also semester outcomes for Mathematics, Serbian, average and hours of absences. The data set contains personal characteristics, such as gender, year of birth, month of birth and place of birth.[20] School specific data include school size, number of Roma - in both school and class - and whether the school is in a urban setting.

Tables 2 and 3 summarise respectively the averages of the control variables and main outcomes of interest for Roma and Non Roma children in the pre- and treatment year, 2008/2009 and 2009/2010 respectively.[21]

[insert TABLES 2 and 3 here]

In the pre-treatment year the mean characteristics of the schools that were enrolled in the programme later (column 2) resemble those of the schools that enrolled first (columns 1). The

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

[19] 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.

[20] It is worthy to mention that 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-Jugoslavia. Many Roma refugees in Serbia, for instance, come from Kosovo.

[21] The same tables are obtained for the years 2006/2007 and 2007/2008 and they are available upon request.

tables show 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 Enrollees* schools is found also in the treatment year, providing some support for our claim that *Early Enrollees* and *Late Enrollees* are comparable.[22] The same holds for the assistants: *Early Enrollees* and *Late Enrollees* assistants are comparable in terms of observable characteristics (see Table 4).[23]

[insert TABLE 4 here]

Three important aspects need to be stressed when comparing Roma and Non Roma children. On a grading scale of 1 to 5, the difference of almost two grades between Roma and Non Roma pupils in Serbian and Mathematics is very large: for instance, the average in Mathematics for Roma in *Late Enrollees* is 2.37 in 2008/2009 whereas it is 4.17 for Non Roma; for Serbian it is 2.55 for Roma and 4.33 for Non Roma. Moreover, dropouts seem to be almost exclusively of Roma children: in 2008/2009 in *Late Enrollees* schools 1.9% among Roma children dropout while among Non Roma children only 0.06% did it. Lastly, Roma children show to be absent from school approximately three to four times as much as Non Roma children. In terms of absences, a Non Roma child is absent from school approximately 40 hours in a year, while a Roma child misses school somewhere between 118 and 155 hours in a year.

By simply comparing the averages of outcomes of pre- and treatment year in the two types of schools, we can see that for Roma children there is both a slight improvement in all marks and a decrease in the percentage of students with an insufficient average. These effects are larger in *Early Enrollees* than in *Late Enrollees*. Dropouts almost double in the last year in both types of school. The reason for this sharp increase is likely related to the liberalisation of the visa regime with the European Union which induced a certain number of Roma families to migrate to the EU. Finally, absences increase in 2009/2010 in both *Early Enrollees* and *Late Enrollees* for both Roma and Non Roma, but for Roma they increase by less in *Early Enrollees* schools.

Our data also allow to see whether inequality in marks − the difference between higher and lower marks − decreases as a response to the programme.

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

[23] Unfortunately we have not got demographic information about the assistants besides the gender.

We use both Roma and Non Roma children's marks to calculate our Gini inequality index.[24] As usual, under perfect equality the Gini coefficient equals 0; on the other hand, Gini index is equal to 1 when there is total inequality. To calculate the Gini index we use the common formula for income inequality, but we replace the income with pupils' grades.[25] A general improvement in marks of pupils would translate into a reduction of the Gini coefficient and hence a reduction in inequality.

Figure 2 shows the trends of the Gini coefficient over the four year period − from 2006 to 2010 − in Serbian and Mathematics.

[insert FIGURE 2 here]

The graph suggests that inequality is lower in *Late enrolee* schools for both subjects. An encouraging fact emerging from the graph is that inequality is decreasing in both *Late* and *Early enrolee* schools over the period. Nonetheless, we do not see a sharp reduction in the inequality index among *Early enrolee* schools compared to *Late enrolee* schools in the year of the treatment. It is thus unlikely that the programme was effective in reducing inequality in marks in the first year of its implementation.

5 Identification Strategy

The aim of this paper is to evaluate the effects of the Roma Teaching Assistant Programme on educational outcomes of pupils in the first four grades of school. More precisely, we want to examine the impact of the programme on dropouts, attendance and grades of Roma pupils in the first year of its implementation.[26] We intend to address the following questions:

- Does the programme have an impact on Roma pupils' grades?
- Does the programme reduce dropouts rates of Roma pupils?
- Does the programme increase attendance rates of Roma pupils?

The ideal experiment would require having a random selection of the schools assigned to the programme. Unfortunately, we are not in this setting: schools were not chosen randomly to

[24] Similar results are obtained using the Atkinson, Theil index and decomposing the aforementioned indices in within and between-group inequality.

[25] The Gini formula is given by $G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2n^2\overline{x}}$, where n is the number of individuals, x_i the mark in Serbian or Mathematics of individual i and $\overline{x}$ the mean mark in Serbian or Mathematics of the whole population.

[26] The dataset reduces here to only Roma children. Non Roma children are not included in the following analysis.

participate. Nonetheless, the gradual implementation of the programme allows us to base the evaluation on a comparison of *Early* and *Late Enrollees*. Our treatment group are 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*). There is certainly the concern that schools and assistants starting the programme in the two different years may differ because they had to apply in order to get selected for the programme. Although the selection criteria remain almost the same, we do not know what motivates schools to apply before others and whether these motivations are related to differences in the principle or in school quality.[27] We do know, though, that observable characteristics do not differ between those schools which applied in the first year and those in the second year: 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 10.5% compared to 12.2%.[28] Moreover, we do know that the committee which decided the schools selected - composed by the Minister of Education and other representatives of the Ministry, representatives of National Council, OSCE and of the Ministry for Human and Minority Rights - gave priority to schools in the poorest municipalities or with huge Roma settlements[29] and rated them based on their shown interested and motivation (application) in the same way, in both years. The same holds for the assistants.

Placebo tests are one possibility to ensure the robustness of our results. Another possibility is to compare older cohorts less exposed to the programme (control group) to younger cohorts (treated group) exposed to the programme from *Early Enrollees*-treated schools.

The advantages and disadvantages of both control groups need to be mentioned. The first control group consists of schools which enroled later in the programme. The main advantage of this group is that the impact of the programme would not be confounded with other government policies which 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 disadvantage of this control group, as mentioned before, is that we are not able to control for unobservable differences which

[27] 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 - were encharged to inform schools directly.

[28] Schools which applied in the first year could also apply in the second year. 45% of schools which applied in 2009, and did not get selected, did it again in 2010 and two third of them got selected in the second year (16 out of 24 schools applying in both years). Among these selected schools in 2010, only 2 schools, corresponding to 15%, is present in our subsample.

[29] Subotica, Novi Sad, Niš, Kragujevac, Belgrade.

have led some schools to enter the programme before other schools. Using older cohorts in the treatment schools as a control group obviously eliminates our possible problem of selection bias. Nonetheless, this identification 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 over the years. A possible way to better take into account strengths and weaknesses of both approaches may be therefore combining the *Early - Late Enrollees* analysis with the cohort one: we look at the triple difference between cohorts of treated and control schools.

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

Our first identification strategy exploits the fact that some schools received assistants prior to other schools. We compare *Early Enrollee* schools with *Late Enrollees* schools in the years 2008/2009 - when there was no programme - and 2009/2010 - when the programme got introduced.

5.1.1 Average treatment approach

Our specification (1) is the difference-in-difference model with school fixed effects without any control:

$$Y_{ijt} = \beta_0 + \delta_t + \rho_j + \beta_1 treatment_j * post_t + \varepsilon_{ijt} \tag{1}$$

The outcome variables Y_{ijt} are final marks in Serbian and Mathematics, probability to dropout and hours of absences 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.

Specification (2) includes control variables X'_{ijt}:

$$Y_{ijt} = \beta_0 + \delta_t + \rho_j + \beta_1 treatment_j * post_t + \beta_2 X'_{ijt} + \varepsilon_{ijt} \tag{2}$$

Our 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 s/he attends school). With school

fixed effects we are able to control for time-invariant unobservable school characteristics as well as unobservable geographical characteristics.

The coefficient of interest 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.

The results of the regressions for the different outcomes of interest are reported in Table 5.[30] 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). It is reasonable to expect differences in the impact of government interventions due to different scholastic achievements by gender and different social roles attached to the different sexes in Roma culture. For this reason we also control for the gender of the assistant, but it does not turn out to be significant in any specification.[31] Then we split the sample by grade to examine the presence of possible differential effects (columns 5 to 8). We expect that pupils from the first grade are the most responsive ones for two reasons. First, the assistants work mostly with the first graders: they are lagging less than higher graders. Second, first graders are the ones who do not have any habits about school (e.g. attendance and doing homework), that is, they are the ones who can be influenced most by regular work.[32]

[insert TABLE 5 here]

Overall, the results suggest that the programme had only a statistically significant impact on hours of absences: pupils exposed to the programme were on average 17 hours, corresponding to 0.12 standard deviations, less absent from school than pupils not exposed to it.[33] This is especially the case for male, whose reduction in absences is of 26 hours or equivalently 0.18 standard deviation. Marks in Mathematics and Serbian would suggest that the programme had a positive impact on Roma pupils, but coefficients do not show to be significant. The fact that the inclusion of control variables does not change the magnitude of our coefficients suggests that our coefficients are very robustly estimated.

[30] The coefficients for control variables are not reported. Overall, the higher the percentage of Roma in a class, the worse their average marks and the higher is their hours of absences. Class size is statistically insignificant in all regressions, but school size turns out to be significant in some specifications and, as we would expect, it has a negative impact on marks. Complete results are available upon request.

[31] Results are available upon request.

[32] In an attempt to better understand who exactly has been targeted by the programme, we have split our sample based on whether a child has enrolled at school at the right age or at a higher age. We expected to find that children who started school later were targeted by the programme, but we did not find support for this hypothesis.

[33] On average, Roma pupils are absent from school 143 hours in a year.

We do not believe that there is interaction between children in *Early* and *Late Enrollee* schools. In fact only in three cases children from a control and children from a treatment school live in the same settlement. If one were to believe that there are spillover effects from treated children on children from control schools, this would only imply that coefficients in our regressions are underestimated.

5.1.2 Intensity of treatment approach

It is worth investigating the intensity of the programme. Each school has only one assistant implying that the higher the number of Roma children the less intense is the programme. If the assistant has to follow a high number of students, it is likely that she could follow less each of them: she would be less present both in regular classes and in helping them with their homework and assignments.

The following specification is a variation of the previous approach and it exploits the within school variation of Roma and the fact that the programme intensity depends on the number of Roma in a school. Schools are divided in two groups and each group has an equal number of schools.[34] One group contains 19 schools with less Roma and the other group contains 19 schools with a higher number of Roma.[35] The main difference to the prior model is that we interact the dummy (*more_Roma*) with treatment and time. The coefficient of interest is now β_6.

The intensity of treatment is modeled as follows:

$$\begin{aligned} Y_{ijt} = \beta_0 + \delta_t + \beta_1 treatment_j + \beta_2 treatment_j * post_t + \beta_3 more\_Roma_{jt} + \\ + \beta_4 more\_Roma_{jt} * post_t + \beta_5 more\_Roma_{jt} * treatment_j + \\ + \beta_6 more\_Roma_{jt} * treatment_j * post_t + \varepsilon_{ijt} \end{aligned} \tag{3}$$

The regressions results are in Table 6. Again, we have two specifications: without controls (1) and with controls (2) and we look at the impacts by gender.[36]

[34] Similarly, we have tried to divide the schools in four quantiles. However, with four quartiles the data did not pass balance tests for pretreatment year and we have therefore decided to group schools in less - two groups.

[35] The average number of Roma between the two years - pre- and treatment year - is used in order to define the two groups. The threshold is here 50 pupils. We believe that differenciating the schools in groups helps to better understand the role of the school size on the impacts of the programme. Alternatively, we use the share of Roma in each school and we obtain comparable results.

[36] We are not able to split our sample by grades because our data does not pass some balance tests for the pretreatment year if we divide schools in two groups and by grades.

[insert TABLE 6 here]

The intensity of the programme clearly plays a role in explaining its effects. The lower is the number of Roma in a school, and thus the more the assistant can follow them, the higher is the impact on the outcomes of interest. Absences, for instance, reduce on average by 40 hours in a year in schools with less Roma, compared to *Late Enrollee* schools. These effects disapper 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 remain 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 in Mathematics and Serbian increase respectively by 0.511 (0.44 standard deviations) and 0.345 (0.29 standard deviations).[37]

Note that Roma attending schools with a lower percentage of Roma could be different than Roma in schools with a higher share. One could, for instance, argue that Roma in schools with a lower share are more willing to adapt and assimilate to the majority population. We do believe that there are systematic differences between schools with a lower and a higher share of Roma and therefore 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.

5.1.3 Placebo Regressions

The difference-in-difference approach relies on the parallel trends assumption. That is, 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. The most obvious way to examine the robustness of our results is to run the same regressions (regression (1), (2), and (3)) for the years 2006/2007 versus 2007/2008 and for the years 2007/2008 versus 2008/2009 (Table 9). These two placebo tests allow us to test if treatment and comparison schools are comparable; in other words, in this way we can test if the outcomes in the two groups of schools had a parallel trend before the introduction of the programme.

Significant difference-in-difference coefficients in the case of average treatment approach in the years prior to the introduction of the programme would question the adequacy of our comparison group. The two placebo tests suggest that our results are robust.

[37] Their average mark in Mathematics is 2.56 and in Serbian 2.82 in schools with less Roma.

5.2 Second approach: Cohort regressions and triple difference

In our second approach we try to circumenvent the problem of possible selection bias by using as control schools older cohorts from treated schools. 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. We know that assistants worked mostly with the first grade and our results from the average treatment approach are suggestive that the first grade has benefitted most from the programme. Hence in this section we use a specification which informs us whether the programme was successfull for the children enroled in the first grade.

We first estimate the following regression for *Early Enrollees*:

$$Y_{ijt} = \beta_0 + \beta_1 young_i + \beta_2 post_t + \beta_3 young_i * post_t + \varepsilon_{ijt} \tag{4}$$

where Y_{ijt} are again final marks in Serbian and Mathematics, probability to dropout and hours of absences 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 perfomered compared to the older grades.

The same regression (4) 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:

$$\begin{aligned} Y_{ijt} = \beta_0 + \beta_1 young_i + \beta_2 post_t + \beta_3 young_i * post_t + \gamma_1 treatment_j * post_t + \\ + \gamma_2 young_i * treatment_j + \gamma_3 young_i * post_t * treatment_j + \varepsilon_{ijt} \end{aligned} \tag{5}$$

The regressions are estimated without and with controls, as in previous specifications. We also look at the impacts by gender. Results are in Tables 7 and 8. T below the column number in the table stands for the subsample with only *Early Enrollee* schools, while C stands for the subsample with only *Late Enrollee* schools.

[insert TABLES 7 and 8 here]

When comparing first graders with older pupils in only *Early Enrollee* schools, our coefficients of interest have, with the exception of absences, the correct sign, but are not statistically significant. Pupils in the first grade exposed to the programme get higher marks than first

graders in control schools with respect to their older mates. On average, being in a *Early Enrolles* school increases marks in Mathematics and Serbian by respectively 0.28 and 0.35 standard deviations for first graders. Moreover, it reduces their probability to dropout on average by 6.6 percentage points. The coefficient of interest of the cohort analysis in the *Late Enrolles* schools (β_3 in column 5) suggest that in the absence of the programme dropouts take place more among first graders than among older graders, confirming the importance for the assistants to focus primarily on this group of students. One might argue that this phenomenon leads to a selection of better pupils in higher grades so that the two groups of young cohorts and old cohorts would not be comparable. Test statistics for the difference in average dropouts between the two groups suggest that it is not the case here.[38] Participation of assistants in regular lessons, organisation of additional classes, their assistance with homework and visits to parents help children to perform better at school. Surprisingly, hours of absences increase in both treated and control schools: we would have expected them to reduce because of the programme. A possible reason for this finding is that some Roma families have migrated to European countries and have withdrawn their children from school as a consequence of the new visa liberalisation regime. However, we find that the increase in absences is smaller in *Early Enrollees* than in *Late Enrollees* schools, confirming the results obtained by the average treatment approach also in the magnitude of the difference (32 hours). Although the coefficient γ_3 of column 6 is not statistically significant, it suggests that, without the programme, the absences would have possibly had an even larger increase. Overall, the effects on the triple difference between cohorts in treated and control schools seem to be mainly driven by differences between schools than differences between cohorts in the same school.

5.2.1 Placebo Regressions

Again, we need to control for the robustness of our results by running regressions (regression (4), (5)) for the years 2006/2007 versus 2007/2008 and for the years 2007/2008 versus 2008/2009 (see Table 10). The two placebo tests suggest that our results are robust.

[38] The p-value of the test for the difference between the means in dropouts is 0.3860 for the control schools and 0.1593 for the treated schools.

5.3 Spillover effects

It remains to investigate whether this programme affected also Non Roma pupils. We employ both identification strategies and their combination and find that neither marks improved nor absences reduced for Non Roma students. The presence of a Roma assistant do not improve Non Roma schooling outcomes.[39]

[insert TABLES 11 and 12 here]

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.

6 Conclusion

In this paper we estimate the impact of the Roma Teaching Assistant Programme in its first year of implementation on different outcomes of interest. We use a difference-in-difference approach. Our first estimation strategy exploits the fact that the introduction of assistants in schools was gradual: some schools entered the programme before others. Given that schools and assistants needed to apply to the programme, a problem of potential selection bias may arise: there are unobservable characteristics, likely related to differences in the principle or in school quality, we cannot control for. In order to circumvent this problem, we use a second identification strategy. We compare pupils of the first grade from treated schools with older cohorts from the same schools. This identification strategy controls well for schools specific characteristics, but we are not able to control for government interventions which might have taken place over the period. Therefore, we 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.

The comparison of Roma and Non Roma schooling outcomes suggest that the current differences in marks between Roma and Non Roma are enormous. Moreover, Roma children are much more absent from school than their Non Roma colleagues. The stark differences in all main schooling outcomes underline the necessity in this context for government programmes aiming at helping Roma pupils.

Results of our analysis suggest 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.

[39] Placebo tests for the pre-treatment year (2007/2008) are here reported.

There is evidence that children in treated schools went on average at least 17 hours more to school. Our results suggest that, on average, marks have improved by almost 0.28 standard deviations in mathematics, 0.35 standard deviation in Serbian and dropouts have reduced by 6.6 percentage points for those children exposed to the programme in their first grade. Higher and more systematic impacts are obtained in schools with a lower number of Roma: the higher is their number, the less the assistant can follow them, and the lower is the impact of the programme on the outcomes of interest. This seems to be especially the case for females, for whom being in a school with a lower number of Roma turns out to be more favourable. Marks of girls in schools with less Roma improve by 0.44 standard deviations in mathematics and 0.29 standard deviations in Serbian.

While first graders in treated school perform better than their older colleagues, overall the programme does not seem to have a significant impact on pupils' achievement. This is likely due to the fact that assistants work mainly with lower grades and that 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 has looked 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. Nevertheless, our results suggest that the programme is more effective in schools with less Roma. It would be worth rethinking to assign more than one assistant to schools with a large number of Roma.

References

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

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

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

Boyd, D., P. Grossman, H. Lankford, S. Loeb, and J. Wyckoff (2008). Teacher Preparation and Student Achievement:. *Working Paper no.14314*.

Das, J., S. Dercon, J. Habyarimana, P. Krishnan, K. Muralidharan, and V. Sundararaman (2011). School Inputs, Household Substitution, and Test Scores.

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

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

Hanushek, E. A., J. F. Kain, and S. G. Rivkin (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.

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

Jacob, B. A. and L. Lefgren (2004). Remedial education and student achievement: A regression discontinuity analysis. *Review of Economics and Statistics 86*(1), 226–244.

Kertesi, G. and G. Kezdi (2011). The roma/non-roma test score gap in hungary. *American Economic Review: Papers & Proceedings 101:3*, 519–525.

Lauer, P. A., M. Akiba, S. B. Wilkerson, H. S. Apthorp, D. Snow, and M. L. Martin-Glenn (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. and A. Schlosser (2005). Targeted Remedial Education for Underperforming Teenagers: Costs and Benefits. *Journal of Labour Economics 23*(4), 839–874.

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

Open Society Institute (2007). Equal Access to Quality Education For Roma.

Open Society Institute (2008). International Comparative Data Set.

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

Roland G. Fryer, J. (2010). Racial Inequality in the 21st Century: The Declining Significance of Discrimination. *NBER, Working Paper 16256*.

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

Schultz, P. T. (2004). School subsidies for the poor: evaluating the Mexican Progresa poverty program. *Journal of Development Economics 74*(1), 199–250.

Schweinhart, L. J., H. V. Barnes, and D. P. Weikart (1993). Significant Benefits: The High/Scope Perry Preschool Study Through Age 27. Technical report, High Scope Press, Ypsilanti, MI.

Vermeersch, C. and M. Kremer (2004). School Meals, Educational Achievement and School Competition: Evidence from a Randomized Evaluation. *World Bank Policy Research Working Paper No. 3523*.

7 Tables

Table 1: Programme timeline

	2009 Early Enrollees	2010 Late Enrollees
Number of schools applying to the programme	128	252
Number of assistants applying to the programme	158	329
Number of schools joining the programme	26	77
Number of schools in our sample	23	15

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

	Pretreatment year			Treatment year		
	Early Enrollees (1)	Late Enrollees (2)	Difference (1-2) (3)	Early Enrollees (1)	Late Enrollees (2)	Difference (1-2) (3)
Female:						
Roma	0.5	0.47	0.03	0.49	0.47	0.02
			(0.02)			(0.26)
Non Roma	0.49	0.48	0.01	0.47	0.49	-0.02
			(0.014)			(0.013)
Born in same town:						
Roma	0.86	0.81	0.05	0.88	0.81	0.07*
			(0.04)			(0.35)
Non Roma	0.92	0.91	0.01	0.93	0.92	0.01
			(0.011)			(0.011)
Roma per School	0.22	0.19	0.03	0.19	0.23	-0.04
			(0.06)			(0.06)
School size	305	361	-56	301	363	-62
			(52.96)			(56.04)
No. of Roma per Class	4.91	4.39	0.52	5.25	4.49	0.76
			(1.33)			(1.48)
No. of Roma per Class	5.56	4.64	0.92	5.9	4.6	1.3
(if at least a Roma)			(1.35)			(1.48)
Class size	22.16	23.97	-1.8	22.44	24.21	-1.77
			(1.42)			(1.38)
Number of schools	23	15		23	15	
Number of Roma pupils	1241	811		1268	847	
Number of Non Roma pupils	4303	3374		4122	3514	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors corrected for clustering at the school level are reported in parentheses.

Table 3: Averages of outcomes in pre- and treatment year for Roma and Non Roma pupils

	Pretreatment year			Treatment year		
	Early Enrollees (1)	Late Enrollees (2)	Difference (1-2) (3)	Early Enrollees (1)	Late Enrollees (2)	Difference (1-2) (3)
Mathematics:						
Roma	2.28	2.37	-0.9	2.36	2.40	-0.04
			(0.05)			(0.05)
Non Roma	4.25	4.17	0.07	4.3	4.2	0.01
			(0.02)			(0.02)
Serbian:						
Roma	2.43	2.55	-0.12	2.49	2.56	-0.7
			(0.05)			(0.05)
Non Roma	4.4	4.33	0.07	4.43	4.34	0.09
			(0.02)			(0.02)
Dropout:						
Roma	0.015	0.019	-0.004	0.026	0.035	-0.009
			(0.006)			(0.007)
Non Roma	0.001	0.0006	0.0004	0.001	0	0.001**
			(0.006)			(0.0005)
Absences (hours):						
Roma	118	125	-7	134	155	-21
			(6.51)			(6.74)
Non Roma	39	36	3	42	40	2
			(1.02)			(0.97)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors corrected for clustering at the school level are reported in parentheses.

Table 4: Characteristics of the assistants

	Early Enrollees (1)	Late Enrollees (2)	Difference (1-2) (3)
Female	0.52	0.5	0.02
			(0.18)
Maximum level of education			
Secondary school	0.65	0.58	0.07
			(0.18)
University	0.35	0.33	0.02
			(0.17)
Experience with Roma	0.91	0.67	0.24
			(0.15)
Experience in NGO	0.91	1	-0.9
			(0.06)
Number of assistants	23	12	

Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We could not get information about three assistants of the *Late Enrollee* schools.

Table 5: Average treatment approach

	(1) without controls	(2) with controls	(3) female	(4) male	(5) first grade	(6) second grade	(7) third grade	(8) fourth grade
MATHEMATICS								
post	0.051	0.065	0.096	0.041	-0.076	0.040	0.208*	0.168*
	(0.069)	(0.062)	(0.080)	(0.056)	(0.175)	(0.126)	(0.117)	(0.085)
treatment*post	0.046	0.030	0.015	0.053	0.213	-0.085	0.040	-0.156
	(0.081)	(0.077)	(0.091)	(0.085)	(0.221)	(0.192)	(0.146)	(0.158)
SERBIAN								
post	0.039	0.046	0.079	0.027	-0.048	0.099	0.222***	0.109
	(0.060)	(0.048)	(0.055)	(0.050)	(0.132)	(0.133)	(0.079)	(0.076)
treatment*post	0.044	0.012	-0.035	0.058	0.121	-0.130	-0.110	-0.075
	(0.069)	(0.066)	(0.075)	(0.080)	(0.177)	(0.207)	(0.097)	(0.177)
No. observations	4085	3961	1916	2045	989	1111	988	873
DROPOUT								
post	0.017**	0.015**	0.001	0.027**	0.067***	-0.014*	0.011***	-0.007
	(0.007)	(0.006)	(0.010)	(0.012)	(0.016)	(0.008)	(0.004)	(0.011)
treatment*post	-0.006	0.003	0.028*	-0.018	-0.037	0.027*	-0.004	0.045**
	(0.009)	(0.009)	(0.014)	(0.014)	(0.025)	(0.013)	(0.011)	(0.020)
No. observations	4167	4039	1951	2088	1005	1140	1009	885
ABSENCES								
post	31.236***	32.853***	22.456***	42.034***	64.693**	19.900	21.515*	-0.007
	(7.856)	(9.078)	(10.797)	(10.764)	(28.341)	(23.608)	(17.049)	(0.020)
treatment*post	-17.299**	-16.679*	-4.713	-26.119**	-28.336	-19.406	-10.048	0.045**
	(7.856)	(9.078)	(10.797)	(10.764)	(28.341)	(23.608)	(17.049)	(0.020)
No. observations	3980	3868	1871	1997	945	1084	980	885

Robust standard errors corrected for clustering at the school level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Intensity of treatment

	(1) without controls	(2) with controls	(3) female	(4) male
MATHEMATICS				
treatment*post	0.402***	0.339***	0.511***	0.173*
	(0.129)	(0.113)	(0.149)	(0.100)
more_Roma*treatment*post	-0.454***	-0.417***	-0.680***	-0.161
	(0.152)	(0.143)	(0.179)	(0.138)
SERBIAN				
treatment*post	0.303*	0.250**	0.345**	0.149
	(0.150)	(0.121)	(0.159)	(0.113)
more_Roma*treatment*post	-0.330*	-0.307**	-0.500**	-0.109
	(0.170)	(0.150)	(0.186)	(0.151)
No. observations	4085	3961	1916	2045
DROPOUT				
treatment*post	0.001	-0.000	0.053	-0.046
	(0.024)	(0.022)	(0.042)	(0.032)
more_Roma*treatment*post	-0.008	-0.001	-0.044	0.038
	(0.026)	(0.023)	(0.044)	(0.035)
No. observations	4167	4039	1951	2088
ABSENCES				
treatment*post	-28.012	-40.478*	-7.635	-68.816*
	(25.997)	(22.323)	(34.882)	(34.882)
more_Roma*treatment*post	19.312	36.390	9.521	59.759**
	(27.112)	(24.179)	(36.097)	(24.684)
No. observations	3980	3868	1871	1997

Robust standard errors corrected for clustering at the school level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Cohort regression - A

	without controls			with controls			female			male		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
MATHEMATICS	T	C	ALL	T	C	ALL	T	C	ALL	T	C	ALL
young	0.108	0.172	0.213	-0.141	-0.104	-0.077	-0.141	-0.063	-0.105	-0.119	-0.161	-0.043
	(0.146)	(0.127)	(0.152)	(0.150)	(0.145)	(0.132)	(0.134)	(0.224)	(0.215)	(0.251)	(0.230)	(0.171)
post	0.060	0.121*	0.162	0.083	0.106	0.132**	0.116**	0.165*	0.190**	0.067	0.072	0.098*
	(0.048)	(0.061)	(0.103)	(0.050)	(0.062)	(0.054)	(0.049)	(0.087)	(0.071)	(0.070)	(0.062)	(0.058)
young*post	0.070	-0.324**	-0.365**	0.082	-0.241	-0.291*	-0.019	-0.344	-0.426*	0.180	-0.205	-0.239*
	(0.131)	(0.147)	(0.177)	(0.116)	(0.143)	(0.152)	(0.132)	(0.234)	(0.241)	(0.155)	(0.142)	(0.135)
treatment*post			-0.127			-0.059			-0.080			-0.045
			(0.145)			(0.074)			(0.088)			(0.086)
young*treatment			-0.131			-0.102			-0.041			-0.157
			(0.241)			(0.198)			(0.240)			(0.250)
young*post*treatment			0.460*			0.381*			0.428			0.412*
			(0.235)			(0.194)			(0.279)			(0.207)
SERBIAN												
young	0.173	0.290***	0.340***	-0.073	-0.011	0.035	-0.175	-0.035	-0.060	0.040	0.004	0.130
	(0.110)	(0.078)	(0.087)	(0.118)	(0.154)	(0.119)	(0.117)	(0.160)	(0.145)	(0.192)	(0.278)	(0.202)
post	0.047	0.123**	0.173	0.051	0.077	0.118**	0.037	0.104	0.159**	0.082	0.079	0.106**
	(0.053)	(0.056)	(0.110)	(0.056)	(0.050)	(0.045)	(0.051)	(0.072)	(0.063)	(0.081)	(0.049)	(0.041)
young*post	0.058	-0.391***	-0.441***	0.079	-0.255**	-0.300***	0.052	-0.284	-0.352**	0.101	-0.295*	-0.328**
	(0.115)	(0.109)	(0.152)	(0.102)	(0.104)	(0.101)	(0.113)	(0.173)	(0.170)	(0.161)	(0.154)	(0.139)
treatment*post			-0.158			-0.079			-0.131			-0.043
			(0.154)			(0.074)			(0.086)			(0.086)
young*treatment			-0.199			-0.147			-0.123			-0.168
			(0.167)			(0.138)			(0.154)			(0.222)
young*post*treatment			0.531**			0.382**			0.416*			0.423*
			(0.212)			(0.149)			(0.215)			(0.212)
No. observations	2462	1624	4086	2395	1567	3962	1180	736	1916	1215	831	2046

Robust standard errors corrected for clustering at the school level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Cohort regression - B

	without controls			with controls			female			male		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
DROPOUT												
young	-0.020***	-0.014*	-0.012**	-0.002	0.011	0.002	-0.002	0.013	0.009	-0.005	0.008	-0.008
	(0.004)	(0.007)	(0.006)	(0.006)	(0.014)	(0.007)	(0.013)	(0.017)	(0.017)	(0.010)	(0.015)	(0.009)
post	0.003	-0.004	-0.002	0.015**	-0.002	-0.003	0.021*	-0.019	-0.016	0.009	0.010	0.007
	(0.006)	(0.006)	(0.007)	(0.006)	(0.010)	(0.006)	(0.010)	(0.012)	(0.010)	(0.009)	(0.015)	(0.014)
young*post	0.030**	0.080***	0.078***	0.014	0.079***	0.080***	0.031	0.071***	0.077***	-0.003	0.087***	0.083***
	(0.014)	(0.024)	(0.023)	(0.013)	(0.019)	(0.018)	(0.020)	(0.019)	(0.018)	(0.012)	(0.028)	(0.028)
treatment*post			0.005			0.018**			0.038**			0.003
			(0.009)			(0.009)			(0.015)			(0.016)
young*treatment			-0.009			0.005			-0.003			0.015
			(0.006)			(0.006)			(0.016)			(0.012)
young*post*treatment			-0.047*			-0.066***			-0.047*			-0.087***
			(0.027)			(0.022)			(0.027)			(0.029)
No. observations	2509	1658	4167	2438	1601	4039	1200	751	1951	1238	850	2088
ABSENCES												
young	-21.835***	-1.674	-0.141	3.123	37.687*	18.656	-8.373	39.502	26.379	12.221	32.323	8.644
	(6.679)	(14.939)	(13.725)	(8.355)	(20.554)	(17.503)	(11.240)	(23.527)	(21.174)	(14.033)	(23.473)	(17.397)
post	6.732	19.031**	20.564	10.022	30.724***	20.728**	7.655	21.190*	8.570	13.007	37.766***	30.488***
	(5.386)	(8.406)	(13.534)	(5.975)	(9.912)	(7.856)	(7.172)	(10.707)	(8.771)	(8.865)	(11.015)	(9.396)
young*post	37.800***	45.788*	44.255*	23.579*	57.493**	54.639**	40.707**	67.618**	67.603**	9.740	52.423**	49.480**
	(11.139)	(24.209)	(24.838)	(11.537)	(21.263)	(22.235)	(16.000)	(28.938)	(26.469)	(14.055)	(22.036)	(24.287)
treatment*post			-14.801			-10.409			-0.785			-17.036
			(18.481)			(10.354)			(11.898)			(12.891)
young*treatment			-22.663			-0.868			-19.597			16.848
			(17.445)			(16.384)			(19.622)			(16.766)
young*post*treatment			-5.487			-31.867			-28.524			-40.337
			(28.803)			(24.945)			(30.772)			(27.852)
No. observations	2394	1586	3980	2336	1532	3868	1152	719	1871	1184	813	1997

Robust standard errors corrected for clustering at the school level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Average treatment approach - Placebo

	2006/2007 and 2007/2008			2007/2008 and 2008/2009		
	(1)	(2)	(3)	(4)	(5)	(6)
	all	female	male	all	female	male
MATHEMATICS						
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)
SERBIAN						
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)
No. observations	3585	1750	1835	3816	1876	1970
DROPOUT						
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

Robust standard errors corrected for clustering at the school level in parentheses:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Cohort regression - Placebo

	2006/2007 and 2007/2008			2007/2008 and 2008/2009		
	(1)	(2)	(3)	(1)	(2)	(3)
	T	C	ALL	T	C	ALL
MATHEMATICS						
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)
SERBIAN						
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)
No. observations	2232	1354	3586	2364	1482	3846
DROPOUT						
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

Robust standard errors corrected for clustering at the school level in parentheses:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Average treatment approach - Non Roma

				placebo 2007/2008		
	(1)	(2)	(3)	(4)	(5)	(6)
	Mathematics	Serbian	Absences	Mathematics	Serbian	Absences
post	0.011	-0.011	5.025***	0.058**	0.052**	-0.518
	(0.018)	(0.031)	(1.185)	(0.027)	(0.025)	(2.874)
treatment*post	0.029	0.054	-1.586	-0.043	-0.029	3.749
	(0.025)	(0.034)	(1.725)	(0.033)	(0.029)	(3.106)
No. observations	14981	14982	14686	15345	15345	15052

Robust standard errors corrected for clustering at the school level in parentheses:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Cohort regression - Non Roma

	Mathematics			Serbian			Absences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	T	C	ALL	T	C	ALL	T	C	ALL
young*post	0.044	-0.062	-0.070	0.033	-0.056	-0.066	0.658	3.178	3.259
	(0.058)	(0.085)	(0.079)	(0.072)	(0.077)	(0.069)	(2.682)	(3.274)	(3.057)
young*post*treatment			-0.010						
			(0.010)						
No. observations	8232	6749	14981	8232	6750	14982	8099	6587	14686
				placebo 2007/2008					
	Mathematics			Serbian			Absences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	T	C	ALL	T	C	ALL	T	C	ALL
young*post	-0.099*	-0.056	-0.053	-0.118**	-0.089	-0.087	0.610	-1.718	-2.440
	(0.050)	(0.104)	(0.099)	(0.047)	(0.101)	(0.097)	(2.847)	(3.922)	(3.935)
young*post*treatment			-0.049			-0.037			2.787
			(0.111)			(0.107)			(4.986)
No. observations	8573	6772	15345	8573	6772	15345	8435	6617	15052

Robust standard errors corrected for clustering at the school level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 1: Location of the schools with assistants

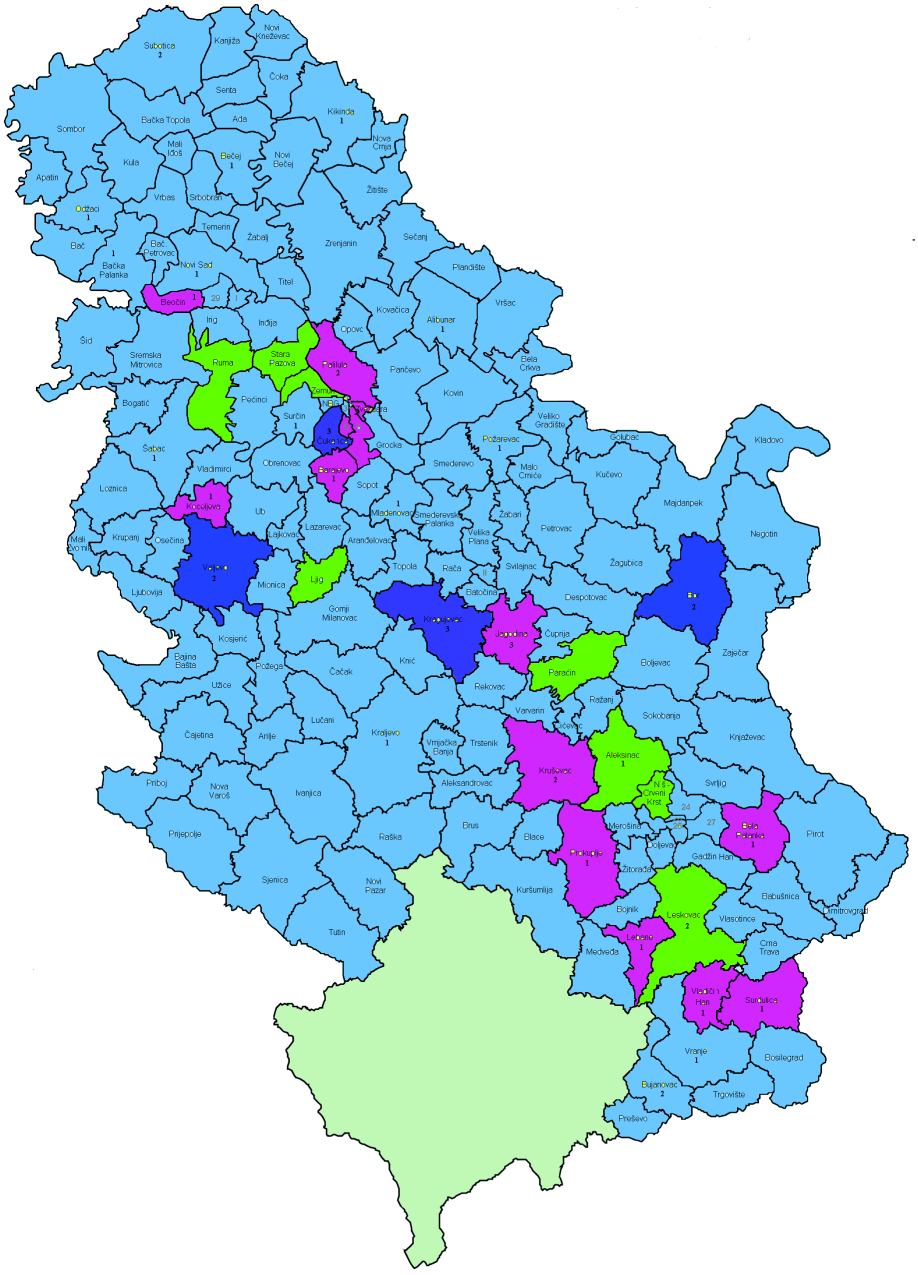

Figure 2: Gini coefficient of inequality in Serbian and Mathematics