Equal Access to Education: Remedial Education and Perceived Returns to Education of Roma People

Marianna Battaglia[1] Lara Lebedinski[2]

[1]University of Alicante

[2]FREN

FEDEA Workshop Remedial Education in Europe
19th January 2017

Facts

Roma people are the largest ethnic minority in Europe:

- poorer than other population groups
- suffer severe social exclusion
- **attain very low education**
  enrollment rates in primary school: around 60% (Brueggemann, 2012)
  completion rates in primary school: 30-40%

In Serbia: ▸ Serbia

- 60% of Roma younger than 18 have not completed primary education (country avg.: 20%) ▸ Education
- after the first 3 years of school, **Roma lag 2.2-2.5 years behind the average students** (Baucal, 2009).

Aim

Examine the impact of a remedial education program targeting primary school-age children of a minority group

1. **in the short-term**
   a. on schooling outcomes (dropouts, attendance and marks) in the first year of its implementation. *Equal Access to Education: An Evaluation of The Roma Teaching Assistant Program in Serbia, WD (2015)*

Aim

Examine the impact of a remedial education program targeting primary school-age children of a minority group

1. **in the short-term**
   a. on schooling outcomes (dropouts, attendance and marks) in the first year of its implementation. *Equal Access to Education: An Evaluation of The Roma Teaching Assistant Program in Serbia, WD (2015)*
   b. on parental expectations (educational achievement and return to education). *The Curse of Low Expectations: Remedial Education and Perceived Returns to Education of Roma People.*

Aim

Examine the impact of a remedial education program targeting primary school-age children of a minority group

1. **in the short-term**
   a. on schooling outcomes (dropouts, attendance and marks) in the first year of its implementation. *Equal Access to Education: An Evaluation of The Roma Teaching Assistant Program in Serbia, WD (2015)*
   b. on parental expectations (educational achievement and return to education). *The Curse of Low Expectations: Remedial Education and Perceived Returns to Education of Roma People.*
2. **in the medium-term**
   a. on schooling outcomes at the end of primary school (dropouts/change of school and marks) and on standardized test scores.
   b. peer effects for Non-Roma. *Medium-term Effects of The Roma Teaching Assistant Program in Serbia*

Papers' contribution

- There are not systematic studies in economic literature that try to investigate how to improve life circumstances of Roma, in particular kids *(O'Higgins and Brueggemann (2014); Rostas and Kostka (2014); Garaz (2014); Baucal (2006); Kertesi and Kezdi (2011)).*
- They add evidence on short-term and medium-term effects of **remedial education targeting a stigmatized ethnic group**.
- They offer **primary data** in a context where they are scarce (few official data on Roma people).

Papers' contribution

- There are not systematic studies in economic literature that try to investigate how to improve life circumstances of Roma, in particular kids *(O'Higgins and Brueggemann (2014); Rostas and Kostka (2014); Garaz (2014); Baucal (2006); Kertesi and Kezdi (2011)).*
- They add evidence on short-term and medium-term effects of **remedial education targeting a stigmatized ethnic group**.
- They offer **primary data** in a context where they are scarce (few official data on Roma people).
- They suggest the importance of the remedial education mechanism, **together with the provision of a role model**, in increasing households' current investment in education.

Literature

- **remedial education programs** targeting underachieving students: *Lavy and Schlosser (2005); Hanushek et al. (2002); Jacob and Lefgren (2004); Banarjee et al. (2007)*.
- **programs aiming at improving schooling outcomes of minority communities and the poor**: after-school programs *(Lauer et al., 2006)*, merit pay for principals, teachers and students *(Podgursky and Springer, 2007; Roland G. Fryer, 2010)*, professional development for teachers *(Boyd et al., 2008)*, involvement of parents *(Domina, 2005)*, buses to better schools *(Angrist and Lang, 2004)*, alter the neighborhoods where to live *(Jacob, 2004; Sanbonmatsu et al., 2006)*.
- **role models for minorities**: *Dee (2004); Akerlof and Kranton, (2000); Krishnan and Krutikova (2010); Ladson-Billings (1994).*
- **subjective expected returns to education**: *Nguyen, 2008; Kaufmann and Attanasio, 2009; Jensen, 2010.*

- The **Roma Teaching Assistant Program** is the main program in Central and Eastern Europe targeting Roma inclusion in education.
  1. Each school receives **one teaching assistant**.
     Assistants' duties (especially in lower grades):
     1. Help children during regular classes;
     2. Organize additional classes;
     3. Help with homework and assignments;
     4. Once a week visit parents.
  2. They are Roma and live in the same municipality of the school.

- The **Roma Teaching Assistant Program** is the main program in Central and Eastern Europe targeting Roma inclusion in education.
  1. Each school receives **one teaching assistant**.
     Assistants' duties (especially in lower grades):
     1. Help children during regular classes;
     2. Organize additional classes;
     3. Help with homework and assignments;
     4. Once a week visit parents.
  2. They are Roma and live in the same municipality of the school.

- Gradual implementation:
  - **Early Enrollees**: 26 schools (out of 78) in **September 2009**;
  - **Late Enrollees**: 77 more schools (out of 252) in **November 2010**.

- Schools and assistants needed to apply. Criteria for selection:
  - **schools**: percentage of Roma students (5%-40%) and, preferably, availability of preschool program in the school;
  - **assistants**: secondary school attainment, knowledge of Romani, and experience in working with children.

- Schools and assistants needed to apply. Criteria for selection:
  - **schools**: percentage of Roma students (5%-40%) and, preferably, availability of preschool program in the school;
  - **assistants**: secondary school attainment, knowledge of Romani, and experience in working with children.

- The phasing in of the program can be treated as it was exogenous:

1. The same selection criteria in both years and the selection committee rated schools in the same way;
2. schools and assistants could apply in both years and those applying before and after do not differ (around 13% of Roma [p-value=0.458]).
3. schools and assistants selected in the first year are not different in observable characteristics from schools selected later. ▸ Characteristics

- One concern: motivation may differ between schools applying before and schools applying later. If these motivations are related to differences in principle or school quality, possible selection bias.

- One concern: motivation may differ between schools applying before and schools applying later. If these motivations are related to differences in principle or school quality, possible selection bias.
4. this concern is mitigated since school quality is not different between selected schools (average marks, absences and dropout rates of previous years).
- Principal quality should be reflected in assistant quality to affect parents' expectations but assistants are not chosen by the principal.

- One concern: motivation may differ between schools applying before and schools applying later. If these motivations are related to differences in principle or school quality, possible selection bias.
4 this concern is mitigated since school quality is not different between selected schools (average marks, absences and dropout rates of previous years).
- Principal quality should be reflected in assistant quality to affect parents' expectations but assistants are not chosen by the principal.

- There is not selective sorting into *Early Enrollees* school.

A. In the short-term, data from administrative records on 38 schools:

- **Early Enrollees** - in 3 schools we were not allowed to collect data;
- **Late Enrollees** - we selected 15 schools out of 77 based on the following criteria:
  1. the same district;
  2. rural/urban municipality;
  3. similar school size;
  4. similar percentage of Roma pupils of a *Early Enrollees*. ▸ Map

Table 1: Program timeline ▸ Characteristics

	2009 Early Enrollees	2010 Late Enrollees
Number of schools joining the programme	26	77
Number of schools in sample	23	15

- From 2006/2007 until 2009/2010; 1 to 4 grades of primary school.

B. In the short-term, survey first-hand collected data:

- survey with 300 Roma households in Belgrade ▸ Figure;
- in Fall 2010: one year after the implementation of the program in *Early Enrollees* and before *Late Enrollees* received the assistants;

Table 2: Households interviewed

	2009 Early Enrollees	2010 Late Enrollees
No. schools in the program	5	8
No. households interviewed	122	178

▸ Characteristics

- pupils randomly selected among students attending schools;
- at least one child in the lower four grades of primary school in 2009/2010.
- Information on expectations. ▸ Measures

C. In the medium-term, administrative data (Sept. 2008- 2016):

- from administrative records - marks for the 8th grade and likelihood of finishing the same primary school on time for the same 38 schools.
- from a national examination - standardized test scores at the end of primary school (8th grade) for the same schools.

Table 3: Program timeline

	2008	2009	2010	...	2012	...	2014	2015
Early Enrollees	1st grade	**2nd grade**	3rd grade	...	5th grade	...	7th grade	8th grade
	2nd grade	**3rd grade**	4th grade	...	6th grade	...	8th grade	
	...	....	...	...	...	...		
	4th grade	5th grade	6th grade	...	8th grade			
	...	....	...	...				
	8th grade							
Late Enrollees	1st grade	2nd grade	**3rd grade**	...	5th grade	...	7th grade	8th grade
	2nd grade	3rd grade	**4th grade**	...	6th grade	...	8th grade	
	...	....	...	...	...	...		
	4th grade	5th grade	6th grade	...	8th grade			
	...	....	...	...				
	8th grade							

A. Econometric Specification

Ideal experiment: random selection of the schools assigned to the program. It is not the case. So, as **Control Groups**:

1. *Late Enrollees*: exploit the gradual implementation of the program
   - Advantage: the impact of the program is not confounded with other government policies of 2009 (free books and repetition rates);
   - Disavantage: we are not able to control for unobservable differences which have led some schools to enter the program before others.

A. Econometric Specification

Ideal experiment: random selection of the schools assigned to the program. It is not the case. So, as **Control Groups**:

1. *Late Enrollees*: exploit the gradual implementation of the program
   - Advantage: the impact of the program is not confounded with other government policies of 2009 (free books and repetition rates);
   - Disavantage: we are not able to control for unobservable differences which have led some schools to enter the program before others.
2. *Older cohorts* - less exposed to the program. Assistants work mainly with first grades.
   - Advantage: solve possible selection bias;
   - Disavantage: it relies on the strong assumptions that there were no government interventions over the period and regular trends of the outcomes.

1. Average treatment approach - *Early versus Late Enrollees*

$$Y_{ijt} = \beta_0 + \delta_t + \rho_j + \beta_1 treatment_j * post_t + \gamma_1 X'_{ijt} + \varepsilon_{ijt}$$

2. Cohort regressions - *younger (grade 1) versus older cohorts (grades 2, 3 and 4) in pre- and treatment year in treated and control school*

$$Y_{ijt} = \beta_0 + \beta_1 young_i + \beta_2 post_t + \beta_3 young_i * post_t + \gamma_1 X'_{ijt} + \varepsilon_{ijt}$$

3. Triple interaction: *within and between schools*

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

Table 4: Short-term: schooling outcomes (1)

	Average treatment	Cohort regression		Triple Interaction
	(1)	(2)	(3)	(4)
	All	Early Enrollees	Late Enrollees	All
		Dropout[a]		
post	0.015**			
	(0.006)			
treatment*post	0.003			
	(0.009)			
young*post		0.014	0.079***	0.080***
		(0.013)	(0.019)	(0.018)
young*post*treatment				**-0.066*** **
				(0.022)
		Absences		
post	32.853***			
	(9.078)			
treatment*post	**-16.679***			
	(9.078)			
young*post		23.579*	57.493**	54.639**
		(11.537)	(21.263)	(22.235)
young*post*treatment				-31.867
				(24.945)
Max no. observations	4039	2438	1601	4039
School FE	Yes	Yes	Yes	Yes

Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%. Controls included are school size (squared), number of Roma in school (squared), percentage of Roma per class, class size (squared), female (=1), age (squared), and migrant (=1).

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

Table 5: Short-term: schooling outcomes (2)

▸ By gender

	Average treatment	Cohort regression		Triple Interaction
	(1)	(2)	(3)	(4)
	All	Early Enrollees	Late Enrollees	All
		Serbian[a]		
post	0.046			
	(0.048)			
treatment*post	0.012			
	(0.066)			
young*post		0.079	-0.255**	-0.300***
		(0.102)	(0.104)	(0.101)
young*post*treatment				**0.382****
				(0.149)
		Mathematics[b]		
post	0.065			
	(0.062)			
treatment*post	0.030			
	(0.077)			
young*post		0.082	-0.241	-0.291*
		(0.116)	(0.143)	(0.152)
young*post*treatment				**0.381***
				(0.194)
Max no. observations	4039	2438	1601	4039
School FE	Yes	Yes	Yes	Yes

Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%. Controls included are school size (squared), number of Roma in school (squared), percentage of Roma per class, class size (squared), female (=1), age (squared), and migrant (=1).

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

Additional results

1. Intensity of treatment approach: whether the effect of the program varies with the number of Roma per school. Each school has only one assistant: the higher the number of Roma per school, the less intense is the program.

Additional results

1. Intensity of treatment approach: whether the effect of the program varies with the number of Roma per school. Each school has only one assistant: the higher the number of Roma per school, the less intense is the program. Positive impact on every outcome in schools with fewer Roma. The intensity of the program affects especially girls.

Additional results

1. Intensity of treatment approach: whether the effect of the program varies with the number of Roma per school. Each school has only one assistant: the higher the number of Roma per school, the less intense is the program. Positive impact on every outcome in schools with fewer Roma. The intensity of the program affects especially girls.

2. Spillover effects: Non Roma pupils. Neither absences reduced nor marks improved for Non Roma students. The program is succeeding in reducing the gap between Roma and Non Roma children:

   Before:

   - on a grading scale of 1 to 5, a difference of almost 2 **grades** in Mathematics and Serbian.
   - Roma children were **absent** from school approximately 3 to 4 times as much as Non Roma children.

B. Econometric specification on parental expectations

Underinvestment in education due to:

- Supply constraints - absence of documents, language barriers, discrimination from teachers and pupils (last row and special schools);
- Demand constraints - financial constraints, **expected returns**:

Parents attach little value to schooling: too high costs as compared to **perceived** discounted stream of expected future benefits.

B. Econometric specification on parental expectations

Underinvestment in education due to:

- Supply constraints - absence of documents, language barriers, discrimination from teachers and pupils (last row and special schools);
- Demand constraints - financial constraints, **expected returns**:

Parents attach little value to schooling: too high costs as compared to **perceived** discounted stream of expected future benefits.

Reasons to believe that the lack of goals and expectations is important:

1. a large percentage of Roma lives in segregated settlements;
2. they rarely perform jobs requiring high education levels;
3. mean earnings of Roma workers are lower than Non-Roma. ▸ Figure

- Objective returns vs. **Subjective returns**

- Objective returns
  1. **among Roma** large differences in average earnings across different education levels:
     - average wages with secondary vs primary educ.:
       21% higher for girls and 27% for boys (our sample).
  2. the higher the education level, the better the jobs performed ▸ Figure

- Subjective returns
  Parents are not aware of the situation. ▸ Figure

1. *Early Enrollees* vs. *Late Enrollees* schools ▸ Means of control variables

$$Y_{(i)j} = \beta_0 + \beta_1 treatment_j + \beta_2 X'_{(i)j} + \varepsilon_{(i)j}$$

$treatment_j$: =1 if there is at least one child i in the household j who goes to a treated school (*Early Enrollees*); ▸ Results

2. Pupils helped by the assistant (**selected subset** in *Early Enrollees*)

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

$assistant_j$: =1 if there is at least one child i in the household j who goes to a treated school (*Early Enrollees*) and who is **helped** by the assistant;

1. The assistant chooses the pupils she works with: **selected subset** of those in *Early Enrollees* schools ➔ IV strategy ▸ Characteristics

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

- *Early Enrollees* (ITT) is the instrument for treatment received - being helped by the assistant (LATE = treatment on the treated)

  θ_1 captures LATE = the effect of treatment on the treated

  ▸ OLS and IV ▸ First-stage

Table 6: Helped by the assistant

Gender	Boys		Girls	
Max. education	Primary school	Secondary School	Primary school	Secondary school
	(1)	(2)	(3)	(4)
Probability to find a job				
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
Expected log earnings				
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
Secondary school as the highest expected education level				
assistant		**0.260***		0.007
		(0.136)		(0.176)
controls[b]		yes		yes
No. observations		232		221

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] Additional control variables are age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.

- Which are the channels?
  1. remedial education (*test scores*) ▸ questions
  2. role model

Table 7: Test scores

Gender	**Boys**	**Girls**	**Boys**	**Girls**
	(1)	(2)	(1)	(2)
	Maths score		**Serbian score**	
assistant	0.133	0.085	**0.280***	**0.304***
	(0.149)	(0.158)	(0.154)	(0.159)
controls	yes	yes	yes	yes
No. observations	189	153	185	153

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

- Which are the channels?
  1. remedial education (*test scores*) ▸ questions
  2. role model

Table 7: Test scores

Gender	**Boys**	**Girls**	**Boys**	**Girls**
	(1)	(2)	(1)	(2)
	Maths score		**Serbian score**	
assistant	0.133	0.085	**0.280***	**0.304***
	(0.149)	(0.158)	(0.154)	(0.159)
controls	yes	yes	yes	yes
No. observations	189	153	185	153

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

- If among the **worst performers** expectations have increased for those helped by the assistant, not only remedial education channel.

Table 8: Helped by the assistant. Worst performers.

Gender	**Boys**		**Girls**	
Max. education	Primary school	Secondary School	Primary school	Secondary school
	(1)	(2)	(3)	(4)
Probability to find a job				
assistant	-0.259	-0.025	-0.156	0.026
	(0.177)	(0.151)	(0.215)	(0.138)
controls	yes	yes	yes	yes
No. observations	151	151	145	145
Expected log earnings				
assistant	0.580	0.146	0.321	0.084
	(0.438)	(0.160)	(0.270)	(0.155)
controls	yes	yes	yes	yes
No. observations	77	121	60	118
Secondary school as the highest expected education level				
assistant		**0.392*** **		0.022
		(0.180)		(0.233)
controls		yes		yes
No. observations		128		117

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

Additional Results

1. **Pupils'** expected likelihood of achieving a secondary education level is in line with their parents. ▸ Table23

2. **Heterogeneous effects**
   1. Effects for secondary school completion mainly driven by Non-Muslim households;
   2. Revision of expectations mainly for younger kids: especially girls. ▸ Table24

C. Econometric Specification: Medium-term

Data

1 Intensity of the average treatment approach - *Early versus Late Enrollees* in September 2016 (and versus schools which applied and did not get selected)

$$Y_{ijt} = \beta_0 + \beta_1 treatment_j + \gamma_1 X'_{ijt} + \varepsilon_{ijt}$$

Y_{ijt}: Standardized test scores, marks and dropout/change of school.

2 Cohort regressions - within *Early Enrollees*, outcomes of 8th graders 2016 (treated for 7 years) versus 8th graders 2012 (never treated because in the 5th grade when the program started). Within *Late Enrollees*, outcomes of 8th graders 2016 (treated for 6 years) versus 8th graders 2012 (never treated, already at the 6th grade)

$$Y_{ijt} = \beta_0 + \beta_1 young_i + \beta_2 post_t + \beta_3 young_i * post_t + \gamma_1 X'_{ijt} + \varepsilon_{ijt}$$

3 Triple interaction: *within and between schools*

$$Y_{ijt} = \beta_0 + \beta_1 young_i + \beta_2 post_t + \beta_3 young_i * post_t + \\ + \gamma_1 treatment_j * post_t + \gamma_2 young_i * treatment_j + \\ + \gamma_3 young_i * post_t * treatment_j + \gamma_1 X'_{ijt} + \varepsilon_{ijt}$$

4 Effect of treatment on teachers' behavior. *post* is 2016 (versus 2012)

$$Testscore_{ijt} - Marks_{ijt} = \beta_0 + \delta_t + \rho_j + \beta_1 treatment_j * post_t + \gamma_1 X'_{ijt} + \varepsilon_{ijt}$$

Preliminary summary statistics

1. at the end of primary school, Roma pupils get higher test scores in *Early* than in *Late Enrollees* schools. It is not the case for Non-Roma.

Preliminary summary statistics

1. at the end of primary school, Roma pupils get higher test scores in *Early* than in *Late Enrollees* schools. It is not the case for Non-Roma.

2. in *Early Enrollees* schools, Roma pupils are slightly more likely to reach the 8th grade on time. There are not statistically significant differences for Non-Roma.
   - Almost 60% of Roma pupils who should have finished primary school in 2016 have not finished it in the same school: either they migrated, they dropout or they are repeating some grades.

Table 9: Preliminary Summary Statistics

	Early Enrollees (1)	**Late Enrollees** (2)	**Difference** (1)-(2) (3)	**P-value** (4)
	Standardized Test Scores at the 8th grade - max 30 points			
Roma	10.773	8.627	2.146	[0.0055]
	(5.529)	(4.256)		
Non-Roma	17.343	17.419	-0.076	[0.7807]
	(5.837)	(5.598)		
	Standardized Test Scores in Math at the 8th grade - max 10 points			
Roma	1.871	1.343	0.528	[0.0431]
	(1.923)	(1.298)		
Non-Roma	4.3	4.397	-0.097	[0.4312]
	(2.595)	(2.559)		
	Standardized Test Scores in Serbian at the 8th grade - max 10 points			
Roma	4.255	3.422	0.833	[0.0104]
	(2.227)	(2.047)		
Non-Roma	6.793	6.738	0.055	[0.5821]
	(2.102)	(2.095)		
	Likelihood of Reaching the 8th grade on time			
Roma	.4093	.324	.0853	[0.0300]
	(.0251)	(.0296)		
Non-Roma	.8667	.8577	.009	[0.5585]
	(.3401)	(.3495)		

Standard deviation in parentheses.

Table 10: Intensity of the average treatment approach (Roma pupils)

Standardized Test Scores at the 8th grade	
Early Enrollees	**2.045***
	(1.174)
Standardized Test Scores in Math at the 8th grade	
Early Enrollees	0.554
	(0.374)
Standardized Test Scores in Serbian at the 8th grade	
Early Enrollees	**0.812***
	(0.435)
Likelihood of reaching the 8th grade on time	
Early Enrollees	0.0815
	(0.051)
Marks in Math at the 8th grade	
Early Enrollees	0.172
	(0.175)
Marks in Serbian at the 8th grade	
Early Enrollees	**0.401*****
	(0.142)
Max no. observations	239
Controls	Yes

Robust standard errors corrected for clustering at the school level in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%. Control variables included are age of child, age of child squared, gender, class size (squared), % of Roma in class, whether the kid was born somewhere else.

Conclusions in the short-term

1. Modest effect of the program:
    1. Marks improve and dropouts reduce for those children treated in their first grade. On average, all children treated go more to school.
    2. Positive impact in schools with fewer Roma, especially for girls.

Conclusions in the short-term

1. Modest effect of the program:
    1. Marks improve and dropouts reduce for those children treated in their first grade. On average, all children treated go more to school.
    2. Positive impact in schools with fewer Roma, especially for girls.

2. Positive effect on expectations.

- Remedial education programs with a teacher of the same social background can help create role models
    - and reduce information gap on returns to education;
    - and **by opening opportunities previously unexpected to a group**;
- **increasing current investment in education**;

Conclusions in the short-term

1. Modest effect of the program:
    1. Marks improve and dropouts reduce for those children treated in their first grade. On average, all children treated go more to school.
    2. Positive impact in schools with fewer Roma, especially for girls.

2. Positive effect on expectations.

- Remedial education programs with a teacher of the same social background can help create role models
    - and reduce information gap on returns to education;
    - and **by opening opportunities previously unexpected to a group**;
- **increasing current investment in education**;

We are now working on the effects in the medium-term

- In Serbia, Census 2011: 147,604 (2.05% of the population); estimates Open Society Institute: 350,000-500,000 (6%).
- In Belgrade, Census 2011: 27,325 (1.65% of the population); estimates Open Society Institute: roughly 80,000 (5%). Back

Table 11: Characteristics of Roma versus Non-Roma

Demographic characteristics	Roma	Non-Roma
Household size	4.5	3.2
Number of children younger than 18 (in household)	2.4	0.9
Age (average)	25	42
Population younger than 18	**0.40**	**0.22**
Employed (males)	0.69	0.72
Employed (females)	0.34	0.53
Individuals below the poverty line	0.64	0.10
Education		
Children between 6 and 15 not enrolled in school	**0.29**	**0.01**
Unfinished primary school	0.36	0.14

Source: Serbia - LSMS 2003

School system in Serbia

- School is compulsory until the age of 15. Children enrol at primary school if they are aged at least 6.5 years at the start of the scholastic year (September).
- Since 2007: compulsory preschool for 6 months.

- **8 years of primary education(4+4)** - in the short-term we only look at the first 4 years/grades.
- **In the first 4 grades: only 1 teacher for all subjects**.
- **Range of marks: from 1 (worse) to 5 (best)**; if one insufficient vote, either the kid passes or s/he takes the retake exam in August.

- No school fees, but other costs such as books and other school material (since 2009/2010 free text books for the first grade). Back

Table 12: Characteristics of the selected schools in Serbia

Back

		pre-treatment year			
		Early Enrollees	**Late Enrollees**	**Difference (1-2)**	**(s.e.)**
		(1)	(2)	(3)	(4)
Class size		22.161	23.966	-1.804	(1.424)
School size		304.937	361.506	-56.569	(52.963)
% Roma per class		0.221	0.183	0.038	(0.056)
% of Roma per school		0.223	0.193	0.030	(0.056)
Age	*Roma*	8.748	8.675	0.073	(0.089)
	Non-Roma	8.421	8.400	0.0216	(0.066)
Female	*Roma*	0.502	0.471	0.031	(0.023)
	Non-Roma	0.487	0.477	0.010	(0.014)
Born in the same town	*Roma*	0.867	0.814	0.053	(0.038)
	Non-Roma	0.922	0.912	0.010	(0.011)
Number of schools		23	15		
Number of Roma pupils		1241	811		
Number of Non-Roma pupils		4303	3374		

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

Table 13: Characteristics of the selected schools in Belgrade

Back

	All	Early Enrollees	Late Enrollees	Difference (1-2)	P-value
	(1)	(2)	(3)	(4)	(5)
Class size	21.55	20.60	22.75	-2.15	[0.505]
No. Roma per class	4.67	3.80	5.75	-1.95	[0.566]
No. Roma per class (if at least a Roma)	5.25	4.75	5.75	-1.00	[0.764]
No. of Roma per school (%)	0.23	0.23	0.24	-0.01	[0.952]
Female					
Roma	0.49	0.53	0.47	0.06	[0.209]
Non-Roma	0.47	0.48	0.47	0.01	[0.544]
Born in the same town					
Roma	0.77	0.79	0.75	0.04	[0.358]
Non-Roma	0.90	0.90	0.90	0.00	[0.722]
Number of schools	9	5	4		
Number of Roma pupils	581	231	350		
Number of Non-Roma pupils	2133	927	1206		

Figure 1: Schools in our sample

Back

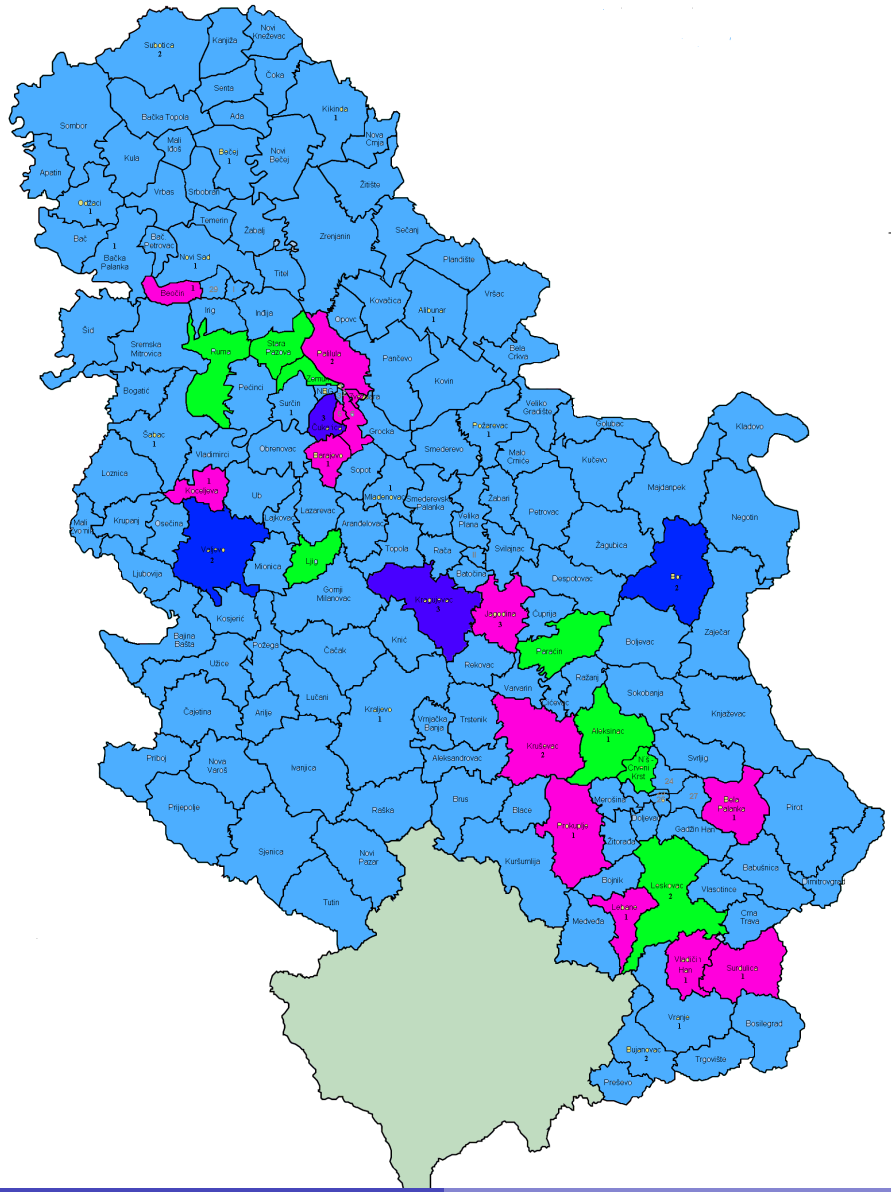


Figure 2: Settlements in Belgrade

Back

Measures of expectations: ▸ Data

1. **Expected likelihood of getting a job**

   Assume that your oldest boy has finished (primary) secondary school -and that is his highest degree- and he is 25-30 years old: how certain are you that he will get any kind of job? ▸ Likert Scale

Measures of expectations: ▸ Data

1. **Expected likelihood of getting a job**

   Assume that your oldest boy has finished (primary) secondary school -and that is his highest degree- and he is 25-30 years old: how certain are you that he will get any kind of job? ▸ Likert Scale

2. **Expected (log) mean amount of earnings, once employed**

   Assume that your oldest boy has finished (primary) secondary school 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?

Measures of expectations: ▸ Data

1. **Expected likelihood of getting a job**

   Assume that your oldest boy has finished (primary) secondary school -and that is his highest degree- and he is 25-30 years old: how certain are you that he will get any kind of job? ▸ Likert Scale

2. **Expected (log) mean amount of earnings, once employed**

   Assume that your oldest boy has finished (primary) secondary school 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?

3. **Secondary school as the highest expected educ. level:**

   What level of formal education do you think that (name) will complete?

   each child aged 6 to 15 ▸ Summary statistics

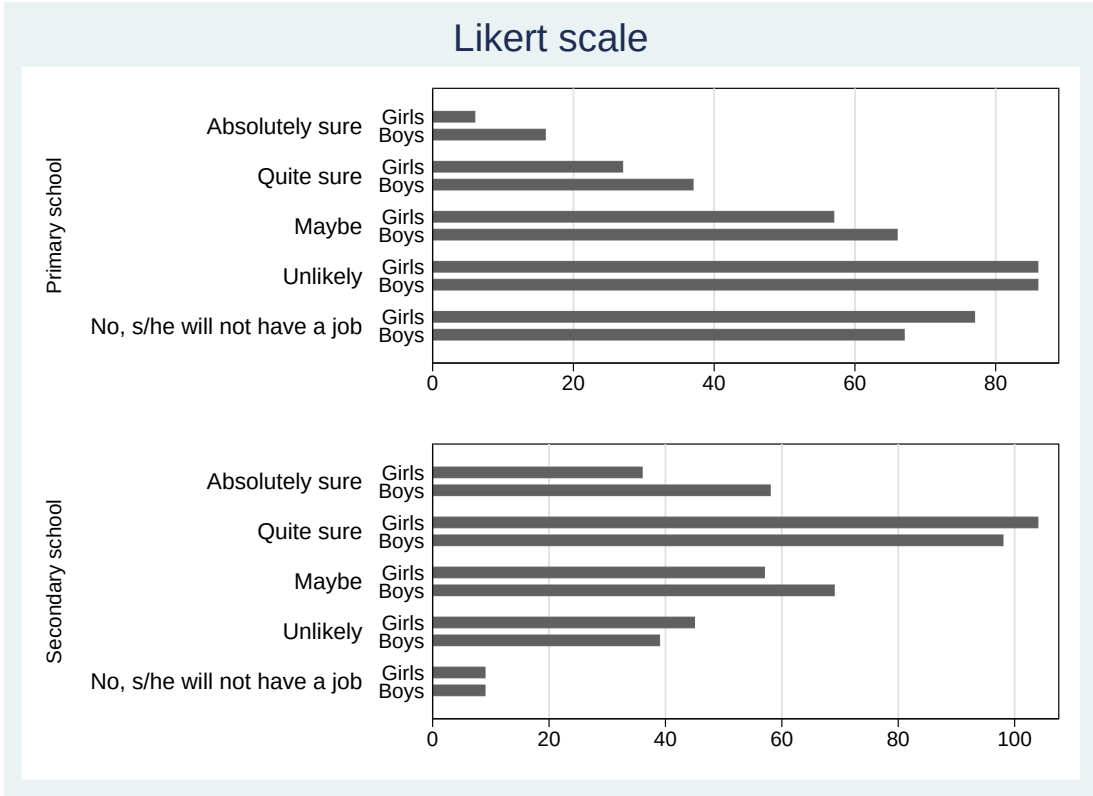
Expected likelihood of finding a job:

Absolutely sure, *Quite sure*, *Maybe*: =1

Unlikely, *No, s/he will not find a job*: =0

Back

Figure 3: Likert scale



Back

Table 14: 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)	0.42	0.35	0.48	-0.13**
				(0.06)
With secondary school (=1)	0.82	0.82	0.82	0.00
				(0.05)
Probability to find a job: Girls				
With primary school (=1)	0.35	0.31	0.39	-0.08
				(0.06)
With secondary school (=1)	0.79	0.74	0.82	-0.07
				(0.05)
Expected mean log earning: Boys				
With primary school	9.91	9.97	9.87	0.10
				(0.06)
With secondary school	10.21	10.24	10.18	0.06*
				(0.19)
Expected mean log earning: Girls				
With primary school	9.82	9.90	9.78	0.12*
				(0.07)
With secondary school	10.14	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%.

Table 15: Average treatment approach

	All	All	Girls	Boys
	(1)	(2)	(3)	(4)
			Dropout[a]	
post	0.017**	0.015**	0.001	0.027**
	(0.007)	(0.006)	(0.010)	(0.012)
treatment*post	-0.006	0.003	**0.028***	-0.018
	(0.009)	(0.009)	(0.014)	(0.014)
			Absences	
post	31.236***	32.853***	22.456***	42.034***
	(7.856)	(9.078)	(10.797)	(10.764)
treatment*post	-17.299**	-16.679*	-4.713	**-26.119****
	(7.856)	(9.078)	(10.797)	(10.764)
			Serbian[b]	
post	0.039	0.046	0.079	0.027
	(0.060)	(0.048)	(0.055)	(0.050)
treatment*post	0.044	0.012	-0.035	0.058
	(0.069)	(0.066)	(0.075)	(0.080)
			Mathematics[b]	
post	0.051	0.065	0.096	0.041
	(0.069)	(0.062)	(0.080)	(0.056)
treatment*post	0.046	0.030	0.015	0.053
	(0.081)	(0.077)	(0.091)	(0.085)
Max no. observations	4167	4039	1951	2088
School FE	No	Yes	Yes	Yes

This table reports the effect of the program on dropouts, absences and Serbian and mathematics. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%. Controls included are school size (squared), number of Roma in school (squared), percentage of Roma per class, class size (squared), female (=1), age (squared), and migrant (=1).

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

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

Table 16: Cohort regression

Back

	All			Girls			Boys		
	Early Enrollees	Late Enrollees	All	Early Enrollees	Late Enrollees	All	Early Enrollees	Late Enrollees	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Dropout[a]				
young*post	0.014	0.079***	0.080***	0.031	0.071***	0.077***	-0.003	0.087***	0.083***
	(0.013)	(0.019)	(0.018)	(0.020)	(0.019)	(0.018)	(0.012)	(0.028)	(0.028)
young*post*treatment			-0.066***			**-0.047***			**-0.087*****
			(0.022)			(0.027)			(0.029)
					Absences				
young*post	23.579*	57.493**	54.639**	40.707**	67.618**	67.603**	9.740	52.423**	49.480**
	(11.537)	(21.263)	(22.235)	(16.000)	(28.938)	(26.469)	(14.055)	(22.036)	(24.287)
young*post*treatment			-31.867			-28.524			-40.337
			(24.945)			(30.772)			(27.852)
					Serbian[b]				
young*post	0.079	-0.255**	-0.300***	0.052	-0.284	-0.352**	0.101	-0.295*	-0.328**
	(0.102)	(0.104)	(0.101)	(0.113)	(0.173)	(0.170)	(0.161)	(0.154)	(0.139)
young*post*treatment			0.382**			**0.416***			**0.423***
			(0.149)			(0.215)			(0.212)
					Mathematics[b]				
young*post	0.082	-0.241	-0.291*	-0.019	-0.344	-0.426*	0.180	-0.205	-0.239*
	(0.116)	(0.143)	(0.152)	(0.132)	(0.234)	(0.241)	(0.155)	(0.142)	(0.135)
young*post*treatment			0.381*			0.428			**0.412***
			(0.194)			(0.279)			(0.207)
Max no. observations	2438	1601	4039	1200	751	1951	1238	850	2088
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

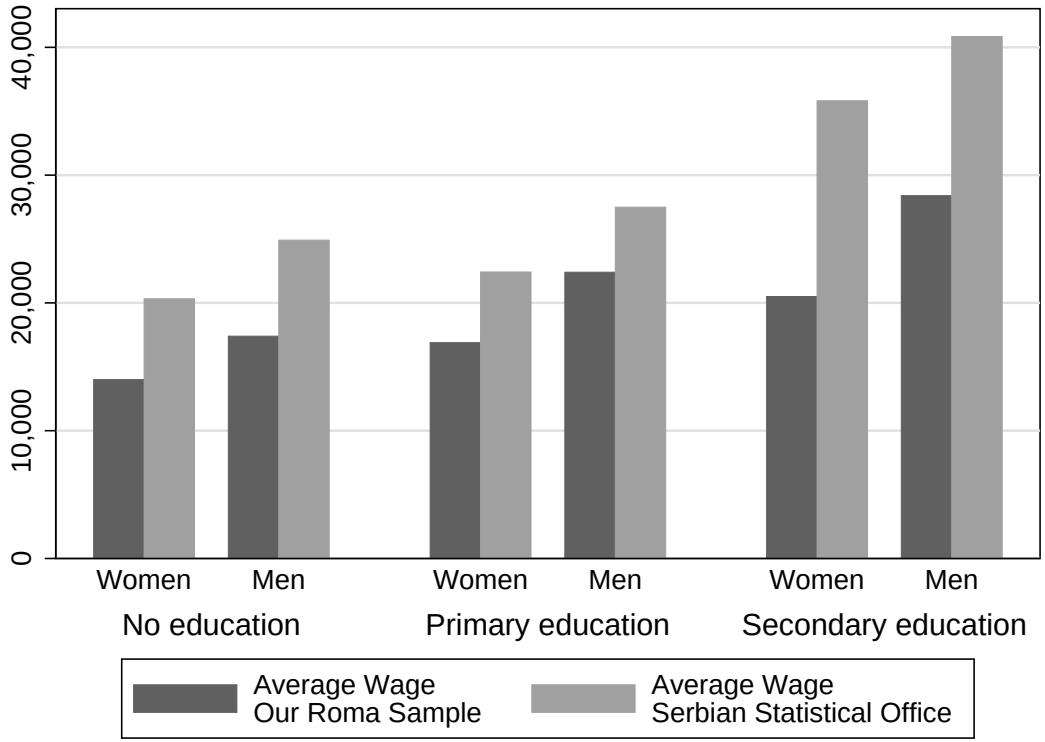
This table reports the effect of the program on dropouts, absences, Serbian and mathematics, using the cohort comparison methodology. Robust standard errors corrected for clustering at the school level are reported in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%. Controls included are school size (squared), number of Roma in school (squared), percentage of Roma per class, class size (squared), female (=1), age (squared), and migrant (=1).

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

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

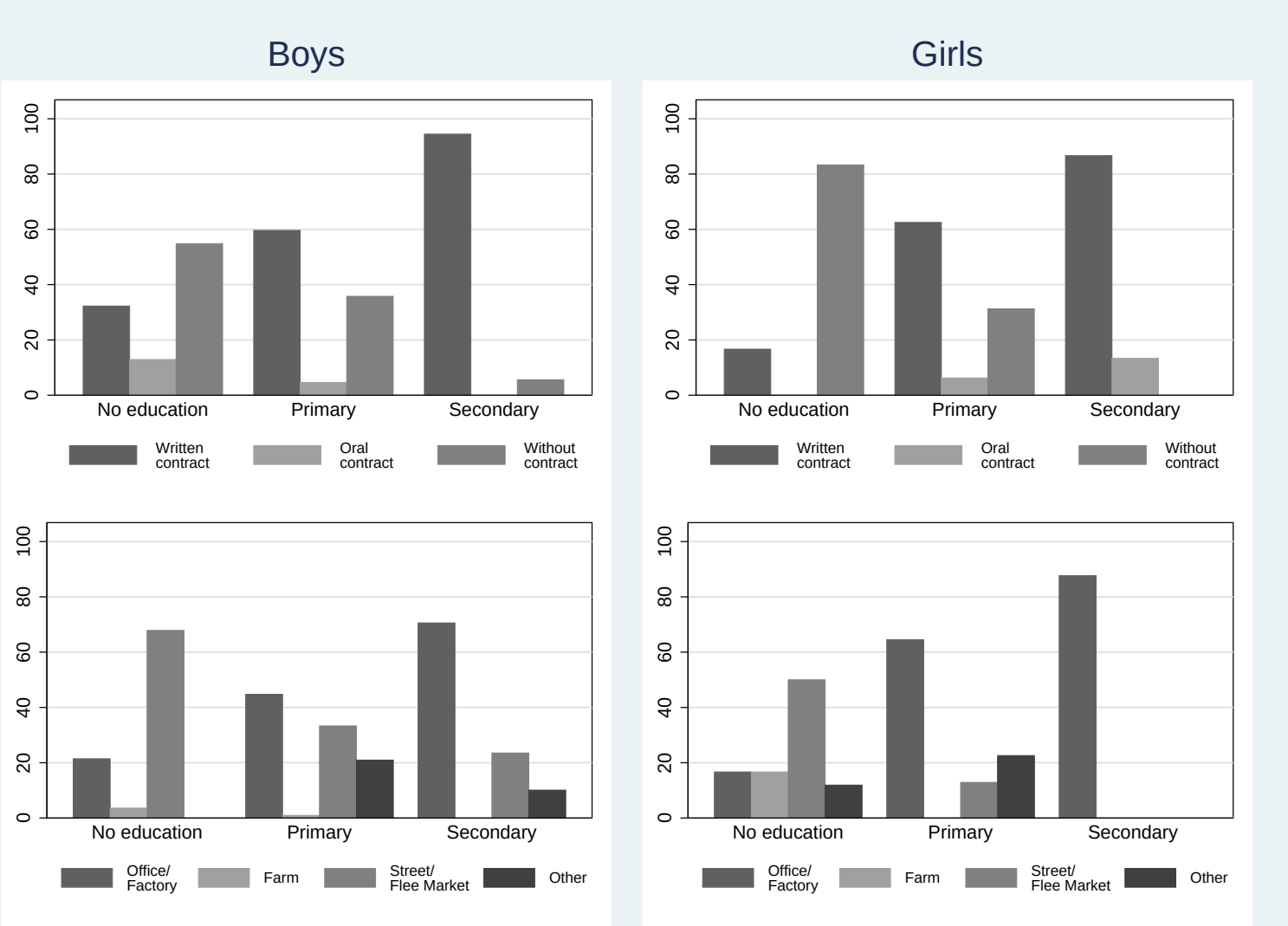
Back

Figure 4: Comparison of real returns to education for Roma and Non-Roma in Belgrade



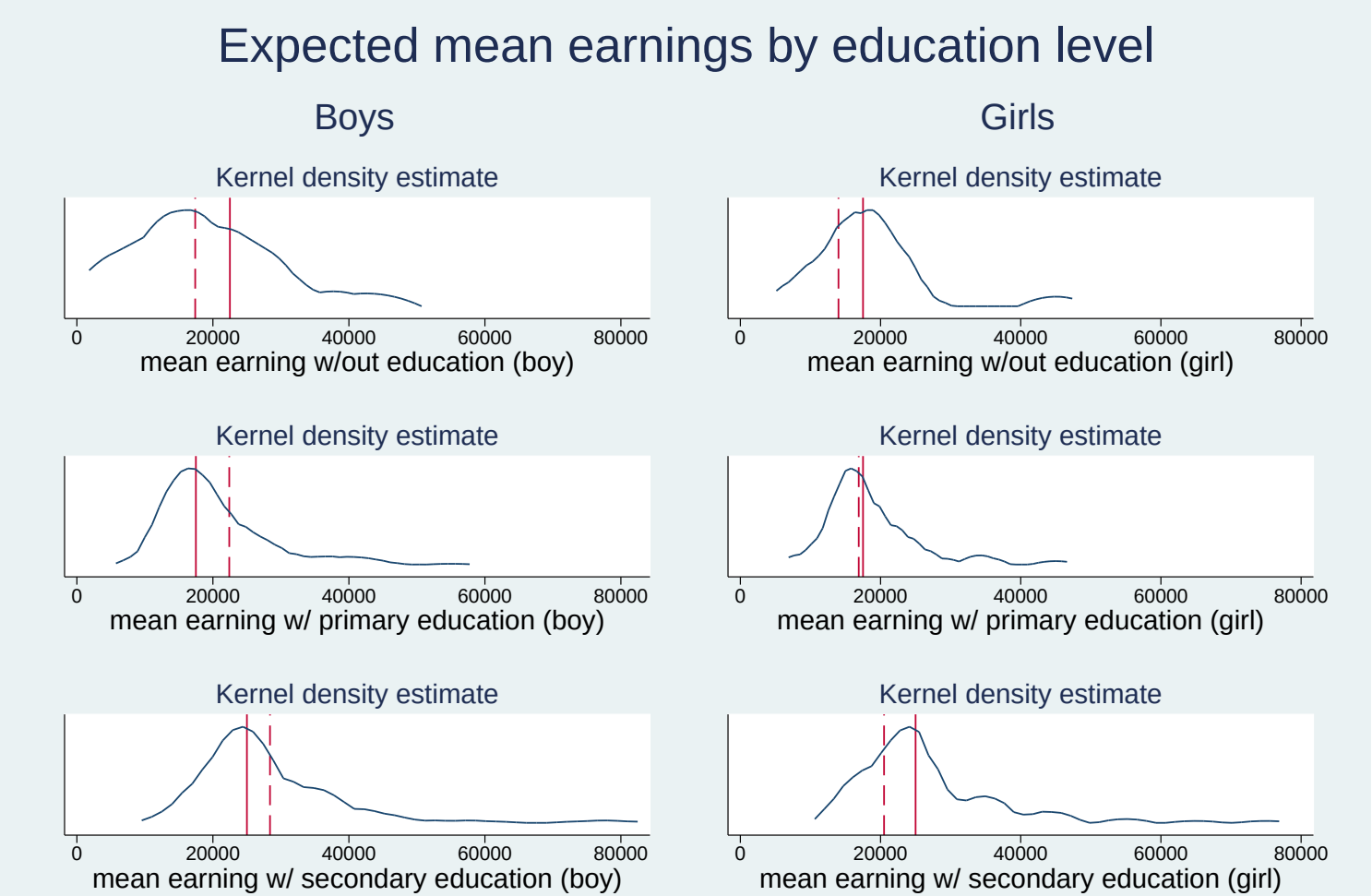
Back

Figure 5: Jobs characteristics by education levels



Back

Figure 6: Expected and actual earnings



The vertical dashed (solid) lines mark the estimated (expected) average returns in dinars.

Table 17: Means of control variables in treated and control households

Variables at the household level	All	Treatment	Control	Difference	s.e.
Wealth	0.08	-0.14	0.22	-0.36	(0.39)
Monthly Total income (in dinars)	28949.47	28224.39	29453.33	-1228.94	(2574.97)
Informal (=1)	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)	0.21	0.28	0.16	0.12	(0.07)
No schooling/Unfinished primary school (=1)	**0.07**	**0.07**	**0.07**	**0.00**	**(0.03)**
Finished primary school (=1)	**0.74**	**0.69**	**0.76**	**-0.07**	**(0.05)**
Finished secondary school (=1)	**0.19**	**0.23**	**0.16**	**0.07**	**(0.05)**
Muslim (=1)	0.71	0.57	0.80	-0.23***	(0.10)
Number of children under 5	0.72	0.75	0.70	0.05	(0.10)
Number of female children between 6 and 18	1.65	1.73	1.59	0.14	(0.13)
Number of male children between 6 and 18	1.75	1.80	1.80	0.10	(0.12)
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					
Male (=1)	0.52	0.50	0.54	-0.04	(0.04)
Age of child	9.89	10.11	9.74	0.37	(0.28)
Rank among siblings	2.20	2.33	2.11	0.22**	(0.10)
Mark in Mathematics	2.77	2.86	2.70	0.16	(0.11)
Mark in Serbian	2.85	2.94	2.79	0.15	(0.11)
max no. observations	673	280	393		

Back

Table 18: *Early Enrollees*: Means of controls for treated and untreated households

Variables at the household level	Assistant	No assistant	Difference	s.d.
Wealth	-0.38	0.13	-0.51	(0.46)
Total income (in dinars)	27905	29052	-1147	(3270)
Informal (=1)	0.39	0.25	0.14*	(0.081)
Urban (=1)	0.66	0.29	0.37***	(0.07)
Only Roma in neighborhood (=1)	0.35	0.21	0.14	(0.09)
No schooling/unfinished primary school	0.11	0.03	0.08	(0.06)
Finished primary school (=1)	0.63	0.75	-0.12*	(0.07)
Finished secondary school (=1)	0.25	0.21	0.03	(0.07)
Muslim (=1)	0.68	0.47	0.21**	(0.09)
Number of children under 5	0.87	0.66	0.21	(0.16)
Number of female children between 6 and 18	1.82	1.67	0.15	(0.16)
Number of male children between 6 and 18	1.86	1.75	0.11	(0.18)
Number of adults	2.53	2.38	0.15	(0.17)
max no. observations	65	56		
Variables at the individual level - children				
Male (=1)	0.5	0.5	0	(0.05)
Age of child	9.73	10.44	-0.69**	(0.31)
Rank among siblings	2.34	2.33	0.01	(0.17)
Mark in Mathematics	2.91	2.81	0.10	(0.14)
Mark in Serbian	3.03	2.87	0.16	(0.14)
max no. observations	148	130		

Back

Table 19: OLS and IV

Back

	OLS		ITT		IV	
Gender	Boys	Girls	Boys	Girls	Boys	Girls
	(1)	(2)	(3)	(4)	(5)	(6)
Probability to find a job						
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	yes	yes	yes	yes	yes	yes
Expected log earnings						
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	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	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%.

Table 20: First-stage

Back

Gender	Boys	Girls
Max. level of education	Secondary School	Secondary school
	(1)	(2)
First stage - Being helped by the assistant		
treatment	0.472***	0.465***
	(0.058)	(0.073)
controls[a]	yes	yes
No. observations	232	221
F-statistic on treatment	84.14	47.21

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

[a] The coefficients are estimated both with the controls used with the first two outcomes and with the third one. The estimates reported are obtained by using the outcome "secondary school as the highest expected level of education". Therefore, 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 21: Early Enrollees vs. Late Enrollees

▸ Pooled ▸ Back

Gender	Boys		Girls	
Max. education	Primary school	Secondary School	Primary school	Secondary school
	(1)	(2)	(3)	(4)
Probability to find a job				
treatment	-0.068	0.012	-0.114	-0.037
	(0.078)	(0.055)	(0.078)	(0.054)
controls[a]	yes	yes	yes	yes
No. observations	276	276	268	267
Expected log earnings				
treatment	**0.107***	**0.092***	0.123	**0.100***
	(0.059)	(0.050)	(0.085)	(0.056)
controls[a]	yes	yes	yes	yes
No. observations	119	224	98	216
Secondary school as the highest expected education level				
treatment		**0.123***		0.003
		(0.066)		(0.086)
controls[b]		yes		yes
No. observations		232		221

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] Additional control variables are age of child, age of child squared, rank among siblings, demeaned mark in Mathematics and demeaned mark in Serbian.

Table 22: Pooled sample

Back

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

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

- **Maths score**
  1. Please tell me how much is 5+4?
     (correct; not correct)
  2. Please tell me how much is 23+12?
     (correct; not correct)

- **Serbs score**
  1. **Able to read**: Could you please read me the letters, the word and the sentence on this card?
     (doesn't know letters; recognises only letters; knows to read the word, but can't read; knows to read the sentence)
  2. **Able to write**: Please write the following sentence.
     (doesn't know to write; wrote the sentence with mistakes; wrote the sentence correctly)

▸ Back

Table 23: Pupils' expectations

Gender	Boys	Girls
	(1)	(2)
Dep.var.:	**Secondary school**	
treatment	0.010	-0.039
	(0.027)	(0.057)
controls	yes	yes
No. observations	135	126
assistant	0.021	-0.076
	(0.053)	(0.105)
controls	yes	yes
No. observations	135	126

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

Back

Table 24: Heterogeneous effects

Back

Gender	Boys	Girls	Boys	Girls	Boys		Girls	
Max. level of education				Secondary school				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep.var.:	**Probability to find a job**		**Expected log earnings**		**Highest expected education level**			
treatment	-0.031	-0.045	0.013	0.001	**0.214****	0.093	0.163	-0.159
	(0.068)	(0.076)	(0.085)	(0.122)	(0.094)	(0.093)	(0.143)	(0.116)
muslim	-0.008	0.022	0.077	0.065	-0.134		-0.093	
	(0.067)	(0.078)	(0.082)	(0.116)	(0.093)		(0.131)	
treatment*muslim	0.066	0.012	0.121	0.150	-0.146		-0.223	
	(0.098)	(0.107)	(0.109)	(0.137)	(0.138)		(0.158)	
young						0.055		**-0.378****
						(0.116)		(0.144)
treatment*young						0.056		**0.279****
						(0.108)		(0.133)
controls	yes	yes	yes	yes	yes	yes	yes	yes
No. observations	276	268	224	216	232	221	221	221

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