The Curse of Low Aspirations: Remedial Education and Perceived Returns to Education of Roma People *

Job Market Paper

Marianna Battaglia

Bocconi University

Lara Lebedinski

Bocconi University

Abstract

We examine how a remedial education program for primary school-age children affects parental aspirations about their children's future. Using original survey data we collected in Serbia, we investigate whether expectations on labor market perspectives and educational achievement change as a consequence of exposure to the Roma Teaching Assistant Program. We argue that these changes are likely to occur mainly through a role model mechanism: in the program all the assistants are Roma and from the same social background of the pupils they help. The presence of a person belonging to the same community, who proved to be successful, motivates parents to believe their children can succeed. Our results show that parents of pupils in treated schools expect higher returns to education for their kids. They are also more likely to expect them to achieve a secondary level of education.

Keywords: aspirations, perceived returns to education, program evaluation

JEL classification codes: I25, J13, D04

1   Introduction

Aspirations for the future consistently affect choices made in the present. The social environment where one lives plays a role in shaping aspirations. The aim of this paper is to examine the impact on parental aspirations of a remedial education program for primary school-age children targeting the marginalised Roma minority group. We investigate whether expectations on labor market perspectives and educational achievement change as a consequence of exposure to the Roma Teaching Assistant Program (RTA), a remedial education program introduced in Serbia in 2009. We find that parents whose children are exposed to the program expect higher returns to education for their kids. They are also more likely to expect them to achieve a secondary level of education. Moreover, an examination of heterogeneous effects suggests the following. First, results on highest expected level of education are driven by responses from Non-Muslim parents and parents living in *mixed (Roma and Non-Roma)* neighborhoods. Second, parents revise their expectations in response to the program mainly for younger kids (6 to 10 years). We argue that these changes are likely to occur through a role model mechanism. In the RTA program, all the assistants are Roma and from the same social background of the pupils they help. They can share their successful experience with them and with their parents who will be motivated to believe that their children can achieve analogous results. Households may thus respond to changes in perceived returns when making schooling decisions and increase current investment in education. The rise in aspirations can shape educational and career choices.

Roma people attain very low education. In most countries enrollment rates in primary school among Roma are in the range of 40% to 60%. Their completion rates are even lower: only 30% to 40% of Roma adults have completed compulsory primary education.[1] Underinvestment in education can be due to financial constraints. However, Roma people

---

[1]There are reasons to believe that these numbers are upper bounds. First, some schools keep children who do not come to schools in their school books. Second, a large number of Roma finish evening schools or special schools which count as finished primary schools although the requirements in these schools are much lower. 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 they are mainly attended by pupils who are late in enrolling and by pupils who decided to return to school after dropping out.

may not invest in education because they do not expect schooling to give them enough future opportunities. In the formal job market there is often discrimination against minority groups. The informal job market does not often require any level of education. Roma people are mainly involved in casual and seasonal jobs, performed without any written contract, e.g. they collect rubbish, sell goods on the market or do low skilled jobs. The cost of investing in education is too high as compared to the discounted stream of expected future benefits. Therefore, there is no incentive to invest. A vicious cycle of low aspirations arises.

For the purpose of our analysis, we have conducted an extensive survey with 300 Roma households in Belgrade. In Fall 2010 we interviewed both parents and their children in 12 different settlements of the city. The RTA program started in 2009 and we look at its impact a year after its implementation. The program was introduced gradually: some schools received their teaching assistant before others. Parents and children who attend schools with the teaching assistant in Fall 2009 are our treated group. Parents and children who attend schools that received the teaching assistant at a later point in time (Fall 2010) are our control group.[2] The allocation of Roma teaching assistants was not randomised: schools and assistants needed to apply to be part of the program. We therefore need to tackle the possible problem of selection bias.

The importance of role models for minorities is not new in the education literature. A series of researchers and policy makers in the 90s was pushing for an increased hiring of minority teachers in the US (Graham, 1987; Ladson-Billings, 1994). In fact, the importance of having a teacher with the same background has been found significant in improving the achievement gap for minorities (Dee, 2004). The Black-White mark gap has been intensively investigated in the United States[3] and the inclusion of identity in

---

[2] All schools involved in the program are public schools.

[3] The first and most known program is the Perry Preschool program introduced in 1962: it targeted children from disadvantaged socioeconomic backgrounds and consisted of a 2-5-hour daily preschool program for children aged three years old and weekly home visits by teachers. Attempts have been also made during the primary school through the introduction of after-school programs (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 interventions is mixed: certain programs have left the racial

economics has been widely recognised (Akerlof and Kranton, 2000). However, to the best of our knowledge, this is the first study addressing changes in aspirations of minority groups arising as a consequence of having a teacher from the same community. Only Krishnan and Krutikova (2010) evaluate the long-term effects of an after-school program for children living in the slums of Bombay and find rather weak evidence on expected life evaluation and aspirations. Nonetheless, they do not look at a minority group and elicit directly from children their role models, whereas we argue that assistants of the RTA program are perceived as such. Our paper, together with its companion paper (Battaglia and Lebedinski, 2011), adds evidence on short-term effects of remedial education programs on minority groups. It also suggests replicable examples in contexts where minorities suffer low attainment rates and social exclusion. For Roma people, for instance, this is the case in many other European countries and so far there are few attempts to investigate how to improve their life circumstances, in general, and of children, in particular. Furthermore, we contribute to the existing literature by providing primary data in a context where data are scarce.

Our paper is in line with the contributions of Ray (2004), Genicot and Ray (2010), Nguyen (2008), and Beaman et al. (2012). We know that individuals' desires and their standards of behavior depend, in part, on past experiences and from observing their peers. In societies where the poor do not observe someone with their similar background succeeding, downward mobility and underinvestment in education are expected. A reasonable distance between one's current standards of living and where one wants to be motivates her to believe she can succeed. Our paper is also linked to the strand of literature on subjective expectations and information gap between perceived and actual returns to schooling.[4] Standard economic theory suggests that, in the presence of perfect information, individuals choose their level of education by equating the marginal benefits of education to its marginal costs. Underinvestment in education can be due to credit

---

achievement gap essentially unchanged.

[4]Literature suggests that this gap can be filled also by providing additional information through statistics (Jensen, 2010). These tools turn out to be most cost−effective solutions than incentives, like cash transfers or private school vouchers.

constraints, high discount rates and low school quality.[5] However, several works emphasised the importance of subjective expectations (Manski, 1993; Jensen, 2010; Nguyen, 2008; Kaufmann and Attanasio, 2009). Returns perceived by individuals affect schooling decisions. Yet, perceptions may be inaccurate, due to limited or imperfect information.

Finally, our work is related to the literature on residential segregation and neighbourhood effects that studies the relevance of neighbourhoods and one's peers in influencing socioeconomic outcomes.[6] For instance, segregation of the African Americans has been identified as one of the reasons for the persistence of inner city poverty in the US (Cutler and Glaeser, 1997). Moreover, the neighbourhood where one lives can clearly affect one's labor market (Clark and Drinkwater, 2002; Edin et al., 2003; Bayer et al., 2008; Boeri et al., 2011) and educational outcomes (Card and Rothstein, 2007). Lastly, the ethnic composition of a municipality can be important for the quality of local public goods such as schools (Alesina et al., 1999; La Ferrara and Mele, 2006).

The rest of the paper is organised as follows. Section 2 gives information on the background and the way the survey has been designed. Section 3 describes data and provides some descriptive statistics. Section 4 presents the estimation strategy and results. Section 5 discusses findings and concludes.

2 Background: Context, Roma Teaching Assistant Program and Survey Design

2.1 Roma in Serbia

Roma people are the largest ethnic minority in Europe.[7] They are poorer than other population groups and more likely to fall into poverty and remain poor. They have been experiencing discrimination for centuries in all the countries where they live. Specifically,

---

[5]See Glewwe (2006) for an extensive summary on education in developing countries.

[6]For an excellent review of the literature on neighbourhood effects see Durlauf (2004) and Blume and Durlauf (2006).

[7]The Roma people are mainly located in South Eastern Europe: Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Czech republic, Hungary, Kosovo, Republic of Macedonia, Moldova, Montenegro, Romania, Slovakia, Slovenia and Ukraine.

Roma suffer severe social exclusion in terms of overrepresentation among low skilled jobs and no participation in the political and cultural life and this is persistent over time.

Official data on Roma in Serbia are scarce and inaccurate.[8] Roma people often do not declare themselves as belonging to the Roma minority in surveys. Most of them consider themselves both Roma and Serbian and the question of nationality allows only one answer.[9] Thus, the 2002 Census counts 108,000 Roma, corresponding to 1.44% of the total Serbian population, while estimates suggest a number between 350,000 and 500,000, approximately 4-6% of the overall population (Open Society Institute, 2007).

The Living Standard Measurement Survey (LSMS) from 2003 provides rich information on the living conditions of the Roma population in the country. It is important to note that this survey includes only Roma living in segregated settlements, which according to the 2002 Census is the case for 83% of the Roma population. The numbers from the LSMS are alarming. Two out of three Roma households are poor: their average consumption is below the absolute poverty line.[10] Almost half of the Roma population (40%) is younger than 18 years old[11] and only 71% of children from Roma settlements aged 6 to 15 attend school. Among the adults, 25% have no schooling at all and another 36% have not finished primary school.[12] Conversely, 99% of Non-Roma aged 6 to 15 are enrolled in school and only 13% of adults have not completed primary school. The employment rate among Roma males is very similar to that of the Non-Roma population (69%), but the female employment rate is very low with 34% versus 53%. The LSMS confirms that Roma live in difficult conditions and that they constitute a marginalised

---

[8]This is the case for most Central and Eastern European countries where the majority of the Roma population lives.

[9]The most appropriate approach when asking for one's identity would be to allow for multiple identities, but this approach has been rather uncommon in this type of surveys.

[10]The percentage of the extremely poor among the Roma interviewed is 11.9%. Those who are considered extremely poor are those who cannot satisfy even their basic needs for food.

[11]The average age of Roma people is 25, whereas the average age in Serbia is 42. The average number of children younger than 18 years old is 2.4 per Roma households, while the population average is only 0.9. The average household size of Roma population is of 4.5 household members; the national average of 3.2.

[12]In Serbia, school is compulsory until the age of 15. Children enroll 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 program is compulsory; in 2010 the length of the compulsory preschool has been extended to 9 months. 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 get one teacher per subject.

minority.

Data on education and job market are in line with those of other countries. They underinvest in education due to credit constraints and the existence of barriers of access to education. Roma people often lack the required ID and face financial constraints. On average, in Serbia costs associated with schooling (books and other school material) correspond to almost 2% of yearly household income (LSMS 2003). In our sample of Roma people, they correspond to almost 6% of yearly household income.[13] Some children face difficulties at school due to language barriers, they are engaged in child labor and suffer discrimination from teachers and pupils.[14] However, there are many reasons to believe that the lack of goals and aspirations is an important factor influencing the educational decision of Roma people. First, a large percentage of them live in segregated settlements. Since they are isolated from the mainstream society, they do not often have different models to which they can relate to in their immediate neighbourhood (this idea is line with the argument of Wilson (1987)). According to the 2002 Census, 83% of self-declared Roma live in census tracts with at least 7% of the Roma population in Serbia. Second, it is extremely rare that Roma people perform jobs for which high levels of education are required. For instance, in Serbia there are usually no teachers of Roma origin working in schools. They can barely be found in any public office.[15] Third, there is evidence that the mean earnings of Roma workers are lower than those of Non-Roma workers, especially for higher levels of education. Figure 1 reports descriptive data for the city of Belgrade. We needed to use two sources of data because the Serbian Statistical Office only collects data for jobs in the formal sector that are usually not performed by Roma. There are no official data on wages of Roma people given they mainly perform informal activities. Therefore, data for Roma are obtained from our sample, while data for Non-Roma are obtained from the Serbian Statistical Office. The figure intends to simply provide a picture of the

---

[13]For 10% of them these costs even ranged between 12% to 25% of yearly household income.

[14]Some children have a limited knowledge of Serbian: in a survey conducted by UNICEF - Multiple Indicator Cluster Survey, 2006 - only 10% of Roma declare Serbian to be their mother tongue. Moreover, Roma pupils may face discrimination from teachers and other pupils in schools: they are often seated in the last row, teachers do not read their homework and do not encourage them in their studies. Frequently they are also sent to special schools with consequences in future employment opportunities.

[15]In our sample only 36 women (out of 487) and 27 men (out of 427) in working age perform jobs with a full time contract, in the formal sector.

context and does not pretend to be indicative of the exact amounts.

[insert FIGURE 1 here]

Nonetheless, we argue that, since Roma people underestimate the outcomes of investing in education, a policy intervention would be successful in increasing their aspirations. Among Roma there are large differences in average earnings across different education levels. In our sample, for instance, average wages with secondary education are 27% higher than with primary education for boys and 21% for girls, and average wages with primary education are 29% higher than with uncompleted primary for boys and 21% for girls.[16] Moreover, the higher is the education level achieved, the better are the job market perspectives, both in terms of type of contract and place where to perform the job. Data from LSMS (2003) reported in Figure 2 suggest that these differences are substantial for Roma living in Serbia. The top panel presents their types of contract by education level. 90% of boys and 80% of girls with a secondary education level have got a written contract, while almost none works without a contract. Conversely, among those with only primary school almost 30% perform their activities without a contract. The bottom panel reports the places where jobs are performed. The percentage of those who work in the street or in flee market reduces drastically with the level of education. This is even more evident for girls. The pattern for office/factory as the place where to work is the opposite.

[insert FIGURE 2 here]

Therefore, conditional on the fact that Roma people's earnings are lower than Non-Roma's ones, there is still room for improvement based on education among Roma. The more one studies the higher are the wages and the better are the conditions to perform the job. It is crucial to understand whether parents are aware of the actual returns to schooling. We could not conduct a baseline survey before the introduction of the program. Thus, we need to look at data of parents not affected by the program. Figure 3 reports the distributions of expected returns to education for parents whose children attend the

---

[16]For Non-Roma the average wages with secondary education are 49% higher than with primary for boys and 60% for girls. For Non-Roma we use 2011 data for the city of Belgrade obtained from the Serbian Statistical Office.

schools that received the assistant in the second year of the RTA. Their averages are given by the solid lines. The dashed lines correspond to average wages of people in our sample by education. There are few women who completed primary and, especially, secondary school. Thus, results for females are less informative. Official data do not provide this information. The first panel reports the expected wage distributions, conditional on not having achieved any level of education. These distributions are more concentrated on the right of the dashed lines of actual average returns. Parents expect for their children higher returns when no level of education is achieved. The second and third panels of the figure report the expected wage distributions, conditional on having a primary and a secondary level of education, respectively. For male these distributions tend to be more concentrated on the left of the dashed line of actual average returns. Parents expect for their sons less than what people with that education levels actually earn. There is limited or imperfect information, and this likely fosters low aspirations for Roma people.

[insert FIGURE 3 here]

2.2 The Roma Teaching Assistant Program

The Roma Teaching Assistant Program is the main program in Central and Eastern Europe aimed at improving inclusion of Roma in education.[17] After the initial pilot phase, the program took off on a larger scale in the scholastic year 2009/2010.[18] In Fall 2009, 26 schools (*Early Enrollees*) entered the RTA program. In the following year an additional 77 schools (*Late Enrollees*) joined. Each school is assigned one teaching assistant. On average, the number of Roma per school is 86 (13% of total pupils enrolled) and those directly helped by the assistant are 27 - almost one third - especially among younger grades. Although schools are somewhat free in allocating the time of the assistant, her major tasks are helping children during regular classes and organising after-school extra

---

[17] For a more extensive description of the program see Battaglia and Lebedinski (2011).

[18] The Roma Teaching Assistant Program started as a pilot program implemented by various NGOs in 2002. 22 schools received an assistant at different points of time between 2002 and 2007. These schools are not the same schools that got the assistant starting from 2009 and are excluded from our analysis. In 2007 the OSCE took over the coordination and financing of the program. Since 2009 it started to have a country coverage and it is now under the coordination of the Ministry of Education.

classes. One day per week the assistant visits parents of children who are not going to school and informs other parents about their children's progress.

The RTA program is not a randomized experiment: schools and assistants had to apply in order to be part of the program. Schools were chosen based on the following two criteria: (1) a percentage of Roma pupils between 5% and 40%, and (2) preferably the availability of a preschool program in the school.[19] The requirements for assistants were the following: (1) secondary school attainment, (2) knowledge of Romani and (3) preferred experience in working with children.[20] It is not explicitly stated that the assistant needs to be Roma: only the knowledge of their language is required. However, all of them are of Roma origin. The selection criteria remain the same in both years and schools and assistants which applied in the first year and do not get selected could also apply in the second year. The only observable characteristic we are informed of does not differ between schools which applied before and schools which applied later: the percentage of Roma pupils is 13.6% in *Early Enrollees* and 13.7% in *Late Enrollees* schools.[21] Unfortunately we do not know what motivates some schools to apply before others and whether these motivations are related to differences in the principle or in school quality.[22] This might cause a selection bias problem. Our estimates can be overestimated: parents' aspirations can be correlated with the quality of the school. If children are going to better schools, parents may reasonably expect for them better educational achievement and better labor market perspectives, unconditional to the program.[23] However, some schools which

---

[19]Information on the existence of a preschool program are available only for the 78 schools applying in 2009. For the 252 schools applying in 2010 this information was not required anymore. In that year 50 assistants were assigned to kindergartens which offer themselves preschool programs. Schools which were not offering the preschool program 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. Since 2007 the attendance of at least 6 months of a cost free preschool program is compulsory; in 2010 its length has been extended to 9 months.

[20]In 2009 among 158 people applying, 26 were selected; in 2010, among 329 people, 77 got the job (and 50 more became assistants in kindergartens). All assistants live in the same municipality of the school they work for. Among those belonging to the same municipality, detailed criteria, based on level of education attained, motivation and experience in working with children, were used to rank assistants.

[21]This is the only information we have got, together with their size, for schools that applied and did not get selected. *Early Enrollees* schools count on average 792 pupils and *Late Enrollees* 894.

[22]In 2009/2010 the program was advertised in newspapers *Politika* and *Prosvetni Pregled*, the last being a newspaper for people working in the education sector; in 2010/2011 schools' directorates - one directorate may be responsible for more than a municipality - were in charge of sending applications directly to schools.

[23]Average marks of previous years in *Early Enrollees* schools do not suggest they are better schools.

applied in 2009 did not apply anymore the year after.[24] Thus, if they were really more motivated and of better quality than those applying later, it is hard to understand why they did not want to be part of the program anymore in 2010. Other schools that applied and met the requirements in both years got selected only in 2010. They should not differ from those selected in 2009. In our sample of Belgrade, among the 4 schools which got the assistant in 2010, one did also apply the year before. We believe that the selection mechanism does not bias substantially our results. Furthermore, schools selected in the first year are not different in observable characteristics from schools selected later. The same holds for the assistants. Table 1 reports the data in our sample.[25] We collected data from 9 schools in Belgrade: 5 schools received the assistant in 2009; 4 schools in 2010.[26]

[insert TABLE 1 here]

Early Enrollees and *Late Enrollees* schools count a similar number of Roma per class, 4.43 and 5.23, and a similar percentage of Roma per school, 20% and 22%. The sex composition among Roma is the same: in *Early Enrollees* schools 52% of students is female and in *Late Enrollees* schools 46%. 40% of Roma are born in Belgrade in *Early Enrollees* schools and 32% in *Late Enrollees* schools. Schools slightly differ only in class size: *Early Enrollees* schools have smaller classes, with 22.24 students versus 25.63 of *Late Enrollees* schools. The characteristics of the assistants in the two types of school are also comparables. Almost all of them are female with experience in NGO. In *Early Enrollees* schools 40% of the assistants got a university degree; in *Late Enrollees* schools 33%.[27]

Yet, we cannot exclude differences in the principle's motivation.

[24]In the whole country less than half of schools and assistants which applied in the first year and did not get selected applied again in 2010.

[25]The same holds for the whole sample of schools involved in the program in Serbia. In the RTA the schools selected in the first year are not different in observable characteristics from the schools selected later. The same holds for the assistants (Battaglia and Lebedinski, 2011).

[26]In Belgrade 6 schools got selected in the RTA program in the first year of its implementation. One school did not provide us the list of students so it is excluded from our sample. 8 schools were selected in 2010. We have got detailed administrative data from a subsample of 4. For the remaining 4 we only know the percentage of Roma per schools. They are comparable to those in our subsample.

[27]Among assistants in *Late Enrollees* schools there is one missing value for the information on the maximum level of education. This explains why the categories *secondary school* and *university* do not sum to 1.

3 Data and Descriptive Statistics

3.1 The Survey Design

We use first-hand collected data obtained through a survey conducted with 300 Roma households in 5 municipalities of Belgrade.[28] The survey took place in Fall 2010, one year after the implementation of the program in *Early Enrollees* and before *Late Enrollees* schools received the assistant. In 2010 schools received the assistant in November/December. The households in our sample have children who were enrolled in both types of schools. Pupils were randomly selected among students attending the schools.[29] Figure 4 displays a map of Belgrade with the 12 settlements were the survey was carried out. Settlements with the numbers 1 to 5 are the ones where assistants were introduced in 2009/2010, that is the settlements with children from *Early Enrollees* schools. Settlements 6 to 13 had assistants starting from 2010/2011.

[insert FIGURE 4 here]

We are interested in the effect of the RTA program on children from the lower four grades, given that the assistants mainly work with them. Our sample is constructed in such a way that all households have at least one child in the lower four grades of primary school in the scholastic year 2009/2010.

Three sets of questionnaires are administered in the survey: a household questionnaire providing information on the household and community characteristics, a questionnaire for the mother or caretaker and a questionnaire for the children. The *mother questionnaire* consists of an extensive series of questions about the education of the children aged 6 to 15 living in the household. In this section some questions on child labor are also asked. Children who attended first to fourth grade of primary school in 2009/2010 are asked information about their school and teachers. The *child questionnaire* also contains quick tests on children's abilities to read and write and to do some mathematics.

---

[28]The five municipalities are Voždovac, Zvezdara, Zemun, Palilula, and Čukarica.

[29]The response rate is 93.46%: 321 households have been contacted and 300 answered. Households were not compensated for their participation.

3.2 The sample

Our sample is divided in two groups. The first group consists of 122 households with children enrolled in 5 schools which got a Roma teaching assistant in 2009/2010. These households are randomly selected among households with at least one child in a treated school and correspond to the treated group.[30] The 178 remaining households were randomly selected from settlements in Belgrade close to 8 schools which received the RTA program in 2010/2011 and they are our control group.[31] The number of households selected from each settlement is proportional to the size of settlement.

We consider the whole household to be treated if at least one child goes to a school with an assistant in the first year of the implementation of the program. We do expect that parents' aspirations are created at the household level: once a child is exposed to the program, expectations on future opportunities change for all children of the same household. Table 2 reports the characteristics of treated and control groups.

[insert TABLE 2 here]

They are comparable in terms of observable characteristics. Their differences in means are not statistically significant in most of the cases. Wealth, monthly income, educational attainments and household composition do not differ between groups.[32] Households are equally located in rural and urban areas.[33] 32% of households in the control group and 31% in the treated group have at least one member working in the informal sector. However, treated households are more in only Roma neighborhoods (28% versus 16%)[34] and among non-treated households there are significantly more Muslim (80% versus 57%). It would be worthy to investigate whether the program impacts differently depending on

---

[30]Pupils selected were enrolled in the schools in year 2009/2010. Students who dropout are therefore included in the sample.

[31]In Belgrade 6 schools received the assistant in 2009/2010 and 9 schools in 2010/2011. In our sample we have 5 schools out of 6 which received the assistant in 2009/2010 and 8 out of 9 among those which received her in 2010/2011. We were not allowed to collect data in the schools excluded. These schools are not different in characteristics from those belonging to the sample.

[32]Rank among siblings is significantly higher among treated households. Nonetheless, we do not believe this would be problematic given that household composition does not differ between the two groups.

[33]We define urban area a municipality with more than 35,000 inhabitants.

[34]We asked households whether in their community/neighborhood (200 meters around their house) there were only Roma or both Roma and Non-Roma. Therefore, neighborhoods do not correspond exactly to the 12 settlements where the survey has been conducted.

the religion and the type of neighborhood (see section 4.3 on heterogeneous effects).[35]

Identification requires the absence of selective sorting into treatment. Field analysis suggests that parents were not aware of the existence of the program before enrolling 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 program.[36] Therefore, we are confident that our analysis is not affected by possible selection of children into schools.

We assume that everyone in a school with the assistant is aware of her presence. Roma people usually live in communities where they know each other and regularly interact. They are even in contact with people living elsewhere belonging to the same community. Family and community ties are strong. It is unlikely that households do not know that there is an assistant in the school their children are enrolled in. However, there is only one assistant per school and not every Roma child in the treated school is helped by her. One can argue that only parents of children directly interacting with the assistant are aware of her presence in the school. Therefore, the role model mechanism can only work for them. We can explore another definition of being treated beside the main one. Our second definition considers a household to be treated only if at least one child is in an *Early Enrollees* school and there is evidence from the survey that the assistant has worked with her. The school does not keep track of the names of the children with whom she interacts. Thus, we obtain the information from the parents. In this case we are sure that they are aware of the presence of the assistant. A household in this case is treated if either (1) parents state that there is someone in the school who helps the kid with her homework or she is following additional classes at school, or (2) there is someone from the school who ever come to her place or call her because of the kid.[37]

Table 4 reports the characteristics of households with treated and not treated children

---

[35]Overall, the characteristics of our sample are in line with official data. Only, somewhat surprising with respect to them, few households have both parents with unfinished primary school (7%) and a relatively large share of households has at least a parent with finished secondary school (19%).

[36]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 2009/2010 these percentages were respectively 1.6% and 1.3%. The number of Roma pupils enrolling at school for the first time reduced between the two years and it did it proportionally in both types of schools.

[37]We decided not to ask explicitly the parents whether the school of their children is in the RTA program, because it was not clear to us whether the parents are aware of the name of the program and how they perceive the teaching assistant, e.g. as assistant, teacher, etc..

in *Early Enrollees* schools.

[insert TABLE 4 here]

The differences in means between those helped by the assistant (*treated*) and those not helped (*untreated*) are not statistically significant in most of the cases. Certainly, the assistant works mainly with younger children, as also suggested by the guidelines of the program. We know that she chooses the pupils to work with: treated children are not randomly selected. Thus, since the group that receives the treatment is a selected subset of those in *Early Enrollees* schools, a simple comparison of pupils helped and not helped by the assistant would incorrectly estimate the gains or loss from the RTA program (those not helped include both untreated in *Early Enrollees* schools and children in *Late Enrollees* schools). In order to separate the effect of the program from the impacts of the selection mechanism, we use an instrumental variable strategy. In this specification, by assumption being in a *Early Enrollees* - treated - school is only capturing the intention-to-treat (ITT) effect. We therefore use assigned treatment as an instrumental variable for treatment received: being in a *Early Enrollees* school is the instrument for being helped by the assistant. The local average treatment effect (LATE) - obtained by considering only those who are helped by the assistant in *Early Enrollees* schools - is here the effect of treatment on the treated. There are no *always-takers* in this case (those helped by the assistant are only in *Early Enrollees* schools): the treated population consists entirely of compliers.

3.3   Outcome variables

We use three different sets of questions to understand whether the program is effective in changing parents' aspirations about their children's future opportunities. We focus on the expectations of parents because we believe that at such a young age (6 to 15) the aspirations of parents are more relevant for a child's educational attainment and more reliable for expected returns to education than child's aspirations.[38]

---

[38]However we ask pupils which is the highest expected level of education they expect to achieve. Two-thirds of parents answer as their children, whereas one-third expect their kids to achieve a lower

The first and second set of questions relates to expected returns to education. They are asked to either mother or father (or caretaker), however the mother is the main interviewee in 92% of cases. They are asked for the oldest boy and the oldest girl in the household.[39]

The first set of questions considers expectations about the likelihood of getting a job given a certain level of education achieved. We ask: "Assume that your oldest boy has finished primary (or secondary) school - and that is his highest degree - and he is 25-30 years old: how certain are you that he will have any kind of job?". We ask the same question for the oldest girl. Although we are mainly interested in the probability of finding a job given a secondary education level, we use for comparison reasons also the probability of finding a job given a primary school education level.[40] The responses to this question come from a five point Likert scale and they are "Absolutely sure", "Quite sure", "Maybe", "Unlikely" and "No, s/he will not find a job". The Likert scale has a disadvantage: different respondents can interpret the scale differently so that other factors such as optimism or education affect the response. Alternatives such as explaining probabilities to interviewees and asking them to express their expectations using a cardinal scale are suggested by the literature (Delavande et al., 2009). However, due to the low educational level of our respondents, this drawback could not be overcome and we decided to offer them the possibility to choose among five different responses.[41] For the purpose of our analysis, we converted the five Likert scale outcomes into a dummy variable. If the respondent declared that it is unlikely or that her child will not find a job

education level than children expect. Only in few cases parents expect more than their children and this happens mainly when children perform well at school. Discrepancy in answers between parents and children is mainly found among poorer and larger families, living in *only Roma* neighborhoods, Muslim and with lower levels of education. Pupils also perform worse at school than their classmates. Results are not reported, but they are available upon request.

[39]In the pilot survey we asked the questions for each child but we realised that there was no variation in the responses between the children of the same sex. As a consequence we decided to pose this question only for the oldest male and for the oldest female child. In only 6% of cases the oldest child may be older than 15 and thus not enrolled in a compulsory school. However, we believe that this is not a problem in our setting given the hypothetical nature of the question and the fact that the treatment is considered at the household level. In the worst case this would imply that our results are underestimating the real effect.

[40]The same question is asked for the case in which the child does not finish primary school. Results are not reported and available upon request.

[41]Between 1% (male) to 5% (female) of households did not answer these questions.

given a certain education level, we set the probability to zero. In the other three cases −
"Absolutely sure", "Quite sure", "Maybe" −, we set the probability to one.[42] We believe
that by aggregating the categories to a dummy we do not lose important information:
almost two-thirds of respondents answered "Unlikely" and "No, s/he will not find a job"
in the case of primary school and one third in the case of secondary.[43]

[insert FIGURE 5 here]

The second set of questions elicits minimum and maximum amounts parents expect
that their children will earn once employed. We ask: "Assume that your oldest boy has
finished primary school (or secondary) and this is his highest degree and he is 25-30 years
old. Think about the kinds of jobs he might be doing in this case. What do you think is the
minimum amount he can earn per month? And the maximum amount?" The interviewees
have been asked explicitly to take into account both regular and irregular types of income.
The same questions have been asked for girls. We obtained the minimum and maximum
earnings and we use their (log) average as our measure of expected earnings.[44]

The third relevant outcome is the highest expected education level of the child. The
exact question is: "What level of formal education do you think that *(name)* will com-
plete?" We create the dummy "secondary as the highest level of education" that takes
the value 1 when it was answered "secondary (or more)" and 0 otherwise. The question
is asked for each child between 6 and 15 years old.[45] For consistency we also estimate the
impact with a reduced sample, corresponding only to the oldest boy and the oldest girl
in the household.[46] Summary statistics for the outcome variables in our sample are re-

---

[42]The results hold in case we set the probability to zero when the respondent declared that "No, s/he
will not find a job", "Unlikely", and "Maybe", while we set the probability to one for "Absolutely sure"
and "Quite sure".

[43]Ordered logit analysis suggests that some categories may not be collapsed (see Table A.1 in Ap-
pendix). For instance, while for secondary school "Absolutely sure" and "Quite sure" can be clearly
collapsed, they should not be in the case of primary school. For this reason, we also keep the variable as
categorical. Estimates with categorical outcomes are available upon request. They confirm the results
reported.

[44]3% of households did not answer in the case of questions referring to a male child, while these
percentage is around 15% for a female child. Estimates with minimum and maximum earnings are
available upon request.

[45]The median number of children aged 6 to 15 per household is 2. There are many missing values for
this outcome of interest. This explains why our sample is as big as with the other outcomes.

[46]Results are not reported, but they are available upon request.

ported in Table 3 and suggest a possible positive impact of the program on both expected salary and level of education achieved. In treated households expected future earnings are higher than in control households for both primary and secondary level of education. Respondents in treated households are also more likely to expect their children to finish secondary school. This difference is only significant for boys.

[insert TABLE 3 here]

4 Estimation Strategy and Results

In the Roma Teaching Assistant Program all the assistants are Roma and from the same social background of the pupils they help. They can act as role models for the kids they work with. The presence of a person sharing her successful story should affect children's and their parents' aspirations about their future in two ways. First, we expect that treated parents think that also their children can succeed. They would be more likely to find a job and a better job with higher salaries, conditional on achieving a specific education level. We do expect larger impacts for those with higher education level. The better jobs are obtained with a higher education level. It is worth investing in education. Second, as a consequence they would more likely expect their children to finish secondary school.

4.1 Early Enrollees versus Late Enrollees

We estimated the impacts of the RTA program on returns to education with the following specification:

$$Y_j = \alpha_0 + \alpha_1 treatment_j + \alpha_2 X_j' + \varepsilon_j \tag{1}$$

where Y_j corresponds to the outcomes of interest for the household j: likelihood of finding a job with primary school as the highest degree achieved, likelihood of finding a job with secondary school as the highest degree achieved, (log) mean amount of earnings per month with primary education and (log) mean amount of earnings per month

with secondary education. $treatment_j$ equals 1 whether there is at least one child in the household who goes to a *Early Enrollees* school. X'_j includes household wealth per capita, whether there is someone in the family who works in the informal sector, whether the household lives in a urban area and in a Roma neighborhood, whether the household is Muslim, the maximum education level of parents and household composition characteristics. For the outcome "secondary school as the highest expected level of education", we have got information for each child between 6 and 15 years old. We introduce a second specification where the dependent variable is at the child level:

$$Y_{ij} = \beta_0 + \beta_1 treatment_j + \beta_2 X'_{ij} + \nu_{ij} \tag{2}$$

$treatment_j$ is defined as above. X'_{ij} also includes age and age squared of the child, her gender, rank among siblings, demeaned mark in Mathematics and Serbian of the previous scholastic year.[47] Standard errors are clustered at the cohort times school level. Regressions are estimated separately for boys and girls because we are interested in the effects for each gender.[48] We do also report results with the pooled sample in Table A.2 in Appendix.

Results for the probability of finding a job, expected earnings and highest expected education level are reported in Table 5. For consistency, all the estimates are OLS.[49]

Columns (1) to (4) show estimates for boys, while columns (5) to (8) for girls. The coefficients for the expected probabilities of finding a job with primary and secondary school as the highest level of education are reported in the top panel of Table 5. They document that the direction of the impact is robust to excluding controls, but the inclusion of controls improves the precision of the estimates. Results are not statistically significant

---

[47]The marks are demeaned from the average school marks (among Roma). For children in their first grade, the average school marks are used.

[48]It is worth investigating whether the gender of the assistant may affect differently boys and girls for our outcomes of interest. Results do not suggest that aspirations change depending on the sex of the assistant. This seems to matter only for the likelihood of getting a secondary education level for girls. Their parents expect them to achieve an higher level of education when the assistant is female than when he is male. However, the caveat here is that among assistants only one is male. Results are not reported, but they are available upon request.

[49]Probit estimates for the two dummy outcomes confirm the results. They are not reported but are available upon request.

in all specifications. However, they suggest a possible positive trend in expectations. For boys the coefficients turn positive with secondary education; for girls they are larger in absolute terms for primary school as highest degree achieved than for secondary school. Given the low statistical significance of the results, however we argue that job market perspective of those exposed to the program remain substantially unchanged. Having at least one child in a treated school does not change parents' expectations about their children's future opportunities to find a job compared to having no children in a treated school.[50]

[insert TABLE 5 here]

The middle part of the table shows the results for the expected (log) mean earnings per month. Parents in treated households expect higher future salaries for both boys and girls. Conditional on having achieved a secondary education level, being in a treated household increases the expected monthly earnings by almost 9.6% for boys and 10.5% for girls, on average.[51] This increase corresponds to almost 26 Euro with respect to an average expected earning in households not exposed to the program of 271 Euro for boys and 255 Euro for girls. Thus, although treated parents do not expect higher employment perspectives for their children, they do expect higher salaries once they obtain a job. This suggests that they likely expect them to get better jobs. For boys this is the case also conditional on having achieved a primary education level: being in a treated household increases the expected monthly earnings by almost 11%, on average. The regression results for secondary education as the highest expected level of education are reported in the bottom part of the table. We find that parents in treated households are more likely to expect their children to finish secondary school. The impact is statistically significant only for boys. On average, parents of pupils in *Early Enrollees* schools are 12.3 percentage points more likely to expect their sons to finish secondary school.[52]

---

[50]The coefficients of controls are not reported, but they are available upon request.

[51]The regression coefficients can be interpreted as semi-elasticities. 0.092 corresponds to $100 * (e^{0.092} - 1)$; 0.100 corresponds to $100 * (e^{0.100} - 1)$. We estimate the effects also with minimum and maximum earnings. Results are similar. They are not reported, but they are available upon request.

[52]If we consider only the oldest boy and the oldest girl in the household we obtain similar results. The magnitude of the coefficients is even higher. The coefficient of treatment for boys is statistically significant at 10%.

4.1.1 Directly helped versus not directly helped by the assistant in *Early Enrollees* schools

There is only one assistant per school and not every Roma child in the treated school is helped by her. One can argue that only parents of children directly interacting with the assistant can be aware of her presence in the school. They are the only ones for which the role model mechanism is expected to work. Our second definition considers a household to be treated only if at least one child is in an *Early Enrollees* school and there is evidence from the survey that the assistant has worked with her.

The assistant chooses the pupils she works with: treated children are not randomly selected. Thus, since the group that receives the treatment is a selected subset of those in *Early Enrollees* schools, a simple comparison between those actually helped and the control group (those not helped in *Early Enrollees* schools and children in *Late Enrollees* schools) is misleading. We use an instrumental variable strategy. By assumption here being in a *Early Enrollees* school is only capturing the intention-to-treat (ITT) effect. Assigned treatment is therefore used as an instrumental variable for treatment received: being in a *Early Enrollees* school is the instrument for being helped by the assistant.

We use the following specification:

$$Y_{(i)j} = \theta_0 + \theta_1 assistant_j + \theta_2 X'_{(i)j} + \varepsilon_{(i)j} \tag{3}$$

where $Y_{(i)j}$ corresponds to the outcomes of interest: likelihood of finding a job with primary school as the highest degree achieved, likelihood of finding a job with secondary school as the highest degree achieved, (log) mean amount of earnings per month with primary education and (log) mean amount of earnings per month with secondary education and secondary school as the highest expected level of education. $assistant_j$ is equal to 1 when there is at least one child in the household who is helped by the assistant and 0 otherwise. Given the problem of selection bias, we know that the error term $\varepsilon_{(i)j}$ is here

composed of two parts:

$$\epsilon_{(i)j} = \eta_{(i)j} + u_{(i)j} \tag{4}$$

where $\eta_{(i)j}$ is an unobservable individual term and $u_{(i)j}$ is a random term. $assistant_j$ depends on some factors captured by $\eta_{(i)j}$. We therefore model $assistant_j$ in a reduced form framework as follows:

$$assistant_j = \gamma_0 + \gamma_1 treatment_j + \gamma_2 X'_{(i)j} + \eta_{(i)j} + v_{(i)j} \tag{5}$$

where $treatment_j$ is equal to 1 if there is at least one child in the household enrolled in a *Early Enrollees* school. Being enrolled in a *Early Enrollees* school is correlated with the fact of being helped by the assistant but uncorrelated with any unobservable attributes that affect the outcomes of interest. The instrument is as good as randomly assigned. It also satisfies the exclusion restriction by assumption: only parents of children directly interacting with the assistant are aware of her presence in the school. The instrument operates only through the fact of being helped by the assistant ($Y_i(d,0) = Y_i(d,1)$ for $d = 0, 1$). The monotonicity assumption needed for heterogeneous IV models holds: while the instrument may have no effect on some people, all those who are affected, are affected in the same way. θ_1 in (3) captures the local average treatment effect (LATE) that in this case is the effect of treatment on the treated. There are no *always-takers* in this case (those helped by the assistant are only in *Early Enrollees* schools): the treated population consists entirely of compliers.

Results for the probability of finding a job, expected earnings and highest expected education level are reported in Table 6.[53] Columns (1) to (2) show estimates for boys,

---

[53]The use of IV to solve selection bias problems is illustrated in Table A.3 in Appendix. Columns (1) and (2) report OLS results. These estimates are misleading because they compare pupils according to the actual treatment received: those helped by the assistant versus those not helped in the same *Early Enrollees* schools and children in control schools. Columns 3 and 4 compare pupils according to whether they are potentially treated: being in a *Early Enrollees* or *Late Enrollees* school. This is the intention-to-treat (ITT) effect. Since *treatment* was as good as randomly assigned, ITT tells us the causal effect of being in a *Early Enrollees* school. It builds in the fact that some pupils in treated schools are not treated. For this reason, it is smaller than the average casual effect on those actually treated. It clearly corresponds to our main specification where we assume instead that everyone in a

while columns (3) to (4) for girls.

[insert TABLE 6 here]

The coefficients for the expected probabilities of finding a job with primary and secondary school as the highest level of education are reported in the top panel. They are similar to those obtained in the main specification. As before, results are not statistically significant in all specifications but they suggest a possible positive trend in expectations. The second part of the table shows the results for the expected (log) mean earnings per month. Parents in treated households expect higher future salaries for both boys and girls, as in the main specification. The impacts are reasonably higher: in this case we are sure that parents know the assistant. Conditional on having achieved a secondary education level, being in a treated household increases the expected monthly earnings by almost 21% for both boys and girls, on average. This increase corresponds to roughly 55 Euro. As before, for boys this is the case also conditional on having achieved a primary education level: being in a treated household increases the expected monthly earnings by almost 33%, on average, corresponding to 89 Euro. The regression results for secondary education as the highest expected level of education are reported in the third part of the table. We find that parents whose children are helped by the assistant are more likely to expect their children to finish secondary school. As in the main specification, the impact is statistically significant only for boys. On average, parents of pupils in *Early Enrollees* schools are 26 percentage points more likely to expect their sons to finish secondary school. The bottom part reports the results for the first-stage. The coefficients of being in a *Early Enrolless* school are positive, as expected, and highly statistically significant. The first-stage results for the instrumental variable estimation show that F-statistics on the incidence of treatment are clearly above 10.

Overall, if we assume that only parents of children helped by the assistant are aware of her presence, we obtain similar results as before. Not surprisingly, these parents revise

---

Early Enrollees school is treated. Columns 5 and 6 measure the effect of treatment on the treated. They do not consistently differ from OLS estimates because treatment and control groups are not so different in observable characteristics. We know that there is a problem of selection bias because the assistant chooses the pupils to work with. Still, her choice seems to be close to a random choice. The selection bias in this case is negative: those who are helped by the assistant tend to be the worst students.

more their expectations: in this case we are sure that they know the assistant and are exposed to the role model mechanism. Nonetheless, the comparison between the main results and the ones obtained here suggests that also parents of pupils in treated schoool who are not helped by the assistant revise their expectations.

4.2   Remedial education program and role model

We cannot exclude that the effect of the program also passes through the remedial education channel, especially for those we know are helped by the assistant. Parents expect their children to go more to school not because of the role model but because they perform better now thanks to the assistant. In our survey we run quick test scores in Mathematics and Serbian. We define *Maths score* equal to 1 when the kid is able to correctly answer both questions − "Please tell me how much is 5+4?"− and − "Please tell me how much is 23+12?"−, and 0 otherwise. We define *Serbian score* equal to 1 when the kid is able to read and write. A child is able to read when she knows to correctly read the sentence written on a card − "Could you please read me the letters, the word and the sentence on this card?"; *Able to read* takes value 0 when she does not know letters, recognises only letters or knows to read the words but cannot read the complete sentence. A child is able to write when she knows to correctly write a proposed sentence − "Please write the following sentence"; *Able to write* takes value 0 when she does not know to write at all or she writes the sentence with mistakes. Results for the LATE are reported in Table 7.

[insert TABLE 7 here]

Pupils who are helped by the assistant perform better than their classmates and pupils in *Late Enrollees* schools in both test scores, although impacts are statistically significant only for Serbian. On average, boys get 0.56 of a standard deviation more; girls get 0.66 of a standard deviation more.[54]

---

[54]We also studied the impact of the program on schooling outcomes in its first year of implementation in a companion study Battaglia and Lebedinski (2011). We find that the program had a positive effect. There is evidence that children exposed to RTA went more to school and that, on average, marks have improved and dropouts have reduced for children in their first grade. Higher and more systematic impacts are obtained in schools with a lower number of Roma, especially if female.

Nonetheless, if among the worst performers we find that aspirations have increased for those helped by the assistant, we have evidence that the effect of the program does not passes through the remedial education channel only. They have been helped by the assistant but they are not doing better at school. Still, their parents believe that their returns to education would be higher and they would achieve a secondary level of education. Therefore, we select only those who wrongly answer the Serbian score and the Maths score. Results are reported in Table 8.[55]

[insert TABLE 8 here]

Previous results are confirmed, although we need to be cautious in the interpretation of the results given the small sample. They suggest a positive trend in expectations for the probability of finding a job, although the probability of finding a job with only primary education is now statistically significant. Conditional on having achieved a certain education level, being helped by the assistant does increase the expected monthly earnings, although the results are not significantly different from 0. Parents revise their expectations about the highest level of education achievable even more than in the previous specification. On average, parents of pupils directly interacting with the assistant are 50.5 percentage points more likely to expect their male children to finish secondary school. Even if their children - who are helped by the assistant - are not performing well at school, parents still believe they will go more to school and have higher returns to education, once achieved a secondary level of education. There is evidence that the program changes parents' expectations not only through the remedial education channel.

4.3 Number of Roma per class

We believe that the number of Roma per class is important to understand our results. A priori it is difficult to predict whether the high concentration of Roma in the same class leads parents to revise their expectations in one direction or another. There are two possible behaviors that arise. On one side, parents of children in classes with a high

---

[55]For robustness check, we estimate the effects also selecting those who wrongly answer the Serbian score and those who wrongly answer the Maths score separately. Results do not change.

number of Roma may perceive to be more segregated and excluded from the society. The impact of the role model would be mitigated: the initial expectations are so low that the presence of a person from the same community is not enough to rise them. On the other side, the more the Roma per class, the more parents may feel to be an important part of the society and expect to have better future opportunities.

In our survey pupils were randomly selected among students in the schools so we have not got precise information on the number of Roma per class. However, we also collected first-hand administrative data of each school involved in the program (Battaglia and Lebedinski, 2011). By merging these data with data obtained through the survey, we are able to control for the number of Roma per class when investigating the impact of the program on parents' aspirations.[56] Results are reported in Tables 9.

[insert TABLE 9 here]

The number (or percentage[57]) of Roma kids per class matters in the creation of aspirations. The higher is their number in treated schools, the lower the expectations to find a job, get a higher salary and achieve a secondary level of education. Few results are statistically significant. Our sample reduces to few observations. However, the trends are suggestive of a negative impact of the number of Roma per class on the outcomes of interest. The inclusion of this information does not change our main results. Overall, the total effects remain positive.[58] Still, the effect of the program is lower when there is a high number of Roma in the same class.

4.4   Heterogeneous effects

In this section we examine heterogeneous effects of the program on Muslim versus Non-Muslim households, households in *only Roma* neighborhoods versus households in less concentrated neighborhoods, and younger (6 to 10 years) versus older children (11 to 15).

---

[56]We get a smaller sample because we collected administrative data on 4 schools out of 8 among the controls.

[57]We also use percentage of Roma per class. Results confirm those obtained here. They are not reported but they are available upon request.

[58]The only exception is secondary school as the highest level of education expected for girls. Female are not affected by the number of Roma per class and - if they were - they would rather be positively affected by it.

4.4.1 Muslim households versus Non-Muslim households

Our main specification is suggestive of the fact that there could be a differential effect of the program on Muslim households. Parents from Muslim households expect higher earnings conditional on finishing secondary school and they expect their children to attain a lower level of education for both genders when compared to Non-Muslim households.[59] Moreover, descriptive statistics show that treated and control groups differ in the number of Muslim households: there are significantly more Muslim families among households with children enrolled in *Late Enrollees* schools. We therefore think it would be worthy to investigate whether the program affects differently Muslims.

We proceed therefore with the following specification (6) which includes the interaction of being in a Muslim household and in a treated household:

$$Y_{(i)j} = \delta_0 + \delta_1 treatment_j + \delta_2 muslim_j + \delta_3 treatment_j * muslim_j + \delta_4 X'_{(i)j} + \epsilon_j \ (6)$$

$Y_{(i)j}$ are the usual outcomes of interest.[60] The coefficient δ_1 captures the effect of treatment on Non-Muslims. The coefficient δ_2 captures the difference between Muslims and Non-Muslims among the *Late Enrollees*, and δ_3 is the differential impact of interest. Our results are reported in columns (1) and (3) of Table 10.

[insert TABLE 10 here]

Overall, estimates suggest that the program does not impact differently Muslim and Non-Muslim in terms of job market perspectives and expected salaries. Nonetheless, Non-Muslim Roma react more in terms of expectations on educational achievement to the presence of a teaching assistant than Muslim Roma. Non-Muslims who are in *Early Enrollees* schools are on average 21.4 percentage points more likely to expect their male children to finish secondary education compared to Non-Muslims in control schools. We do not know the religion of the assistants in order to investigate further.

---

[59]Results are not reported but they are available upon request.
[60]$X_{(i)j}$ does not include now whether the household is Muslim.

4.4.2 Households in only Roma neighborhoods versus households in less concentrated neighborhoods

Descriptive statistics show that treated households are more in *only Roma* neighborhoods. It is worthy to investigate whether the program affects differently those in *only Roma* neighborhoods and those in *mixed (Roma and Non-Roma)* neighborhoods. We therefore proceed with specification (7) which includes the interaction of being in a *only Roma* neighborhoods and being treated:[61]

$$Y_{(i)j} = \phi_0 + \phi_1 treatment_j + \phi_2 onlyroma_j +$$
$$+ \phi_3 treatment_j * onlyroma_j + \phi_4 X'_{(i)j} + \eta_j \tag{7}$$

Our results are reported in columns (2) and (4) of Table 10. Overall, estimates suggest that the program does not impact differently households in *only Roma* neighborhoods and households in *mixed* neighborhoods in terms of job market perspectives and expected salaries. The coefficients are not statistically significant in all the cases. Nonetheless, parents of children living in *mixed* neighborhoods who are in *Early Enrollees* schools are on average 13.6 percentage points more likely to expect their male children to finish secondary education compared to parents of children living in *mixed* neighborhoods in control schools. Our overall results on expected level of education are mainly driven by responses from parents living in *mixed* neighborhoods.

4.4.3 Young versus old kids

Younger children aged 6 to 10 may respond differently to the program from older children aged 11 to 15. There are two reasons to expect this to be the case. First, assistants were explicitly asked to work more with younger children. Second, we know that the gap in knowledge between Roma and Non-Roma children is present already when children enrol in primary school and that it increases over time. Under such circumstances, it might be easier to influence expectations of parents for younger children than for the older ones.

---

[61] $X_{(i)j}$ does not include now whether the household is in a *only Roma* neighborhood.

We estimate the following regression by gender. We have individual outcomes only for the expected education level and we estimate only this outcome. *young* is equal 1 if the kid is aged 6 to 10.

$$Y_{ij} = \vartheta_0 + \vartheta_1 treatment_j + \vartheta_2 young_{ij} + \vartheta_3 treatment_j * young_{ij} + \vartheta_4 X'_{ij} + \tau_{ij} \quad (8)$$

The results are shown in columns (5) and (6) of Table 10. Our coefficients are not statistically significant when we compare boys in *Early Enrollees* and *Late Enrollees* schools, although the magnitude and direction are still suggestive of the effect. We find that there is little difference between younger and older boys. The program affects the probability to finish secondary school similarly for both groups, although the effect is slightly higher for younger kids. We find a different effect for girls instead. Young girls in *Early Enrollees* schools are on average 27.9 percentage points more likely to be expected to finish secondary school than older female schoolmates.

5   Conclusion

There are reasons to believe that the lack of goals and aspirations is an important factor influencing the educational decision of Roma people. They often perceive low benefits of going to school compared to the respective costs and underinvest in education. Nonetheless, although there is evidence that the mean earnings of Roma workers are lower than those of Non-Roma workers, among Roma the difference between average earnings from one education level and another is high. The problem is that they are not aware of the situation. The provision of a role model can reduce the information gap between perceived returns to schooling and actual returns, which likely fosters low aspirations for Roma people. The Roma Teaching Assistant Program offers a perfect example. All its assistants are Roma and from the same social background of the pupils they help.

We exploit the gradual implementation of the RTA program to identify its impact on aspirations. Our data, collected one year after the first implementation, suggest that parents of children exposed to the program expect higher returns to education for their

29

kids. They are not more likely to expect them to find a job, but once employed they are expected to get higher salaries. This suggests that they might expect for them better jobs. On average, being in a treated household increases the expected monthly earnings by almost 9.6% for boys and 10.5% for girls. Moreover, on average, parents of pupils in *Early Enrollees* schools are 12.3 percentage points more likely to expect their sons to finish secondary education than parents of pupils in *Late Enrollees* schools. There is only one assistant per school and not every Roma child in the treated school is helped by her. One can argue that only parents of children directly interacting with the assistant can be aware of her presence in the school and that the role model mechanism works only for them. Our second definition considers a household to be treated only if at least one child is in an *Early Enrollees* school and there is evidence from the survey that the assistant has worked with her. We obtain results similar to the main specification. The impacts are reasonably higher: in this case we are sure that parents know their children are helped by the assistant and exposed to the role model mechanism. Nonetheless, the comparison between the main results and those obtained in the second specification suggests that also parents of pupils in treated school who are not helped by the assistant revise their expectations. We cannot exclude that the effect of the program passes also through the remedial education channel, especially for pupils we know are helped by the assistant. Parents expect their children to go more to school because they perform better now thanks to the assistant. From our survey we know that those treated do better in test scores. However, if we select those performing badly, our results hold. Furthermore, an examination of heterogeneous effects suggests first that our results on highest expected level of education are driven by responses from Non-Muslim parents and parents of those living in *mixed* neighborhoods. Second, parents revise their expectations in response to the program mainly for younger kids (6 to 10 years). Especially younger girls are more affected by the presence of an assistant: they are on average 27.9 percentage points more likely to be expected to finish secondary school than older female schoolmates.

Overall, these results suggest that the presence of a person of the same social background who showed to be successful motivates parents to believe their children can suc-

ceed. Interventions to raise perceived returns may thus be effective in increasing current investment in education. Remedial education programs featured this way can help create role models by opening opportunities that were previously unexpected to a group. This study suggests replicable examples in contexts where minorities suffer low attainment rates and social exclusion. It shows the importance of a role model mechanism that works, especially if we consider that we are in short-time horizon. One year of a remedial education program may not be enough to break the curse of low aspirations, but encouraging results are found in this direction. Investigating the effects of such programs in the long-run is a central question for future research.

References

Akerlof, G. A. and R. E. Kranton (2000). Economics and Identity. *The Quarterly Journal of Economics 115 (3)*, 715–753.

Alesina, A., R. Baqir, and W. Easterly (1999). Public Goods and Ethnic Divisions. *The Quarterly Journal of Economics 114*, 1243–1284.

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.

Battaglia, M. and L. Lebedinski (2011). Equal Access to Education: An Evaluation of the Roma Teaching Assistant Programme in Serbia.

Bayer, P., S. L. Ross, and G. Topa (2008). Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy 116*, 1150–1196.

Beaman, L., E. Duflo, R. Pande, and P. Topolova (2012). Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India. Sciencexpress.

Blume, L. and S. Durlauf (2006). *Identifying Social Interactions: A Review*. Michael Oakes and Jay Kaufman (eds.), Methods in Social Epidomology, Wiley.

Boeri, T., M. D. Philippis, E. Patacchini, and M. Pellizzari (2011). Moving to Segregation: Evidence from 8 Italian citites. Working Papers 390, IGIER Bocconi University.

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

Card, D. and J. Rothstein (2007). Racial Segregation and the Black-White Test Score Gap. *Journal of Public Economics 91* (11–12), 2158–2184.

Clark, K. and S. Drinkwater (2002). Enclaves, Neighbourhood Effects and Employment Outcomes: Ethnic Minorities in England and Wales. *Journal of Population Economics 7*(5), 603–628.

Cutler, D. M. and E. L. Glaeser (1997). Are Ghettos Good or Bad? *The Quarterly Journal of Economics, 112*, 827–872.

Dee, T. S. (2004). Teachers, Race, and Student Achievement in a Randomized Experiment. *The Review of Economics and Statistics 86*(1), 195–210.

Delavande, A., X. Ginè, and D. McKenzie (2009). Measuring Subjective Expectations in Developing Countries.

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

Durlauf, S. (2004). Chapter 50 Neighborhood Effects. In J. V. Henderson and J.-F. Thisse (Eds.), *Cities and Geography*, Volume 4 of *Handbook of Regional and Urban Economics*, pp. 2173 – 2242. Elsevier.

Edin, P.-A., P. Frederiksson, and O. Åslund (2003). Ethnic Enclaves and the Economic Success of Immigrants: Evidence from a Natural Experiment. *The Quarterly Journal of Economics 118*(1), 329–357.

Genicot, G. and D. Ray (2010). Aspirations and Inequality.

Glewwe, P. M. K. (2006). *Handbook of Economics of Education, Chapter 16.* North Holland.

Graham, P. A. (1987). Black Teachers: A Drastically Scarce Resource. *Phi Delta Kappan 68*(8), 598–605.

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.

Jensen, R. (2010). The (perceived) Returns to Education and the Demand for Schooling. *The Quarterly Journal of Economics 125 (2)*, 515–548.

Kaufmann, K. M. and O. Attanasio (2009, July). Educational Choices, Subjective Expectations, and Credit Constraints. Technical report, NBER Working Paper No. w15087.

Krishnan, P. and S. Krutikova (2010, February). Skill Formation in Bombay's Slums. Technical report, NBER Working Paper No. w15087.

La Ferrara, E. and A. Mele (2006, July). Racial Segregation and Public School Expenditure. Technical report, CEPR Discussion Paper No. 5750.

Ladson-Billings, G. (1994). *The Dreamkeepers: Successful Teachers of African American Children.* Jossy Bass Publishers.

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.

Manski, C. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies 60 (3)*, 531–542.

Nguyen, T. (2008). *Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar.* Ph. D. thesis, Massachusetts Institute of Technology.

Open Society Institute (2007).

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

Ray, D. (2004). Aspirations, Poverty and Economic Change. Technical report, Bureau for Research in Economic Analysis of Development (BREAD) Policy Papers No.442.

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.

Wilson, W. J. (1987). *The Truly Disadvantaged.* University of Chicago Press.

A Tables

Table 1: Characteristics of the schools and assistants (Belgrade)

	All	Early Enrollees	Late Enrollees	Difference (1-2)	P-value
	(1)	(2)	(3)	(4)	(5)
Characteristics of the schools					
Class size	24	22.24	25.63	-3.39	[0.081]
No. Roma per class	4.89	4.43	5.23	-0.79	[0.759]
No. Roma per class	5.1	4.84	5.32	-0.48	[0.854]
(if at least a Roma)					
Roma per school (%)	21	20	22	-2	[0.839]
Female					
Roma	0.49	0.52	0.46	0.06	[0.120]
Non-Roma	0.47	0.48	0.47	0.01	[0.551]
Born in the same town					
Roma	0.35	0.40	0.32	0.08	[0.654]
Non-Roma	0.76	0.81	0.73	0.08	[0.209]
Number of schools	9	5[a]	4		
Number of Roma pupils	581	231	350		
Number of Non-Roma pupils	2133	927	1206		
Characteristics of the assistants					
Female	0.875	0.8	1	-0.2	[0.374]
Maximum level of education					
Secondary school	0.5	0.6	0.33	0.27	[0.543]
University	0.375	0.4	0.33	-0.07	[0.877]
Experience with Roma	0.75	1	0.33	0.67	[0.183]
Experience in NGO	1	1	1	0	[.]
Number of assistants	8	5	3[b]		

[a] *Early Enrollees* schools are 6. One school did not provide us the list of students so it is excluded from our sample.

[b] We could not get information about one assistant in the *Late Enrollees* schools.

36

Table 2: Means of control variables in treated and control households

Variables at the household level	All	Treatment	Control	Difference
Wealth[a]	0.08	-0.14	0.22	-0.36
				(0.27)
Monthly Total income (in dinars)[b]	28949.47	28224.39	29453.33	-1228.94
				(2144.78)
Informal (=1)[c]	0.32	0.31	0.32	-0.01
				(0.05)
Urban (=1)	0.51	0.47	0.53	-0.06
				(0.06)
Only Roma in neighborhood (=1)[d]	0.21	0.28	0.16	0.12**
				(0.05)
No schooling/Unfinished primary school (=1)[e]	0.07	0.07	0.07	0.00
				(0.03)
Finished primary school (=1)[e]	0.74	0.69	0.76	-0.07
				(0.05)
Finished secondary school (=1)[e]	0.19	0.23	0.16	0.07
				(0.05)
Muslim (=1)	0.71	0.57	0.80	-0.23***
				(0.05)
Number of children under 5	0.72	0.75	0.70	0.05
				(0.10)
Number of female children between 6 and 15	1.65	1.73	1.59	0.14
				(0.14)
Number of male children between 6 and 15	1.75	1.80	1.80	0.10
				(0.13)
Number of adults	2.44	2.46	2.44	0.02
				(0.12)
max no. observations	300	122	178	
Variables at the individual level				
Children characteristics				
Male (=1)	0.52	0.50	0.54	-0.04
				(0.04)
Age of child	9.89	10.11	9.74	0.37
				(0.20)
Rank among siblings	2.20	2.33	2.11	0.22**
				(0.09)
Mark in Mathematics[f]	2.77	2.86	2.70	0.16*
				(0.09)
Mark in Serbian[f]	2.85	2.94	2.79	0.15
				(0.09)
max no. observations	673	280	393	

Standard errors in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] The wealth index was calculated with principal component analysis. The index ranges between -5.55 and 3.69.

[b] 28950 dinars corresponds to 279 Euro (1 RSD = 0.009626 Euro, November 2011).

[c] =1 if at least one household member works in the informal sector.

[d] =1 if the respondent declared that the household lives in an exclusively Roma neighbourhood.

[e] It refers to the highest level of education obtained by parents.

[f] We use demeaned mark in Mathematics and Serbian. The mark is demeaned from the average school mark.

Table 3: Means of outcome variables in treated and control households

Variables at the household level	All	Treatment	Control	Difference
Probability to find a job: Boys				
With primary school (=1)[a]	0.42	0.35	0.48	-0.13**
				(0.06)
With secondary school (=1)[a]	0.82	0.82	0.82	0.00
				(0.05)
Probability to find a job: Girls				
With primary school (=1)[a]	0.35	0.31	0.39	-0.08
				(0.06)
With secondary school (=1)[a]	0.79	0.74	0.82	-0.07
				(0.05)
Expected mean log earning: Boys				
With primary school	9.91[b]	9.97	9.87	0.10
				(0.06)
With secondary school	10.21[c]	10.24	10.18	0.06*
				(0.19)
Expected mean log earning: Girls				
With primary school	9.82[d]	9.90	9.78	0.12*
				(0.07)
With secondary school	10.14[e]	10.18	10.11	0.07*
				(0.04)
Variables at the individual level				
Expected to finish : Boys				
Secondary school (=1)	0.61	0.67	0.57	0.10*
				(0.06)
Expected to finish : Girls				
Secondary school (=1)	0.63	0.67	0.60	0.07
				(0.06)

Standard errors in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] Respondent expects the child to find a job given a certain level of education achieved.

[b] The corresponding average earning is 21709 dinars (208 Euro, Nov 2011). For treated households is 22985 dinars (221 Euro, Nov 2011); for control households is 21075 dinars (202 Euro, Nov 2011).

[c] The corresponding average earning is 28654 dinars (276 Euro, Nov 2011). For treated households is 29398 dinars (283 Euro, Nov 2011); for control households is 28141 dinars (271 Euro, Nov 2011).

[d] The corresponding average earning is 19432 dinars (187 Euro, Nov 2011). For treated households is 20915 dinars (201 Euro, Nov 2011); for control households is 18682 dinars (180 Euro, Nov 2011).

[e] The corresponding average earning is 26923 dinars (259 Euro, Nov 2011). For treated households is 27529 dinars (265 Euro, Nov 2011); for control households is 26527 dinars (255 Euro, Nov 2011).

Table 4: *Early Enrollees*: means of control variables for treated and untreated households

Variables at the household level	Treated	Untreated	Difference
Wealth[a]	-0.33	0.21	-0.54
			(0.44)
Total income (in dinars)[b]	29094	27708	1386
			(3215)
Informal (=1)[c]	0.35	0.27	0.08
			(0.088)
Urban (=1)	0.55	0.34	0.21**
			(0.09)
Only Roma in neighborhood (=1)[d]	0.31	0.24	0.07
			(0.082)
No schooling/unfinished primary school (=1)[e]	0.10	0.04	0.06
			(0.05)
Finished primary school (=1)[e]	0.63	0.78	-0.15*
			(0.082)
Finished secondary school (=1)[e]	0.27	0.18	0.09
			(0.075)
Number of children under 5	0.82	0.68	0.14
			(0.16)
Number of female children between 6 and 18	1.79	1.68	0.11
			(0.21)
Number of male children between 6 and 18	1.83	1.76	0.07
			(0.22)
Number of adults	2.48	2.42	0.06
			(0.19)
Muslim (=1)	0.56	0.58	-0.02
			(0.09)
max no. observations	71	50	
Variables at the individual level			
Children characteristics			
Male (=1)	0.49	0.52	-0.03
			(0.06)
Age of child	9.78	10.55	-0.77**
			(0.31)
Rank among siblings	2.37	2.28	0.09
			(0.16)
Demeaned grade in Mathematics[f]	0.05	-0.08	0.13
			(0.14)
Demeaned grade in Serbian[f]	0.10	-0.17	0.27
			(0.13)
max no. observations	166	112	

Standard errors in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

[a] The wealth index was calculated with principal component analysis. The index ranges between -5.55 and 3.69.

[b] 28950 dinars corresponds to 279 Euro (1 RSD = 0.009626 Euro, November 2011).

[c] =1 if at least one household member works in the informal sector.

[d] =1 if the respondent declared that the household lives in an exclusively Roma neighbourhood.

[e] It refers to the highest level of education obtained by a household member.

[f] The grade has been demeaned from the average grade of the school.

Table 5: All outcomes by education level and by gender

Gender	Boys				Girls			
Max. level of education	Primary school		Secondary School		Primary school		Secondary school	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Probability to find a job with primary/secondary school								
treatment	-0.109	-0.068	0.004	0.012	-0.067	-0.114	-0.067	-0.037
	(0.087)	(0.078)	(0.057)	(0.055)	(0.083)	(0.078)	(0.055)	(0.054)
controls[a]	no	yes	no	yes	no	yes	no	yes
No. observations	300	276	300	276	294	268	292	267
R-squared	0.012	0.141	0.000	0.067	0.005	0.121	0.007	0.123
Expected log earnings with primary/secondary school								
treatment	0.128	0.107*	0.079	0.092*	0.149*	0.123	0.079	0.100*
	(0.078)	(0.059)	(0.060)	(0.050)	(0.083)	(0.085)	(0.061)	(0.056)
controls[a]	no	yes	no	yes	no	yes	no	yes
No. observations	129	119	246	224	105	98	232	216
R-squared	0.031	0.199	0.017	0.123	0.050	0.241	0.017	0.147
Secondary school as the highest expected level of education								
treatment			0.097	0.123*			0.067	0.003
			(0.073)	(0.066)			(0.080)	(0.086)
controls[b]			no	yes			no	yes
No. observations			299	232			275	221
R-squared			0.009	0.346			0.005	0.230

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

[a] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults.

[b] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults, age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.

Table 6: Helped by the assistant: all outcomes by education level and by gender

Gender	Boys		Girls	
Max. level of education	Primary school	Secondary School	Primary school	Secondary school
	(1)	(2)	(3)	(4)
Probability to find a job with primary/secondary school				
assistant	-0.135	0.032	-0.224	-0.061
	(0.149)	(0.111)	(0.152)	(0.103)
controls[a]	yes	yes	yes	yes
No. observations	276	276	268	267
R2	0.143	0.070	0.113	0.120
Expected log earnings with primary/secondary school				
assistant	0.285*	0.190*	0.284	0.194*
	(0.165)	(0.109)	(0.189)	(0.116)
controls[a]	yes	yes	yes	yes
No. observations	119	224	98	216
R2	0.162	0.047	0.217	0.094
Secondary school as the highest expected level of education				
assistant		0.260*		0.007
		(0.136)		(0.176)
controls[b]		yes		yes
No. observations		232		221
R2		0.340		0.231
First stage - Being helped by the assistant				
treatment		0.454***		0.459***
		(0.059)		(0.061)
controls[c]		yes		yes
No. observations		224		220
R-squared		0.437		0.442
F-statistic on treatment		58.88		57.14

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

[a] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults.

[b] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults, age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.

[c] The coefficients are estimated both with the controls used with the first two outcomes and with the third one. For those reported here control variables include wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults, age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.

Table 7: Test scores by gender

Gender	Maths score		Serbian score		Able to read		Able to write	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
treatment	0.133	0.085	0.280*	0.304*	0.187	0.107	0.170	0.100
	(0.149)	(0.158)	(0.154)	(0.159)	(0.147)	(0.161)	(0.138)	(0.124)
controls[a]	yes	yes	yes	yes	yes	yes	yes	yes
No. observations	189	153	185	153	189	155	184	154
R-squared	0.210	0.186	0.172	0.046	0.136	0.106	0.228	0.131

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

[a] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults, age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.

Table 8: Worst performers: all outcomes by education level and by gender

Gender	Boys		Girls	
Max. level of education	Primary school	Secondary School	Primary school	Secondary school
	(1)	(2)	(3)	(4)
Probability to find a job with primary/secondary school				
treatment	-0.279*	0.007	-0.035	0.054
	(0.163)	(0.136)	(0.190)	(0.115)
controls[a]	yes	yes	yes	yes
No. observations	121	121	116	116
R-squared	0.208	0.105	0.062	0.233
Expected log earnings with primary/secondary school				
treatment	0.459	0.089	0.282	0.057
	(0.304)	(0.144)	(0.243)	(0.145)
controls[a]	yes	yes	yes	yes
No. observations	63	99	49	98
R-squared	0.394	0.109	0.299	0.109
Secondary school as the highest expected level of education				
treatment		0.505***		0.139
		(0.186)		(0.204)
controls[b]		yes		yes
No. observations		86		62
R-squared		0.286		0.419

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

[a] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults.

[b] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults, age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.

Table 9: Number of Roma per class

Gender	Boys		Girls	
Max. level of education	Primary school	Secondary School	Primary school	Secondary school
	(1)	(2)	(3)	(4)
Probability to find a job with primary/secondary school				
treatment	0.213	0.157	0.351	0.094
	(0.268)	(0.174)	(0.245)	(0.177)
number of Roma per class	0.001	-0.010	0.009	-0.017**
	(0.013)	(0.013)	(0.015)	(0.008)
treatment*	-0.013	-0.007	-0.021	-0.006
number of Roma per class	(0.029)	(0.019)	(0.026)	(0.018)
controls[a]	yes	yes	yes	yes
Total effect	0.200	0.150	0.330	0.088
	(0.242)	(0.16)	(0.221)	(0.16)
No. observations	121	121	120	120
R-squared	0.199	0.189	0.119	0.251
Expected log earnings with primary/secondary school				
treatment	0.262	0.216	0.164	0.229
	(0.269)	(0.148)	(0.335)	(0.159)
number of Roma per class	0.013	0.017**	0.012	0.014
	(0.018)	(0.007)	(0.024)	(0.008)
treatment*	-0.004	-0.027*	0.030	-0.020
number of Roma per class	(0.033)	(0.016)	(0.053)	(0.019)
controls[a]	yes	yes	yes	yes
Total effect	0.258	0.189	0.193	0.209
	(0.242)	(0.130)	(0.288)	(0.142)
No. observations	43	99	34	94
R-squared	0.338	0.256	0.404	0.254
Secondary school as the highest expected level of education				
treatment		0.207		-0.330
		(0.204)		(0.241)
number of Roma per class		0.003		-0.012
		(0.019)		(0.015)
treatment*		-0.028		0.034
number of Roma per class		(0.029)		(0.023)
controls[b]		yes		yes
Total effect		0.179		-0.296
		(0.182)		(0.221)
No. observations		91		100
R-squared		0.441		0.289

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

[a] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults.

[b] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults, age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.

Table 10: Heterogeneous effects: all outcomes for secondary school by gender

Gender	Boys		Girls		Boys	Girls
Max. level of education			Secondary School			
	(1)	(2)	(3)	(4)	(5)	(6)
Probability to find a job with primary/secondary school						
treatment	-0.031	0.057	-0.045	-0.006		
	(0.068)	(0.058)	(0.076)	(0.058)		
muslim	-0.008		0.022			
	(0.067)		(0.078)			
treatment*muslim	0.066		0.012			
	(0.098)		(0.107)			
only Roma in neighborhood		0.079		-0.068		
		(0.090)		(0.093)		
treatment*only Roma in neighborhood		-0.200		-0.132		
		(0.135)		(0.155)		
controls[a]	yes	yes	yes	yes		
No. observations	276	276	268	267		
R-squared	0.144	0.076	0.135	0.127		
Expected log earnings with primary/secondary school						
treatment	0.013	0.070	0.001	0.083		
	(0.085)		(0.122)			
muslim	0.077		0.065			
	(0.082)		(0.116)			
treatment*muslim	0.121		0.150			
	(0.109)		(0.137)			
only Roma in neighborhood		-0.018		0.054		
		(0.087)		(0.106)		
treatment*only Roma in neighborhood		0.102		0.085		
		(0.112)		(0.123)		
controls[a]	yes	yes	yes	yes		
No. observations	224	224	216	216		
R-squared	0.131	0.127	0.158	0.150		
Secondary school as the highest expected level of education						
treatment	0.214**	0.136*	0.163	0.067	0.093	-0.159
	(0.094)	(0.078)	(0.143)	(0.094)	(0.093)	(0.116)
muslim	-0.134		-0.093			
	(0.093)		(0.131)			
treatment*muslim	-0.146		-0.223			
	(0.138)		(0.158)			
only Roma in neighborhood		-0.002		0.063		
		(0.103)		(0.133)		
treatment*only Roma in neighborhood		-0.065		-0.274		
		(0.174)		(0.219)		
young					0.055	-0.378**
					(0.116)	(0.144)
treatment*young					0.056	0.279**
					(0.108)	(0.133)
controls[b]	yes	yes	yes	yes	yes	yes
No. observations	232	232	221	221	232	221
R-squared	0.350	0.347	0.236	0.238	0.348	0.264

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

[a] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults.

[b] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults, age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.

B Figures

Figure 1: Comparison of real returns to education for Roma and Non-Roma (Belgrade)
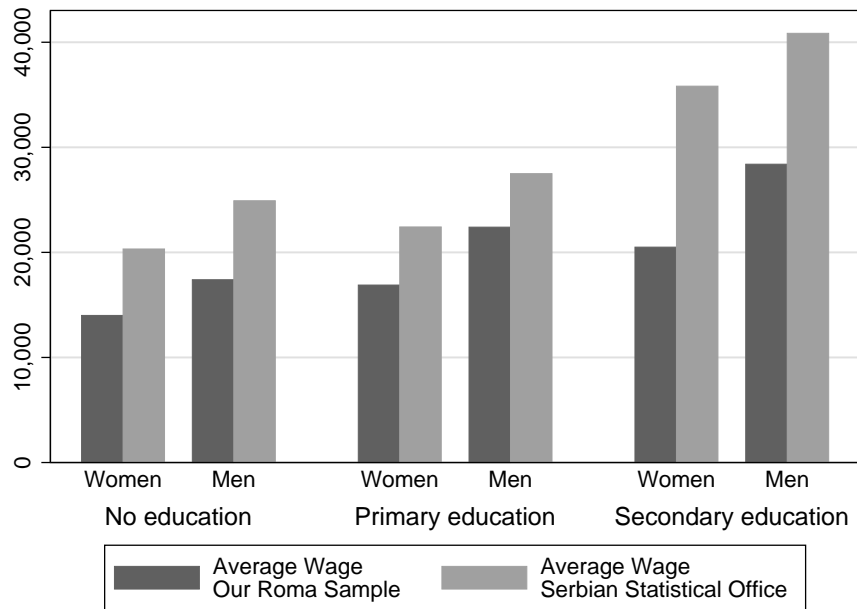
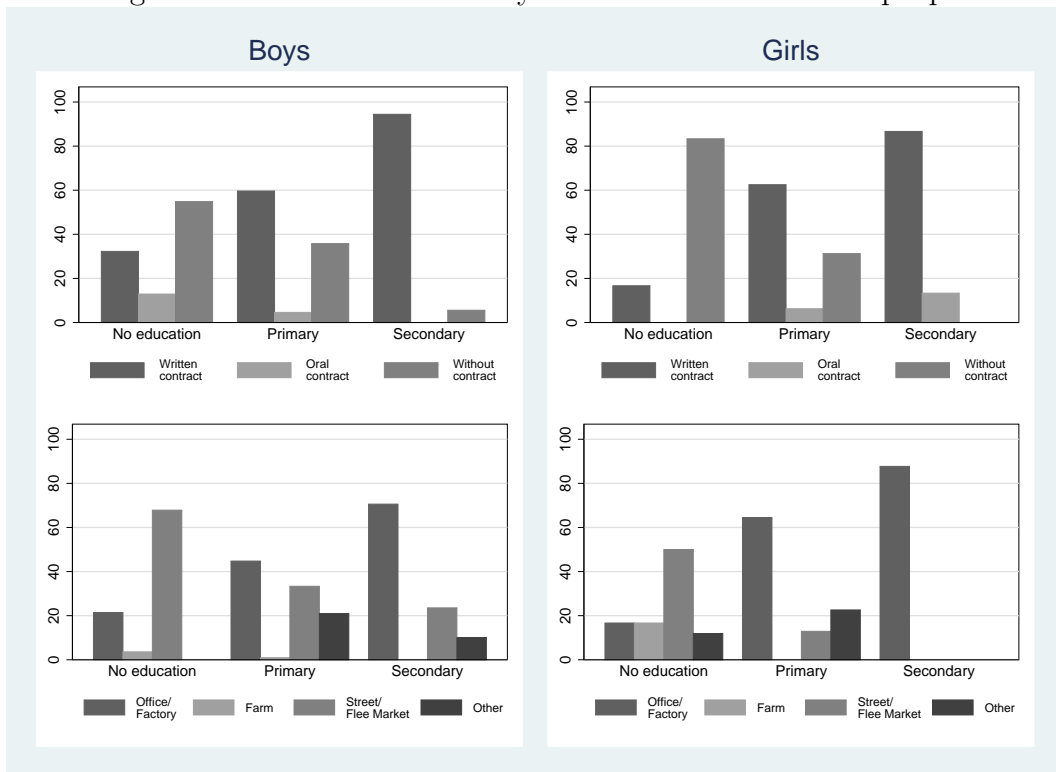Figure 2: Job characteristics by education levels - Roma people

Figure 3: Comparison of real and expected returns to education

The vertical dashed (solid) lines mark the estimated (expected) average returns in dinars.
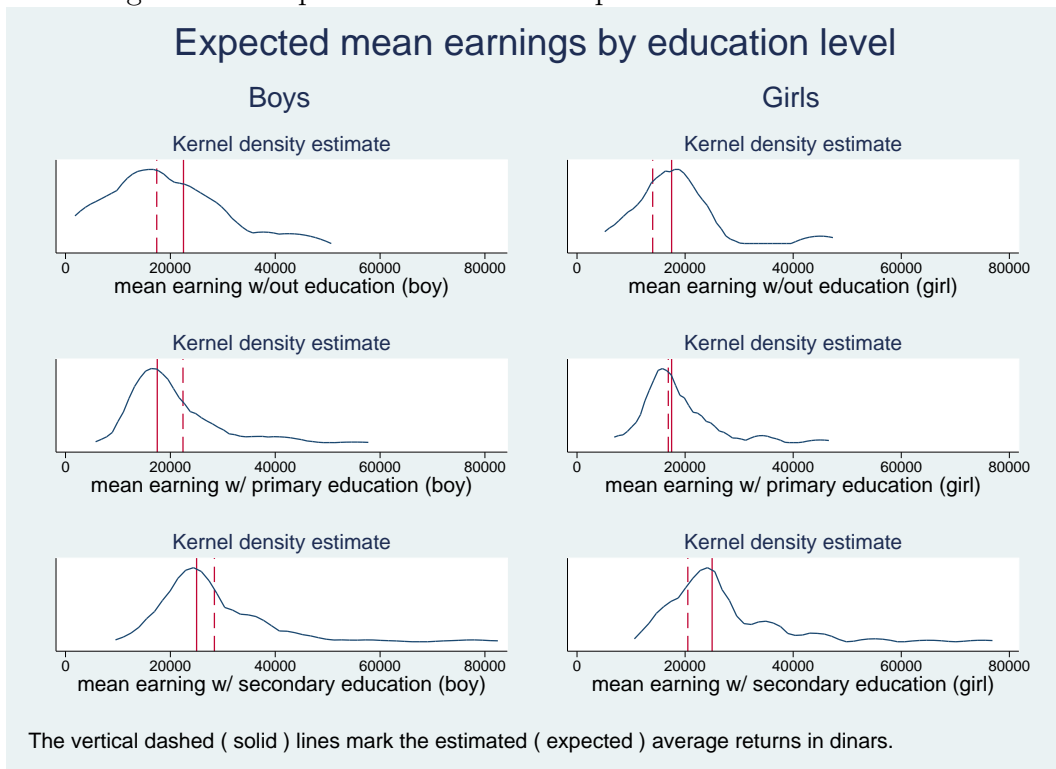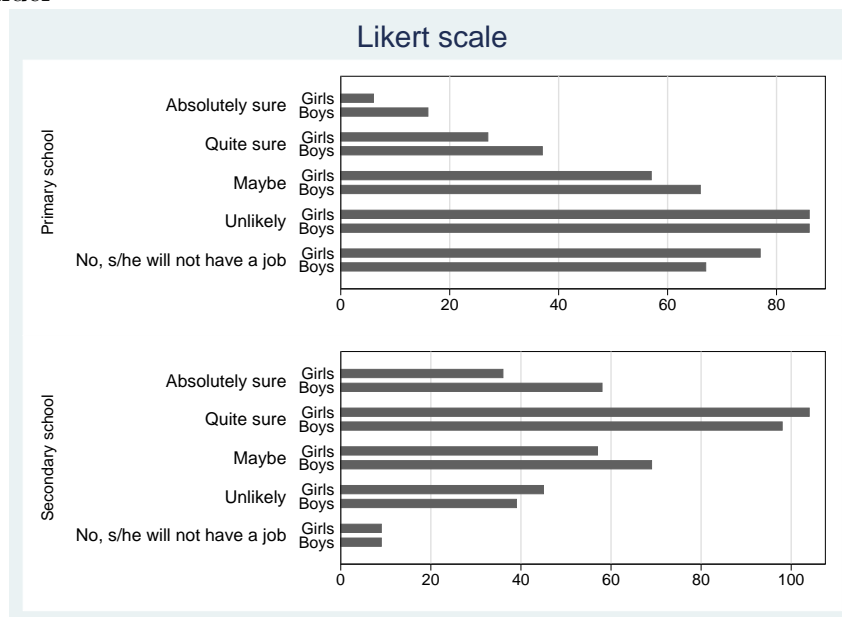
Figure 4: Settlements of the survey

Figure 5: Likert scale for the probability of finding a job with primary and secondary school by gender

A Appendix

Table A.1: Ordered Logit - thresholds among categories

	Boys		Girls	
	Primary	Secondary	Primary	Secondary
	(1)	(2)	(3)	(4)
treatment	0.089	-0.422	0.166	0.083
	(0.309)	(0.372)	(0.390)	(0.349)
controls	yes	yes	yes	yes
cut1				
constant	-2.974***	-0.914	-3.457***	-0.338
	(0.661)	(0.704)	(0.720)	(0.864)
cut2				
constant	-1.535***	0.809	-1.573**	1.876**
	(0.611)	(0.691)	(0.684)	(0.871)
cut3				
constant	-0.131	2.131***	0.059	3.234***
	(0.580)	(0.721)	(0.659)	(0.874)
cut4				
constant	1.563**	4.195***	1.647**	5.669***
	(0.616)	(0.888)	(0.674)	(1.032)
No. observations	276	276	268	267

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

Table A.2: Pooled sample: all outcomes by education level

Max. level of education	Primary school (1)	Secondary school (2)
Probability to find a job with primary/secondary school		
treatment	-0.109	-0.041
	(0.079)	(0.056)
treatment*male	0.037	0.051
	(0.076)	(0.036)
controls[a]	yes	yes
Total Effect	-0.072	0.010
	(0.076)	(0.052)
No. observations	532	531
R-squared	0.134	0.090
Expected log earnings with primary/secondary school		
treatment	0.109	0.096*
	(0.084)	(0.056)
treatment*male	-0.029	-0.011
	(0.081)	(0.039)
controls[a]	yes	yes
Total Effect	0.079	0.085*
	(0.059)	(0.050)
No. observations	209	431
R-squared	0.223	0.139
Secondary school as the highest expected level of education		
treatment		-0.001
		(0.084)
treatment*male		0.123
		0.081
controls[b]		yes
Total Effect		0.122*
		(0.066)
No. observations		454
R-squared		0.286

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

[a] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults.

[b] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults, age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.

Table A.3: OLS and IV estimates: all outcomes by gender

Gender	OLS		ITT		IV	
	Boys	Girls	Boys	Girls	Boys	Girls
	(1)	(2)	(3)	(4)	(5)	(6)
Probability to find a job with primary/secondary school						
Primary school	-0.129	-0.135	-0.068	-0.114	-0.135	-0.224
	(0.084)	(0.088)	(0.078)	(0.078)	(0.149)	(0.152)
Secondary school	0.003	-0.031	0.012	-0.037	0.032	-0.061
	(0.060)	(0.057)	(0.055)	(0.054)	(0.111)	(0.103)
controls[a]	yes	yes	yes	yes	yes	yes
Expected log earnings with primary/secondary school						
Primary school	0.097	0.148*	0.107*	0.123	0.285*	0.284
	(0.079)	(0.079)	(0.059)	(0.085)	(0.165)	(0.189)
Secondary school	-0.005	0.041	0.092*	0.100*	0.190*	0.194*
	(0.045)	(0.054)	(0.050)	(0.056)	(0.109)	(0.116)
controls[a]	yes	yes	yes	yes	yes	yes
Secondary school as the highest expected level of education						
Secondary school	0.152**	0.138	0.123*	0.003	0.260*	0.007
	(0.058)	(0.091)	(0.066)	(0.086)	(0.136)	(0.176)
controls[b]	yes	yes	yes	yes	yes	yes

Robust standard errors corrected for clustering at the school-cohort level in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%. The maximum number of observations with primary school is 276 for boys and 268 for girls; with secondary school is 276 for boys and 267 for girls.

[a] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults.

[b] Control variables included are wealth, informal (=1), urban (=1), only Roma in neighborhood, finished primary school (=1), finished secondary school (=1), muslim (=1), number of children under 5, number of female children between 6 and 15, number of male children between 6 and 15 and number of adults, age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.