ROMA INCLUSION AND IMPACT EVALUATION OF TWO MAINSTREAM EU-FUNDED ACTIVE LABOUR MARKET PROGRAMMES

Anna Adamecz
Bence Czafit
Katalin Bördős
Edit Nagy
Petra Lévay
Ágota Scharle

September 2013

CONTENTS

1. EXECUTIVE SUMMARY

The Budapest Institute evaluated the effects of two mainstream EU-funded active labour market programmes (ALMPs) on Roma inclusion and employment of its uneducated participants. The evaluation was commissioned by the Open Society Foundations (OSF), within the Making the Most of EU Funds for Roma (MtM) initiative. Considering their intended target groups, both selected ALMPs could have covered a substantial share of Roma jobseekers. The evaluation sought to answer two questions: how effectively the programmes actually reached those Roma people who belonged to their target groups; and whether participation in the programmes increased the probability of finding a job.

The *Improvement of employability of the disadvantaged* (SROP 1.1.2)[1] programme targeted various subgroups within the registered unemployed: the uneducated, school leavers, people aged above 50, the long-term unemployed, and those at risk of long-term unemployment. The Roma were not explicitly differentiated as a primary target group; however, they were prioritised within the target groups. The programme provided a personalised combination of subsidies and services, such as labour market counselling, mentoring, vocational training and wage subsidies.

The *One step ahead!* (HRDOP 3.5.3 & SROP 2.1.1) programmes offered general or vocational training to participants who had primary education or less, and in exceptional cases, vocation retraining to those with a vocation considered outdated. Besides training, participants also received a cash transfer during the programme. This programme did not target the Roma either, and contrary to SROP 1.1.2, it did not even prioritise them within the target groups.

The programmes reached only a small percentage of their target groups. Participants are positively selected in terms of their labour market potential: on average they are younger and higher educated than those who were eligible but did not participate in the programmes. However, we cannot tell the reason behind this phenomenon: it may be both self-selection and cream-skimming.

---

[1] SROP stands for Social Renewal Operational Programme (Társadalmi Megújulás Operatív Program, TÁMOP)

Table 1: Details of the two selected ALMPs

	Improvement of employability of the disadvantaged (SROP 1.1.2)	One step ahead! (HRDOP 3.5.3 and SROP 2.1.1)
Programme entry period	2008-2011	2006-2010
Number of participants in the NLO data*	57 894	23 088
Number of Roma participants in the PPR	3 797	n/a
Number of Roma participants – BI estimate**	4 636	2 899
Total budget, million HUF	53 041	18 376
Costs per participant, HUF	916 174	795 911
Found a job+	81%	63%
Costs per participants who found a job, HUF	1 131 079	1 263 35
Comparable cost of public works on 2013 prices, months[2]	13.8	15.4

Notes: *Programme entries before Dec 31, 2010. **We calculated the share of Roma population by settlement, summed these ratios, and multiplied them with a supposed bias of the Census data with respect to the Roma surveys. See in detail in Section 4.2. +The share of those who found a job during the programme or within 6 months afterwards, as a % of the total number of participants. ++ Number of months spent on public works that would cost the same amount per person. 1 HUF roughly equals 300 EUR.

Sources: Official documents, own calculations based on NLO data and Csite et al (2013) on budgets.

We examined the targeting and effects of the programmes using an individual-level dataset consisting of the unemployment and employment history of the participants and comparable control groups. In particular, the programme participation databases and the unemployment registry of the National Labour Office (NLO)[3] linked with the administrative reports of newly hired employees[4] formed the base of our dataset. However, neither the NLO nor the SHLD data contain information about ethnic origin. We obtained ethnicity data from two sources. First, in the case of the *Improvement of employability of the disadvantaged* (SROP 1.1.2) programme, the official Programme Progress Reports (PPR)[5] included some aggregate data about the proportion and performance of Roma participants. Unfortunately, the PPR of *One step ahead!* (HRDOP 3.5.3 & SROP 2.1.1) programme did not include such data. Second, we used the settlement-level ethnicity data of the 2011 Census to create an individual-level proxy variable showing the probability that the individual is Roma based on the share of the Roma population in their home settlement. This indicator has several shortcomings: it assigns the same probability to each

---

[2] Per an employed person.

[3] In Hungarian: *Nemzeti Munkaügyi Hivatal (NMH)*

[4] These reports are to be sent by employers to the tax authority and form the basis of the Standardized Hungarian Labour Dataset (Egységes Munkaügyi Adattár) (SHLD).

[5] In Hungarian: *Program Előrehaladási Jelentések (PEJ)*

resident of a settlement and it underestimates the number of the Roma.[6] However, we are not aware of (and had no access to) a better source of ethnicity data.

According to the PPR, the *Improvement of employability of the disadvantaged* (SROP 1.1.2) programme reached 3,797 Roma individuals, which is very few compared to the estimated number of 2-300 thousand potential Roma participants, or to the total number of programme participants (57 894 persons). The programme was least successful in reaching the Roma in Northern Hungary, in Northern Great Plain and in Southern Transdanubia, i.e. in the regions where the number of Roma people is the highest, and where their labour market chances are the worst. Roma women are less likely to participate in the programme than Roma men; their relative position is the worst in settlements with the highest shares of Roma population. The PPR of the programme shows that Roma participants were just as likely to successfully complete their individual programme plans, and even more likely to complete successfully their training, than non-Roma participants, however, 180 days after the programme they were only half as likely to be employed (16 vs. 32%).

The analysis of NLO data on the *One step ahead!* (HRDOP 3.5.3 & SROP 2.1.1) programme revealed that significantly more people were reached by the programme in those settlements with a higher share of Roma in the population. Our data analysis suggests that this programme, due to the target group being the uneducated, might have covered relatively more Roma people than the *Improvement of employability of the disadvantaged* (SROP 1.1.2) programme, which targeted several other subgroups besides uneducated people.

Based on the Roma population of the settlements from where there were no participants in the two programmes, about 3-5% of the entire Hungarian Roma population were completely left out from the *Improvement of employability of the disadvantaged*, and about 16-17% from the *One step ahead!*. As the data suggests, if a programme is bigger in size, it can reach not simply more people, but smaller settlements as well. This is important, because 16% of the Hungarian Roma population live in small villages with less than 1,000 inhabitants. However, in spite of the fact that the *Improvement of employability of the disadvantaged* programme had participants from a

---

[6] Level of education could be added to improve our estimate of Roma origin, but this would not help in the present analysis as it is focused on uneducated jobseekers.

large number of settlements where altogether about 95-97% of the entire Hungarian Roma population live, the share of its Roma participants was only about 1.3%.

We estimated the causal impact of the programmes by matching a comparable control pair to each uneducated participant based on their observable characteristics (demographics and employment history). The effect of the programmes on the probability of employment is large, positive and significant. Because of the lack of individual-level ethnicity data, we cannot measure the effect on the Roma separately. The uneducated participants of the *Improvement of employability of the disadvantaged* (SROP 1.1.2) found a job during the programme or within half a year afterwards with a 44 percentage points higher probability than their comparable control pairs. However, we probably overestimate the effect of the programme, among others due to the fact that control persons may be more likely to work in the grey economy, which we cannot observe. The effect of the programme on the probability of exit to employment is substantial in case of those participants who did not receive wage or wage cost subsidy and the long-term unemployed as well.

The *One step ahead!* (SROP 2.1.1) programme increased the probability that its uneducated participants find a job by 34– 40 % points (seeTable 2). 57-71% of the participants entered employment at least once during the programme period or within 6 months after completing the programme. The positive impact of the programme is again very similar for the long term unemployed beneficiaries as well.

The budget of the two programmes exceeded 70 billion Forints (see Table 1) over 4 - 5 years. In per capita terms, this amount is equal to 900 thousand HUF per participant in the *Improvement of employability of the disadvantaged* (SROP 1.1.2) programme and 796 thousand HUF per person in the *One step ahead!* (HRDOP 3.5.3 & SROP 2.1.1) programmes. If we consider only those participants who found a job during the programme or within half a year afterwards, per capita costs amount to 1,131 and 1,263 thousand Forints per person, respectively. This per capita budget would be enough to finance 14-15 months of public works per person, calculated at 2013 nominal prices.

Table 2: The effect of the programmes (uneducated men)

Outcome variables	Improvement of employability of the disadvantaged (SROP 1.1.2)		One step ahead! (SROP 2.1.1)	
	% of programme participants	% point effect*	% of programme participants	% point effect*
Exit to employment during the programme or within 6 months afterwards	76	44	71	41
Exit to employment within 6 months after participating in the programme	40	29	12	10
Exit to employment anytime during the observation period (until Oct 2012)	91	49	85	53
No re-entering into unemployment within 6 months after the programme	49	6	67	40
No re-entering into unemployment anytime after the programme during the observation period	75	22	61	20

Source: BI estimates using NLO and 2011 Census data. *Estimated programme effect based on counterfactual impact evaluation. It shows how the programme affected the probability of finding employment/not re-entering unemployment. For example, the SROP 1.1.2 programme increased the probability of finding a job during the programme or within 6 months afterwards with 44%points comparing to a theoretical case in which participants had not participated in the programme.

We compare these results to public works, which is the typical alternative to personalised ALMP for uneducated workers. However, earlier empirical evidence clearly shows that the effect of public works on reemployment in the open labour market is very small, or in some cases even negative,[7] while the programmes evaluated here increased the probability of employment by more than 40%points. This implies that personalised ALMPs can contribute to increasing employment. Considering their direct as well as indirect effects (health benefits, etc.), they can be cost-efficient on the long run; however, it's not straightforward to estimate their cost recovery period.

We conclude that both programmes significantly increased the labour market potential of the participants. This result is especially striking because we evaluated the impact of the programmes on the most disadvantaged jobseekers, the uneducated only. As we documented, the programmes had a positive effect even without wage or wage-cost subsidy and in case of the long term unemployed as well. However, regarding take up of the Roma, the targeting of these programmes could be improved. To learn more about the effects and effectiveness of such programmes in case of Roma people in particular, it would be important to collect and make

---

[7] See overview in Scharle (2011).

available individual-level ethnicity data for research purposes about labour market programme participants and the registered unemployed in general. Alternatively, new methods should be developed to estimate ethnicity status using already available data sources. Without individual level ethnicity data it is impossible to evaluate how effectively the programmes reached Roma people.

Based on these results we are convinced that both types of ALMP's should be continued in the next programming period. However, in order to increase access by Roma participants, we recommend restricting the target groups exclusively to those with at most elementary education. However, we do not suggest the introduction of regional or other quotas regarding the participants of the programmes as this would not necessarily improve targeting at the individual level. In the case of training programs such as the *One step ahead!* we recommend employing strict quality assurance measures and teachers/trainers specialized in adult education. The use of educational materials created specifically for adult learners is also crucial. Sensitive scheduling of the trainings is of utmost importance as well: in high seasons of casual work, usually in the summer, potential participants may be less likely to enter and complete training programmes. Lastly, resources should be allocated across regions based on the number of uneducated jobseekers rather than the number of jobseekers, especially if the budget of such ALMPs is reduced in the next programming period.

2. WHY IS IT IMPORTANT TO MEASURE THE EFFECT OF EMPLOYMENT POLICY PROGRAMMES?

The labour market situation of the Roma in Hungary is extremely poor since the regime change in 1989. Only about 40 per cent of working-age Roma are employed, 10 per cent of whom work temporary or in public works. They are also often involved in unregistered jobs (Kemény-Janky 2003). The ratio of the poor is about four times higher among the Roma compared to the non-Roma; approximately third to half of those living in extreme poverty are Roma people (Ladányi-Szelényi 2002, Havas 2008). Geographic and social mobility of Roma people is very low; most of them have been living without a permanent income source for 15-20 years at the mercy of local politics and legislation. Life expectancy of the Roma is about 10 years lower than the country average due to their bad health conditions, which is further exacerbated by their lower

education levels, worse living conditions and the lack of access to provisions (Janky 2004, Kemény-Janky 2003, Kertesi 2005 (a), Vokó et al 2006).

Government programmes aiming at improving the employment opportunities of Roma people have failed in reaching their goals most of the time. The financing of non-profit organizations and government programmes is disorganized and is always changing. These programmes are usually not followed by any monitoring activity or impact assessment (ÁSZ 2008). Thus, the impact evaluation of mainstream employment programmes explicitly or implicitly targeting Roma workers is beneficial in two ways: it helps uncovering the effect of bigger programmes with broader targeting, and it helps the institutionalization of evidence-based policy making.

3. THE DATASET

The analysis has been conducted using three data sources. General aggregate data about the programmes has been collected from the electronic version of official Program Progress Reports (PPR)[8]. The individual-level dataset of programme participants and controls is comprised of data from the National Labour Office (NLO)[9]. The NLO dataset consists of three parts:

1. Detailed data about participants entering the programmes before Dec 31, 2010;

2. Data from the unemployment registry about individuals who were registered as unemployed for at least one day during the entering period of the programmes;

3. Employment data from the Standardized Hungarian Labour Dataset (*SHLD*)[10] until the end of October, 2012.

---

[8] In Hungarian: Program Előrehaladási Jelentések, PEJ

[9] In Hungarian: Nemzeti Munkaügyi Hivatal, NMH

[10] In Hungarian: Egységes Magyar Munkaügyi Adatbázis, EMMA

Table 3. The details of the programmes

Programme	The official duration of the programme	Duration in the NLO database	No. of participants (NLO data)
Improvement of employability of the disadvantaged	Jan 1, 2008 – Apr 30, 2011	Programme entry: Jan 1, 2008 – Dec 31, 2010 Programme exit: Jun 15, 2008 – Dec 31, 2011	57 894
One step ahead![11]	HRDOP 3.5.3: Jan 1, 2006 – Dec 31, 2008; SROP 2.1.1: Sept 1, 2007 – Nov 4, 2009; SROP 2.1.1/B: Dec 15, 2009 – Dec 30, 2010	Programme entry: Jan 9, 2006 – Dec 11, 2010 Programme exit: N/A.	23 088
Control group pool	-	Registered unemployed for at least a day during the entering period of the programmes.	

Source: NLO and BI calculations based on NLO data.

The following information is included in the unemployment registry, and thus, in our database: entry and exit date of unemployment; reason for exit or pause (new job, public works, retirement, etc.); subsidies received (regular social allowance, different kinds of employment substitute support) and their duration; disability; personal information (gender, date of birth, postal code, educational attainment). We had no data at our disposal about jobs offered by unemployment offices, desired job or reservation wage.

The SHLD dataset includes entry and exit of employment, type of employment and employer code; further information, such as wage or contributions, is not available. The dataset is comprised by the National Tax and Customs Administration of Hungary (NTCA)[12]; thus it includes registered workforce only.

As we have no access to individual-level data about ethnicity of participants or controls, we expanded the database with settlement-level ethnicity data from the 2011 Census.[13] Publicly available Census data include the number of Roma/non-Roma population at each settlements (in case of Budapest: districts) without further information about its distribution with respect to

---

[11] More information about programme phases on page 42.

[12] In Hungarian: *Nemzeti Adó- és Vámhivatal, NAV*

[13] Source: http://www.ksh.hu/nepszamlalas.

gender, age or other variables. Besides that, we also use labour market variables of the Census database (number of working-age population and employment status by gender).

Using the Census data instead of individual ethnicity indicators has several drawbacks. First of all, proportion of Roma in the population is underestimated in Census data, and the size of this bias varies over settlements. (Messing 2011) Since the regime change in 1989, three "Roma-surveys" have been conducted with the aim of revealing the situation of Roma people. In the first and the second surveys - published in 1993 and 2003 – participants were identified as Roma based on the opinion of their non-Roma environment. In 2003, Kemény and Janky (2003) estimated the number of Roma people in Hungary as 600,000 based on a 1% representative sample. In the third survey, published in 2010, the two surveying methods, self-assessment vs. public opinion, were compared (Messing, 2011). Censuses use self-assessment method that underreports the number of the Roma: the 2001 Census reported 190,000, the 2011 Census reported 315,000 Roma people, which is still about half of the supposedly realistic number of 600,000.

Both self- and public assessment might lead to a bias. In case of census data, Roma people might not or falsely answer questions about nationality in the fear of conceived or real consequences. On the other hand, "Roma-surveys" might be biased by the preconceptions and beliefs of respondents about Roma people. In our study, public assessment would seem to be more relevant, because labour market discrimination is more related to the judgment of the environment rather than self-assessment. However, such data is not available on settlement level.

It complicates the problem that the downward bias of the number of Roma in the Census data varies by locations and regions. The uneven nature of the bias might have been increased by the fact that before the 2011 Census in some settlements civil organizations intensively campaigned in favour of higher share of Roma people revealing their nationality in the hope for more accurate data. In these settlements Roma census takers were also hired, because Roma people are more likely to admit their origins if they are asked by a person with Roma background. However, there is no available data about which settlements were involved in this campaign, that is, which

settlements' Census data could be closer to reality. In the absence of further information about regional differences we use Census data in the form as it was published.[14]

4. THE *"IMPROVEMENT OF EMPLOYABILITY OF THE DISADVANTAGED"* PROGRAMME

The *Improvement of employability of the disadvantaged* (SROP 1.1.2)[15] programme was introduced to foster reemployment of the most disadvantaged groups of the unemployed. The programme targeted various subgroups within the registered unemployed: the uneducated; school leavers; people aged above 50; mothers with young children; the long-term unemployed and those at risk of long-term unemployment. The Roma were not explicitly differentiated as a primary target group; however, they were prioritised within the target groups. The programme provided a personalised combination of subsidies and services, such as labour market counselling, mentoring, vocational training and wage subsidies.

4.1. Selection and take up of the uneducated

Based on the NLO database, 57,894 people entered programme until the end of 2010. As mentioned before, according to the eligibility criteria the following groups could have participated in the programme:

- the uneducated (having at most elementary education);
- school leavers;
- people aged above 50;
- parents with young children;
- the long-term unemployed; and
- those at risk of long-term unemployment.

---

[14] In case this research could continue, based on the name and address data of NLO unemployment registry, which is available to the NLO only, and the method of Váradi (2012), the probability of being Roma might be estimated for each person in the register. Although this method has already been used in Hungary on a small scale, it cannot be seen at this point whether the NLO would agree to it.

[15] SROP stands for Social Renewal Operational Programme (in Hungarian: Társadalmi Megújulás Operatív Program, TÁMOP).

Targeting is analyzed by comparing participants to eligible non-participants (potential participants).[16] If actual participants differ from potential participants in some of their characteristics, we can conclude that participants were selected (either through self-selection or through the selection process of unemployment offices) based on those measures.

Table 4. Educational attainment of actual and potential participants of the SROP 1.1.2 programme (men and women together)

Level of education	Potential participants		SROP 1.1.2 participant		Total	
	No.	%	No.	%	No.	%
Below elementary school	74 161	5.7	345	0.6	74 506	5.5
Elementary school	663 689	51.0	14 262	24.6	677 951	49.9
Vocational school	265 720	20.4	14 544	25.1	285 264	21.0
Secondary school	98 572	7.6	8 810	15.2	107 382	7.9
Secondary vocational school	122 580	9.4	11 519	19.9	134 099	9.9
Technical school	35 476	2.7	2 619	4.5	38 095	2.8
College	27 546	2.1	4 380	7.6	31 926	2.3
University	13 244	1.0	1 415	2.4	14 659	1.1
Missing data	431	0.0	0	0.0	431	0.0
Total	1 301 419	100.0	57 894	100.0	1 359 313	100.0

Source: BI calculation from NLO data

On average, 80 per cent of the working-age Roma population has an elementary school degree at most (Kemény et al 2004, Kertesi 2005). The share of the uneducated among participants is about 25%, while the same ratio is well above 57% in the group of potential participants (see Table 5). Thus, programme participants are more educated than the comparable unemployed.

Among uneducated participants of the programme, the ratio of those under 25 is significantly higher (32 vs. 19%) and that of those above 50 is significantly lower (12 vs. 23%) than in the group of potential participants (see Table 5). Programme participants are not just more educated, but uneducated participants are also younger on average than the comparison group.

---

[16] We define potential participants as individuals being registered in the unemployment registry for at least one day under the entering period of the programme and being either aged under 25 or above 50, or uneducated, or long-term unemployed. We cannot identify parents with young children and people at risk of long-term unemployment based on the unemployment registry, so we do not include them in the control group.

Table 5. The age distribution of uneducated SROP 1.1.2 participants

	Participant (%)	Control (%)
Below 25	31.81	18.71
Aged between 25-34	25.59	23.84
Aged between 35-49	30.47	34.15
Above 50	12.12	23.29
Sum	100.00	100.00
Share of the uneducated	25.00	57.00

Source: BI calculation based on NLO data

N=14,601 uneducated participants and 737,699 uneducated potential participants.

Take up of the uneducated varies by regions and settlement size. Table 6 shows the take up rate by the share of the Roma in the population of home settlement by region, whereas Table 7 shows the same by settlement size. Average take up is below 2%; from the 737,699 potential participants 14,601 participated in the programme. With respect to the share of Roma in the population, take up is higher in settlements with above average, but still moderate shares of the Roma (3-9%), but there is no positive correlation between take up and share of the Roma. Theoretically, as the share of the uneducated and unemployed is higher among the Roma, if coverage of the Roma had been efficient, take up should be higher in those settlements where the share of the Roma is higher.[17]

---

[17] We elaborate this argument further in Section 4.3.1.1 Correlation between the share the Roma and the No. of participants

Table 6. Take up of the uneducated in SROP 1.1.2 by regions (%)

Share of Roma in the population	Regions							Total
	Northern Great Plain	Southern Great Plain	Central Hungary	Southern Trans-danubia	Northern Hungary	Central Transdanubia	Western Trans-danubia	
0%	2.22	1.28	1.74	0.73	1.67	1.68	1.58	1.47
0-1%	1.82	1.88	1.40	1.58	2.74	2.05	2.47	1.90
1-2%	2.66	2.34	1.44	1.94	2.83	2.05	2.75	2.02
2-3%	1.90	1.73	1.44	2.30	2.65	1.90	2.04	1.94
3-4%	1.64	2.70	1.51	2.35	2.44	1.66	2.86	2.09
4-5%	1.72	2.71	1.57	2.60	3.31	2.40	2.15	2.19
5-6%	1.75	1.91	0.94	2.01	2.84	1.79	2.51	2.10
6-7%	2.20	2.75	1.94	1.93	3.15	1.32	2.71	2.45
7-8%	2.03	1.38	1.10	1.60	2.01	0.94	1.84	1.89
8-9%	2.49	1.61	1.49	1.80	2.89	1.87	0.45	2.29
9-10%	1.06	3.19	0.48	1.37	3.01	1.31	3.39	1.90
10-20%	1.25	1.28	0.98	2.09	2.37	1.76	2.84	1.84
20-30%	1.97		1.33	2.00	2.14	0.16	2.38	1.99
30-40%	1.45	1.05		1.72	1.93			1.76
40-50%	0.00			2.17	1.93		0.00	1.60
50-100%	0.55			1.71	5.06			1.81
Total	1.78	2.07	1.42	1.92	2.44	1.92	2.27	1.94
Regional distribution of Roma population	*26.10*	*8.20*	*13.00*	*13.80*	*29.60*	*5.10*	*4.20*	*100.00*

Source: BI calculation based on NLO and 2011 Census data

N=14,601 uneducated participants and 737,699 uneducated potential participants.

Note: take up rate is the No. of participants divided by the sum of the No. of potential and actual participants. 0 indicates no participants in that category; an empty cell indicates either no settlement or neither actual nor potential participants in that category. Highlighted cells indicate the categories where at least 2% of the Roma population lives. All highlighted cells together cover 48% of the Hungarian Roma population.

Table 7. Take-up of the uneducated in SROP 1.1.2 by settlement size (%)

Share of Roma in the population	Settlement categories								Total
	Budapest and county centres	Cities, above 10 000 inhabitants	Cities, 6 750-10 000 inhabitants	Villages, 5 000-6 750 inhabitants	Villages, 2 000-5 000 inhabitants	Villages, 1 000-2 000 inhabitants	Villages, 601-1 000 inhabitants	Villages, below 600 inhabitants	
0%			0.67	1.90	1.69	1.83	2.10	1.04	1.47
0-1%	2.00	2.09	1.68	1.48	1.69	1.72	2.21	2.01	1.90
1-2%	1.92	2.10	1.63	2.63	2.28	2.03	2.36	0.75	2.02
2-3%	2.17	1.87	1.87	1.83	2.01	1.65	1.87	1.61	1.94
3-4%	1.82	2.32	2.17	2.49	2.10	1.94	3.66	1.14	2.09
4-5%		1.77	2.81	3.14	1.57	2.84	1.83	3.42	2.19
5-6%	2.53	2.05	2.22	2.03	1.75	2.36	2.69	1.37	2.10
6-7%		2.28	2.77	2.30	2.68	2.23	2.65	2.32	2.45
7-8%		2.31	1.34	3.12	1.57	1.48	1.35	1.56	1.89
8-9%		2.52	6.39	1.88	1.74	2.12	2.33	1.71	2.29
9-10%		3.28	1.23	1.47	2.59	1.83	1.60	0.52	1.90
10-20%		1.79	0.97	1.66	1.58	2.10	2.67	1.78	1.84
20-30%				1.79	2.54	1.62	2.47	1.64	1.99
30-40%					1.59	2.21	1.48	1.65	1.76
40-50%						1.63	1.67	1.54	1.60
50-100%					0.48	1.80	0.94	2.80	1.81
Total	1.99	2.06	2.02	2.05	1.87	1.94	2.22	1.42	1.94
The distribution of the Roma population by settlement size (%)	*15.0*	*19.0*	*5.0*	*5.4*	*20.7*	*18.7*	*7.8*	*8.2*	*100.0*

Source: BI calculation from NLO and 2011 Census data.
N=14,601 uneducated participants and 737,699 uneducated potential participants.
Note: take up rate is the No. of participants divided by the sum of the number of potential and actual participants. 0 indicates no participants in that category; an empty cell indicates either no settlement or neither actual nor potential participants in that category. Highlighted cells indicate the categories where at least 2% of the Roma population lives. All highlighted cells together cover 50% of the Hungarian Roma population.

4.2. Roma inclusion and achievements based on PPR data

Based on data from Program Progress Reports (PPR), 6.54% of participants were Roma.[18] The information on whether participants identify themselves as being Roma stems from personal statements collected by the unemployment offices. It might be biased in two ways. First,

---

[18] The number of participants according to the PPR is 58,051. According to our individual level NLO dataset, this number is 57,894, because it includes only those entering the programme before Dec 31, 2010.

similarly to census inquiry, some people might have not wanted to reveal their ethnicity. Second, unemployment offices might have had interests in including less and reporting more Roma participants. For reaching targeted employment level of programme participants, offices were motivated to include those with better employment chances, which would lead to a higher share of non-Roma people. However, they might have convinced some non-Roma to state that they are Roma to fulfil other criteria. Although we cannot prove such practice from proper referable sources, it may have been in operation.

According to PPR data 3,797 Roma people took part in the programme. As we stated earlier, individual-level ethnicity data is not available in the NLO dataset. Thus, we calculated the proportion of Roma in the population in each settlement from 2011 Census data, and used these shares as a proxy instead of individual-level ethnicity information. For example, according to the 2011 Census data, in Acsa – a settlement in Pest county – 10 per cent of the population is Roma. Consequently, we assigned 10% as the probability of being Roma for each participant who lives in Acsa. The higher the proportion of Roma people in a settlement, the higher value this probability takes, thus the closer the proxy is to 100%. Obviously, using such proxy has many drawbacks, but we have no access to any better measures. If we sum this proxy for all participants, and we multiply it with the ratio of the Roma population according to the Roma surveys and the total number of the Roma according to the 2011 Census, considered as a measure of bias in the census data, we estimate that there are 4,636 Roma people among the participants.[19] This estimation is a little higher than the official 3,797 persons from the PPR. Furthermore, we not only overestimate the number of Roma participants, but we also smooth their spatial distribution.

The 6 per cent Roma participation rate from the PPR is much less than what is predicted by the literature for the proportion of Roma people among the unemployed. According to the "Roma-study" made in 2003, the unemployed working-age Roma population was approximately 300,000 in that year (calculation by Scharle (2011), based on Kemény és Janky (2004)). The number of almost 4,000 Roma participants is very small compared to that figure.

---

[19] The calculation is the following: 2,440*(600,000/315,000)=4,636. See more information about the bias in the The dataset chapter on page 4.

Table 8: Achievements of SROP 1.1.2 participants in the PPR

	All participants		Non-Roma		Roma	
	Capita	% of all participants	Capita	% of all non-Roma participants	Capita	% of all Roma participants
No. of participants involved in training	36 860	63.5%	33 786	62.3%	3 074	81.0%
No. of participants finishing training successfully	33 325	90.4%[1]	30 501	90.3%[2]	2 824	91.9%[3]
No. of participants finishing programme successfully	52 324	90.1%	48 890	90.1%	3 434	90.5%
No. of participants being employed on the 180th day after finishing programme	18 057	31.1%	17 461	32.2%	596	15.7%
Total No. of participants	58 051	100.0%	54 255	93.5%[4]	3 796	6.5%[4]

Source: PPR (NLO)
1: In percentage of the number of all training participants; 2: In percentage of the number of all non-Roma training participants; 3: In percentage of all Roma training participants; 4: In percentage of all participants.

Roma participants were more likely to participate in trainings (81%), than others (63.5%) (Table 8). 20 One reason behind this may be that while more than half of Roma participants (54%) is uneducated, this proportion among non-Roma participants is only 12% (see Table 9). Even though their educational background is different, success rates of training is similar among Roma and non-Roma participants: 90% of non-Roma and 92% of Roma training participants finished training successfully. Similar pattern features the programme as a whole: 90 per cent of both the Roma and the non-Roma completed their individual plans. Despite the fact that Roma participants were as successful as the non-Roma in finishing training, 180 days after the end of the programme they were only half as likely to be employed (16 vs. 32%).

---

[20]Similar data with respect to education levels is not available.

Table 9. Distribution of Roma and non-Roma participants by eligibility criteria in SROP 1.1.2

Eligibility criteria	All participants		Non-Roma		Roma	
	Capita	%	Capita	%	Capita	%
Uneducated	8 673	14.94	6 639	12.24	2 034	53.60
School leavers	13 330	22.96	12 663	23.34	667	17.58
Aged above	5 044	8.69	4 919	9.07	125	3.29
Parents with young children*	2 237	3.85	2 144	3.95	93	2.45
Special regional target group[21]	5 755	9.91	5 438	10.02	317	8.35
Lost their job after Sept 1, 2008.**	23 012	39.64	22 453	41.38	559	14.73
Total	58 051	100.00	54 256	100.00	3 795	100.00

Source: PPR (NLO)
*Received child home care allowance, child raising support, pregnancy - confinement benefit, child care fee, nursing fee within 12 months before getting involved in the programme.
**Due to the economic crisis the government allocated HUF 27 billion further source to the programme in March, 2009. Thus, after April 20, 2009 a new target group emerged besides the already existing 5 target groups: the people who became unemployed after September 1, 2008. (Source: PPR no. 1)

In Table 9, the number and proportion of Roma and non-Roma participants is summarized for each eligibility groups. While 40 per cent of non-Roma participants are from those who lost their jobs after Sept 1, 2008 – most probably due to the crisis –, the majority of Roma participants entered the programme because of being either uneducated (54 %) or school leaver (18%). Thus, most of non-Roma participants are educated or/and have work experience and are only temporarily unemployed due to an economic downturn. On the other hand, most of Roma participants either do not have any work experience or are uneducated having much worse labour market possibilities also in the long run.

---

[21] Special regional target group involves different targeted group of individuals in each region based on the local specialities and differences in labour market conditions.

4.3.Selection and Roma inclusion based on NLO and Census data

As detailed earlier, the NLO database includes information about participants who entered the programme before Dec 31, 2010. Not having individual-level data about whether a participant is Roma or not, we supplemented the dataset with settlement-level ethnicity data from the 2011 Census. Based on this data, 89 % of participants live in settlements where the share of Roma population is below 10 per cent, and this ratio is similar among potential participants, too.[22]

Table 10. Distribution of SROP 1.1.2 participants by share of Roma in the population in home settlement

Proportion of Roma in the population s	Total		Women		Men	
	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	6.09	6.12	6.00	6.28	6.18	5.93
0-1%	23.42	23.84	23.77	25.38	23.11	22.08
1-2%	20.06	21.68	20.39	22.47	19.76	20.76
2-3%	12.67	12.94	12.91	12.84	12.46	13.06
3-4%	7.56	8.18	7.51	8.30	7.60	8.03
4-5%	3.78	4.05	3.76	3.85	3.79	4.27
5-6%	3.66	3.71	3.66	3.66	3.66	3.77
6-7%	2.31	2.67	2.26	2.43	2.36	2.94
7-8%	3.12	2.68	3.04	2.47	3.20	2.93
8-9%	2.18	2.11	2.14	1.83	2.22	2.42
9-10%	1.24	1.19	1.20	1.12	1.28	1.27
Total above 10%	86.10	89.16	86.64	90.63	85.63	87.48
10-20%	8.85	7.15	8.54	6.34	9.13	8.06
20-30%	2.91	2.49	2.80	2.07	3.02	2.97
30-40%	1.19	0.74	1.14	0.61	1.24	0.90
40-50%	0.43	0.26	0.40	0.21	0.45	0.31
50-100%	0.51	0.20	0.49	0.13	0.54	0.28
Total	100	100	100	100	100	100

Source: BI calculation based on NLO and 2011 Census data.
Note: Equality holds at the upper values of categories

---

[22] As before, potential participants are selected from the unemployment registry as being either uneducated, or below 25, or above 50, or long-term unemployed. We could not identify two of the potentially eligible groups from the registry due to lack of data: those raising young children and being at risk of long-term unemployment.

The share of female participants living in settlements with a less than 10% share of Roma population is higher (91%), that of male participants (87%) is lower than the average. Therefore, aggregate data suggest that in settlements with a high share of Roma population women were more likely to be left out of the programme, than men.

The distribution of participants with respect to the share of Roma in the population of home settlement shows regional variation. (Find tables by regions in Appendix 2.1) The difference of distributions is the biggest in Northern-Hungary, Northern-Great-Plain and Southern-Transdanubia, where more actual participants live in settlements with relatively lower share of Roma in the population than potential participants. Thus, these are regions where the programme was less efficient in reaching Roma people.

4.3.1 Selection on settlement level

4.3.1.1 Correlation between the share the Roma and the No. of participants

In this chapter we examine whether the share of participants with respect to working-age non-employed population is correlated with the share of Roma in the population.[23] In the optimal scenario, we would need settlement-level data about the number of working-age non-employed Roma people who are eligible for the programme, and the number of those who participated from this group. We do not have access to such data, thus we construct the following settlement-level measure:

Share of participants = Number of participants / Working-age non-employed population

Number of participants is created by aggregating individual-level data of the NLO database to settlement level. Number of working-age non-employed population is formulated using the 2011 Census as follows:

Working-age non-employed population = Working-age population – Number of employed people.

---

[23] Settlement-level analysis means district-level analysis in case of the capital, Budapest.

Then, we compare our 'Share of participants' measure to the proportion of Roma in the population on settlement level. We assume that if the programme had reached Roma individuals of the target group efficiently, correlation between share of participants and share of Roma in the population would be positive. In other words, we expect relatively more participants to be included in the programme in settlements where relatively more Roma people live. This assumption is quite simplistic, yet it is solid as both the share of the unemployed and the uneducated is higher among the Roma (Kertesi, 2005). Thus, we claim that in case of efficient take up of Roma people, the share of participants from settlements with higher share of Roma population should be higher, because of the higher share of unemployed and uneducated people. Thereupon, the correlation in question should be positive.

Table 11. Distribution of SROP 1.1.2 participants from settlements with fewer than 600 participants by regions and share of Roma in the population (% of Hungarian Roma population)

Proportion of Roma in the population	Regions							
	Northern Great Plain	Southern Great Plain	Central Hungary	Southern Transdanubia	Northern Hungary	Central Transdanubia	Western Transdanubia	Total
0-1%	0.003	0.004		0.009	0.005	0.011	0.005	0.038
1-2%	0.002	0.010	0.003	0.036	0.018	0.008	0.023	0.099
2-3%	0.008	0.011	0.003	0.075	0.026	0.013	0.038	0.176
3-4%	0.007			0.055	0.015	0.007	0.042	0.125
4-5%		0.006		0.057	0.027	0.014	0.034	0.138
5-6%	0.024			0.106	0.016	0.010	0.017	0.173
6-7%	0.034			0.077	0.040	0.037	0.043	0.231
7-8%				0.101	0.030	0.002	0.049	0.182
8-9%				0.106	0.058	0.009	0.039	0.212
9-10%	0.044			0.012	0.017	0.010	0.017	0.100
10-20%	0.281	0.025		0.926	0.432	0.137	0.253	2.055
20-30%	0.086			0.708	0.547		0.068	1.409
30-40%	0.250			0.386	0.462		0.018	1.116
40-50%	0.066			0.280	0.378	0.020	0.081	0.823
50-100%	0.066			0.692	0.550			1.308
Total	0.872	0.058	0.006	3.627	2.621	0.277	0.724	8.186

Source: BI calculations from NLO and 2011 Census data

We estimate linear models in which the dependent variable is our 'Share of participants' measure as indicated above. Separate models are estimated by sex. We also distinguish settlements with

fewer than 600 residents, for the following reasons. First, the labour market situation of Roma people might be different in small settlements. Second, we avoid biases caused by missing or outlier data of very small villages. In smaller than 600-person villages 8.2% of the entire Roma population live (see Table 11). Almost half of them – 3.6% of the entire Roma population – live in Southern-Transdanubia, mostly in villages where share of Roma in the population is above average.

Estimation results are not in line with the expected positive relationship. (see Table 12 and Table 13) After controlling for geographical characteristics and differentiating the potential effect of share of Roma in the population by regions, results suggest that share of participants is in fact lower in settlements with a higher share of Roma in the population. The effect is even more negative in the most disadvantaged regions in case of women.

Table 12. The relationship between settlement-level participation rate and share of Roma in the population – settlements with more than 600 inhabitants

	Women	Women	Men	Men
Proportion of Roma in the population	0.00870	0.0261	0.0498***	0.0202
	(0.00992)	(0.0166)	(0.0124)	(0.0194)
Square of the proportion of Roma in the population	-0.0466*	-0.0223	-0.0939***	-0.0634***
	(0.0248)	(0.0177)	(0.0304)	(0.0207)
Northern Great Plain (NGP)		0.0135***		0.0141***
		(0.00163)		(0.00202)
Southern Great Plain (SGP)		0.00876***		0.00845***
		(0.00119)		(0.00134)
Southern Transdanubia (ST)		0.0110***		0.0116***
		(0.00148)		(0.00186)
Northern Hungary (NH)		0.0132***		0.0139***
		(0.00136)		(0.00174)
Central Transdanubia (CT)		0.00850***		0.00670***
		(0.00115)		(0.00145)
Western Transdanubia (WT)		0.0111***		0.00923***
		(0.00136)		(0.00176)
Capital or county capital		-0.00344**		-0.00487**
		(0.00166)		(0.00211)
City with over 10,000 inhabitants		0.00115		-0.000383
		(0.00157)		(0.00213)
Village, 5 000-6 750 inhabitants		0.00144		-0.000928
		(0.00213)		(0.00251)
Village, 2 000-3 000 inhabitants		-0.00179		-0.00182
		(0.00153)		(0.00208)
Village, 1 000-2 000 inhabitants		-0.00140		-0.000863
		(0.00155)		(0.00211)
Village, 601-1 000 inhabitants		0.00102		0.00180
		(0.00171)		(0.00230)
Share of Roma and NGP interaction		-0.0567***		-0.0236
		(0.0199)		(0.0243)
Share of Roma and SGP interaction		0.00512		0.0286
		(0.0288)		(0.0316)
Share of Roma and ST interaction		-0.0230		0.0239
		(0.0213)		(0.0249)
Share of Roma and NH interaction		-0.0389**		0.000503
		(0.0181)		(0.0218)
Share of Roma and CT interaction		-0.0263		-0.0162
		(0.0330)		(0.0427)
Share of Roma and WT interaction		0.00153		0.0944
		(0.0301)		(0.0619)
Constant	0.0213***	0.0125***	0.0230***	0.0147***
	(0.000455)	(0.00149)	(0.000568)	(0.00201)
Observations	1,929	1,928	1,929	1,928
R-squared	0.007	0.074	0.013	0.075

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Baseline categories: Central Hungary region, City with 6 750 - 10 000 inhabitants, and interaction of share of Roma in the population and Central Hungary. Source: BI calculation from NLO and 2011 Census data.

Table 13. The relationship between settlement-level participation rate and share of Roma in the population – settlements with less than 600 inhabitants

VARIABLES	(1) Women	(2) Women	(3) Men	(4) Men
Proportion of Roma	-0.0218	0.463	0.0600**	-0.425
	(0.0161)	(0.563)	(0.0250)	(0.358)
Square of the proportion of Roma	0.0179	0.00486	-0.0891**	-0.0861**
	(0.0294)	(0.0305)	(0.0378)	(0.0420)
Northern Great Plain (NGP)		0.00771		0.0276***
		(0.0113)		(0.0104)
Southern Great Plain (SGP)		0.00503		0.00766
		(0.0110)		(0.00810)
Southern Transdanubia (ST)		0.00950		0.0250***
		(0.0109)		(0.00798)
Northern Hungary (NH)		0.0178		0.0235***
		(0.0116)		(0.00826)
Central Transdanubia (CT)		0.00964		0.0135
		(0.0110)		(0.00831)
Western Transdanubia (WT)		0.0119		0.0187**
		(0.0109)		(0.00820)
Share of Roma and NGP interaction		-0.480		0.451
		(0.563)		(0.360)
Share of Roma and SGP interaction		-0.564		0.419
		(0.565)		(0.362)
Share of Roma and ST interaction		-0.477		0.475
		(0.563)		(0.359)
Share of Roma and NH interaction		-0.478		0.483
		(0.563)		(0.360)
Share of Roma and CT interaction		-0.522		0.424
		(0.563)		(0.362)
Share of Roma and WT interaction		-0.500		0.394
		(0.563)		(0.359)
Constant	0.0266***	0.0152	0.0230***	0.00893
	(0.00131)	(0.0107)	(0.000568)	(0.00737)
Observations	1,249	1,249	1,929	1,249
R-squared	0.002	0.012	0.013	0.018

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1
Baseline categories: Central Hungary region and interaction of share of Roma in the population and Central Hungary.
Source: BI calculation from NLO and 2011 Census data.

4.3.1.2 Comparing settlements with and without programme participants

Out of the 3,178 settlements of Hungary, women participated in 2,590 and men took part in the programme in 2,558 settlements (see Table 14). Settlements left out from the programme are typically small villages with an average of 319-379 inhabitants and with a higher-than-average share of Roma in the population. Cumulated population of these settlements is about 200,000,

with approximately 10 000 – 15 000 Roma inhabitants according to the Census data. This comprises 3-5% of the entire Roma population.

Table 14. Comparison of settlements with and without SROP 1.1.2 participants

	Settlements with participants			Settlements without participants		
	No. of settlements	Mean	Standard deviation	No. of settlements	Mean	Standard deviation
Women						
Population (capita)	2590	3765	12541	586	319	331
Share of Roma people (%)	2590	3.1	5.3	586	8.0	14.7
Men						
Population (capita)	2558	3793	12616	618	379	481
Share of Roma people (%)	2558	3.1	5.6	618	4.4	9.7

Source: BI calculation based on NLO register and 2011 Census data.

and Figure 2 illustrates whether settlements with and without participants differ with respect to the share of Roma in the population. The 45-degree line represents the theoretical situation in which the distribution of the share of Roma in the population is the same in the two groups of settlements. Area below the line represents that share of Roma in the population is higher in non-participant settlements, and vice versa. Again, we consider above-600 and under-600-inhabitant settlements separately.

As pointed out earlier, efficient targeting of the Roma would imply that due to an overrepresentation of Roma people in the target groups we find higher concentration of the Roma in participant settlements. However, Figure 2 suggests that in case of settlements with a relatively low share of Roma in the population, covered and non-covered settlements do not differ in this respect. Furthermore, settlements with a relatively high share of Roma seem to be more likely to be left out from the programme. This phenomenon is even more pronounced in case of women.

Figure 1. Comparison of settlements with respect to share of Roma in the population – Settlements with more than 600 inhabitants
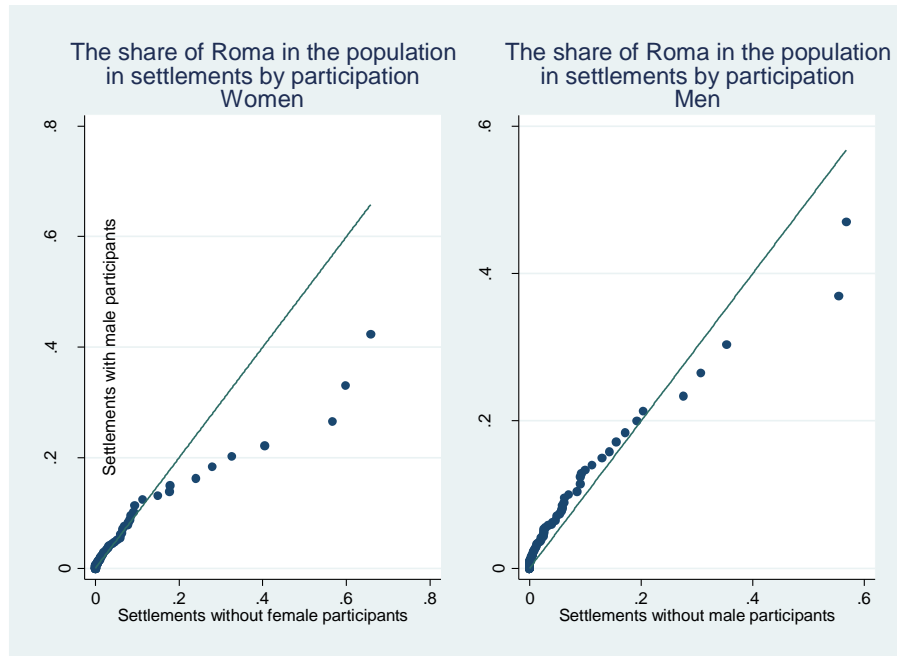
Source: own estimation based on NLO data

Figure 2. Comparison of settlements with respect to share of Roma in the population – Settlements with less than 600 inhabitants
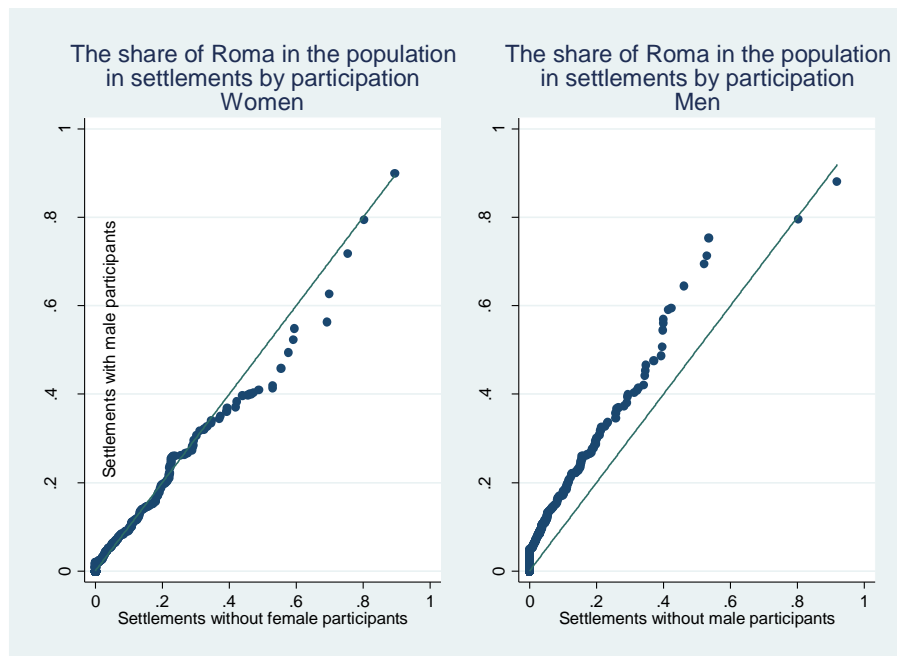
Source: BI estimation based on NLO and 2011 Census data

Among small settlements, share of Roma in the population is slightly higher in case of women while slightly lower in case of men in those with no programme participants. (see Figure 2)

4.3.1.3 Correlation between the share of Roma in the population and the number of participants using additional measures

The correlation between share of Roma in the population and share of programme participants has been investigated by two additional measures as well. In this section settlements are assigned to two categories based on whether the share of Roma in the population is below or above the country average/median. Consequently, we construct the following two binary variables:

Variable 1:

> It takes value 1, if share of Roma in the population is higher than weighted country average (3.2 %) and 0 otherwise;

Variable 2:

> It takes value 1, if share of Roma in the population is higher than weighted country median (1.3 %) and 0 otherwise.

Then we estimate the same models presented in Table 13 using one of these new binary variables as explanatory variables instead of the continuous measure of share of Roma in the population. As usually, we run estimations separately for settlements with fewer and more than 600 inhabitants.

Results are summarized in Table 38 and Table 39 of Appendix 2.2. In case of above-600 settlements the share of participants is higher when share of Roma is above mean or median. However, although the difference is significant, it is extremely small: one thousandth of a percentage point. Also, it varies through regions and in one of the most disadvantaged regions, in Northern Grate Plain it turns to negative.

4.3.2 Selection on individual level

In this chapter we examine whether probability of participation is related to share of Roma in the population of home settlement using individual level data. We compare all participants (57,894 people) to those eligible for but left out from the programme (1,301,419 people). As before, those being eligible are selected from the unemployment registry, and they are either below 25, or above 50, or uneducated or long-term unemployed. [24]

The following selection model was estimated by gender:

$$P(Participation = 1|X) = X'\beta$$

The dependent variable of the model[25] is a binary variable taking the value of 1 if the individual participated in the programme, and 0 if not. *X* refers to the following explanatory variables:

- share of Roma in the population of home settlement;
- settlement type;
- region;
- interaction term of share of Roma in the population and region;
- education level;
- whether the individual is disabled; and
- age.

Results are similar than those from the earlier settlement-level analysis: no positive correlation is found between the share of Roma and the probability of programme participation. Furthermore,

---

[24]To identify people who were eligible for the programme we could not consider all participation conditions. Of the registered unemployed we considered as potential participants only those being long-term unemployed, or aged below 25, or above 50, or having at most elementary school degree. We could not identify parents with young children and those at risk of long-term unemployment. Thus, only 1,333,590-1,301,419=32,171 actual participants belong to the group of potential participants while the remaining 57,894-32,171=25,723 people do not fall below the conditions we could filter for; they most probably entered the programme through those two additional channels. Also, there are 15 people not being included in the unemployment registry and thus not included in the pool of potential participants in the first place.

[25]Linear probability model using standard errors clustered at sub-regional level to control for regionally different random shocks.

in some regions high share of Roma in the population decreases the probability of entering the programme with respect to the baseline.

Table 15. The relationship between participation probability and share of Roma in the population - settlements with less than 600 inhabitants

	Women	Women	Men	Men
Proportion of Roma	-0.00663	1.257	0.0470	-0.564
	(0.0416)	(1.098)	(0.0369)	(0.693)
Square of the proportion of Roma	0.00602	-0.0712	-0.0549	-0.104
	(0.0647)	(0.0727)	(0.0567)	(0.0696)
Northern Great Plain (NGP)		-0.0309***		0.00452
		(0.00833)		(0.0158)
Southern Great Plain (SGP)		-0.0216**		0.00191
		(0.00860)		(0.0152)
Southern Transdanubia (ST)		-0.0310***		0.00347
		(0.00800)		(0.0155)
Northern Hungary (NH)		-0.0183***		0.00702
		(0.00667)		(0.0150)
Central Transdanubia (CT)		-0.0206**		0.00444
		(0.00920)		(0.0155)
Western Transdanubia (WT)		-0.00836		0.0144
		(0.00598)		(0.0151)
Secondary education		0.0302***		0.0208***
		(0.00353)		(0.00222)
Higher education		0.0474***		0.0273***
		(0.00690)		(0.00487)
Disability		-0.0156***		-0.00999***
		(0.00206)		(0.00193)
Age		-1.16e-06***		-2.21e-06***
		(1.55e-07)		(1.82e-07)
Share of Roma and NGP interaction		-1.216		0.604
		(1.099)		(0.693)
Share of Roma and SGP interaction		-1.318		0.580
		(1.103)		(0.696)
Share of Roma and ST interaction		-1.187		0.664
		(1.100)		(0.694)
Share of Roma and NH interaction		-1.195		0.647
		(1.100)		(0.694)
Share of Roma and CT interaction		-1.224		0.635
		(1.102)		(0.695)
Share of Roma and WT interaction		-1.305		0.555
		(1.099)		(0.693)
Constant	0.0283***	0.0521***	0.0216***	0.0359**
	(0.00379)	(0.00541)	(0.00273)	(0.0147)
Observations	54,660	54,660	63,918	63,918
R-squared	0.000	0.022	0.000	0.013

Clustered robust standard errors at the sub-regional level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Source: BI calculation from NLO and 2011 Census data.

Table 16. The relationship between participation probability and share of Roma in the population - settlements with more than 600 inhabitants

	Women	Women	Men	Men
Proportion of Roma	-0.0926***	0.0430	-0.00298	0.0740
	(0.0329)	(0.0406)	(0.0298)	(0.0474)
Square of the proportion of Roma	0.0499	0.0322	-0.0405	0.00446
	(0.0444)	(0.0437)	(0.0440)	(0.0345)
Northern Great Plain (NGP)		0.0169***		0.0156***
		(0.00611)		(0.00520)
Southern Great Plain (SGP)		0.0139***		0.0114***
		(0.00468)		(0.00429)
Southern Transdanubia (ST)		0.0274***		0.0300***
		(0.00853)		(0.00790)
Northern Hungary (NH)		0.0176***		0.0166***
		(0.00561)		(0.00545)
Central Transdanubia (CT)		0.0152***		0.00848**
		(0.00434)		(0.00409)
Western Transdanubia (WT)		0.0312***		0.0217***
		(0.00687)		(0.00483)
Secondary education		0.0459***		0.0336***
		(0.00244)		(0.00140)
Higher education		0.0777***		0.0543***
		(0.00464)		(0.00271)
Capital or county capital		-0.00574		-0.00110
		(0.00458)		(0.00422)
City with over 10,000 inhabitants		-0.00138		-0.00157
		(0.00363)		(0.00321)
Village, 5 000-6 750 inhabitants		0.000214		-0.00302
		(0.00450)		(0.00391)
Village, 2 000-3 000 inhabitants		-0.00536		-0.00465
		(0.00353)		(0.00334)
Village, 1 000-2 000 inhabitants		-0.00705**		-0.00646**
		(0.00349)		(0.00311)
Village, 601-1 000 inhabitants		-0.00221		-0.00344
		(0.00388)		(0.00388)
Disability		-0.0210***		-0.0153***
		(0.00147)		(0.00134)
Age		-2.38e-06***		-3.21e-06***
		(1.81e-07)		(1.08e-07)
Share of Roma and NGP interaction		-0.111**		-0.114**
		(0.0557)		(0.0572)
Share of Roma and SGP interaction		0.0337		0.000180
		(0.0667)		(0.0687)
Share of Roma and ST interaction		-0.0950		-0.102
		(0.0656)		(0.0633)
Share of Roma and NH interaction		-0.0762		-0.0749
		(0.0496)		(0.0532)
Share of Roma and CT interaction		-0.120**		-0.120**
		(0.0556)		(0.0604)
Share of Roma and WT interaction		-0.0327		0.0200
		(0.0708)		(0.0622)
Constant	0.0537***	0.0457***	0.0394***	0.0547***
	(0.00332)	(0.00457)	(0.00277)	(0.00476)

Observations	587,580	587,580	652,946	652,946
R-squared	0.001	0.035	0.000	0.023

Clustered robust standard errors at the sub-regional level are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Source: BI calculation from NLO and 2011 Census data

4.3.3 Conclusions

Both settlement and individual-level data suggest that the relative number of participants is not increasing with the share of Roma in the population. In some especially disadvantaged regions and mostly in case of women, this correlation is in fact negative. Relatively more participants seem to be involved in the programme from settlements where the share of Roma is lower. As we have no access to individual level ethnicity data, we cannot tell whether participants from settlements with a low or high share of Roma in the population are actually Roma or not – we simply assume that participants are more likely to be Roma if relatively more Roma live in their environment. Furthermore, even if the programme was inefficient in reaching the Roma, one cannot distinguish whether the reason behind is selection applied by unemployment offices, or self-selection of potential Roma participants. To develop a deeper understanding on this phenomenon, unemployment offices' selection practices should be surveyed.

There seems to be a difference between participation patterns of Roma women and men: women were more likely to be left out from the programme. Again, it's impossible to distinguish between gender-based discrimination and labour supply characteristics of Roma women; however, there is empirical evidence on lower labour market activity of Roma women. (Janky, 2007) To increase take up of Roma women programmes should target them separately.

We conclude that involvement of the Roma in the SROP 1.1.2 programme was very low. The share of Roma people among participants is far smaller than their share in the target group. To be able to increase Roma take up in such programmes one needs to gain more understanding about reasons behind its low level. One can learn more through two channels:

1. Making evaluation more solid, individual-level ethnicity data of actual and potential participants are needed;

2. Selection practice of unemployment offices has to be surveyed.

4.4. Employment impact of the programme

The causal impact of the programme on the probability of finding employment is the difference between

- the probability of programme participants founding a job (factual results), and

- the probability of that programme participants would have found a job if they had not participated in the programme (counterfactual results).

As we cannot observe counterfactual results, being a purely theoretical concept, we are going to estimate them by constructing a control group using matching methods. Then, we define the causal effect of the programme as the difference between the outcomes of participants and controls. We evaluate the causal effect of the programme in case of uneducated participants only as most Roma participants belong to this group.[26]

We start by comparing raw entry-to-employment rates among uneducated actual and potential programme participants.

Figure 3 shows the ratio of those finding a job before, during and after programme in the two groups.[27] Programme entry started on January 1, 2008, and our database includes entries until December 31, 2010. Between 2006 and 2008 the two series move together, there is no difference in employment entry rates of actual and potential programme participants. However, in the beginning of 2008 employment rate of actual participants starts to increase, while that of the potential participants starts to decrease.

---

[26] By uneducated we mean having an elementary school degree at most.

[27] Potential participants are the uneducated registered unemployed, who were registered as being unemployed for at least one day during the entering period of the programme (January 1, 2008 – December 31, 2010).

Figure 3. Employment entry rates of uneducated actual and potential participants of SROP 1.1.2
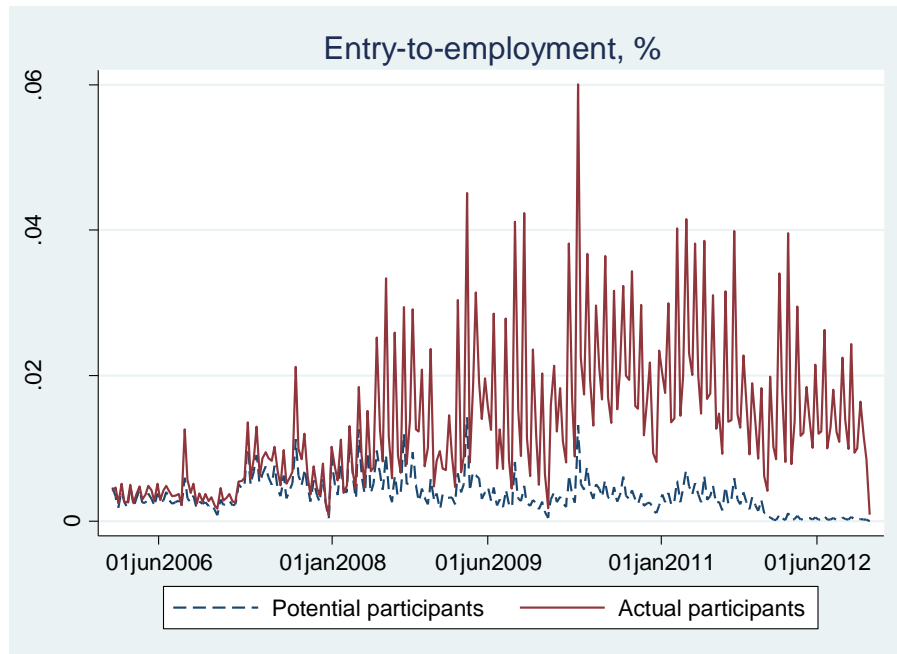
Entry-to-employment, %

Source: BI calculations based on NLO data.
N=14,601 uneducated participants and 737,699 uneducated potential participants. Variable: Number of people entering employment in counted in 10-day subsequent periods, in the percentage of the sample size.

Figure 4 illustrates the share of actual and potential participants entering unemployment between 2006 and 2012. Again, curves go together before 2008 then they start to diverge. However, share of participants entering the registry is actually higher than that of potential participants. Around the end of programme entry period the curves get closer to each other again.

Figure 4. Unemployment entry rates of uneducated actual and potential participants of SROP 1.1.2
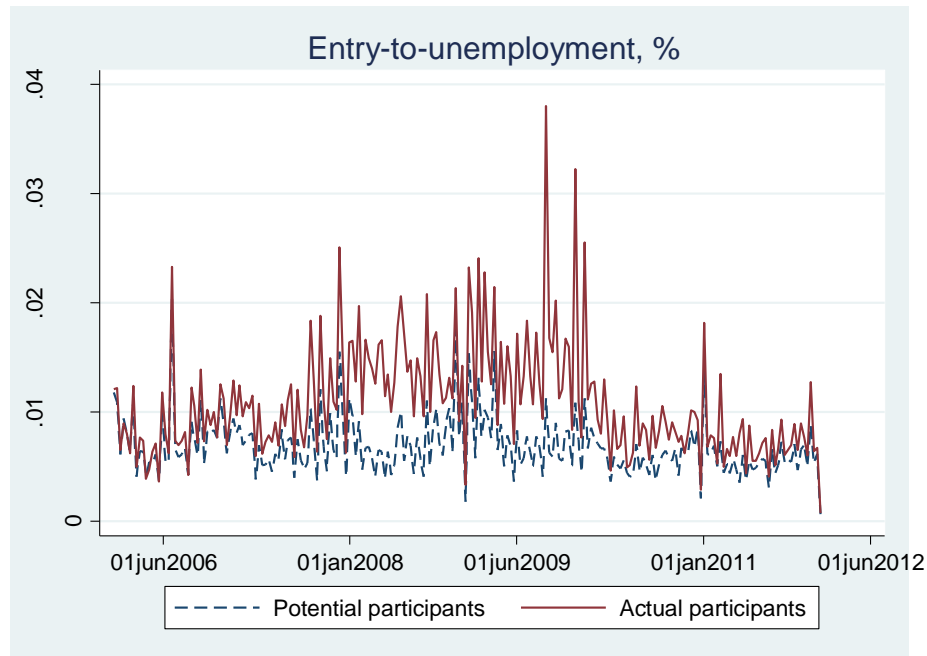
Source: BI calculations based on NLO data.
N=14,601 uneducated participants and 737,699 uneducated potential participants. Variable: Number of people entering unemployment in counted in 10-day subsequent periods, in the percentage of the sample size.

However, raw outcomes of the two groups are not directly comparable, because they differ in their observable characteristics (see Appendix 2.3). Instead, we need to compare factual and counterfactual outcomes. To create a control group we narrow down the pool of actual and potential participants as follows:

1. We restrict both samples to men only. Women with young children might have been involved in the programme; however, we cannot identify parents with young children among potential participants, so we cannot match participants and controls to each other based on that criteria. To handle this problem we have to leave out women from the evaluation. Narrowing down to men only, 7,948 participants and 379,878 potential participants remain in the sample.

2. We exclude people who participated in other programmes during the period of SROP 1.1.2 to effects of SROP 1.1.2 exclusively. The remaining sample thus includes 7,049 participants and 371,789 potential participants.

Next, we choose exactly one pair for each programme participant (called "treat" from now on) based on his observable characteristics. [28] We match treat-control pairs such that they are as similar to each other as possible. In particular,

- they entered unemployment registry within the same +/- 60-day period;

- control pair is unemployed when treat enters the programme;

- they are in the same age range;

- they have the same educational attainment: they both are drop-outs or they both finished elementary school at most;

- they live in the same region and in a same type of settlement;

- their estimated probability of participation (propensity score) is as close to each other within the intersection of all categories above as possible.

Each potential control is matched at most once, so we select our control sample without replacement.[29] If we could not find an appropriate pair for a participant with the above detailed procedure, we exclude that participant form the sample. Using this method we can find a pair to 6,946 participants out of the 7,049, that is only 103 participants are left out because of having no proper pair.

Finally, we have a sample of 6,946 uneducated programme participants and the same amount of non-participant controls. The distributions of observable characteristics in the two groups are not different anymore, so their outcome variables may be compared (see Appendix 2.3, Table 46).

Comparison of outcome variables in treatment and control groups

---

[28] Comparison of the three samples – all participants, narrowed sample and the final matched sample – can be found in Appendix 2.3.

[29] Because we generated the sample of matched nonparticipants without replacement, the order of the observations in the sample might influence final results. (This does not mean that later periods will have only "poorer" controls.) To deal with this problem we generated different set of matched controls. Using the different set of matched controls did not lead to substantially different results.

To evaluate the effect of the programme we construct five outcome variables. Three of these are constructed based on the exit to unemployment data of the SHLD database and the unemployment registry as follows:

> Outcome 1: Exit to employment during or after the programme. It takes on the value 1 if the observed person enters employment according to the SHDL database or exits unemployment to employment according to the unemployment registry during programme participation or within 6 months afterwards; and 0 otherwise.

> Outcome 2: Exit to employment after the programme. Its value is equal to 1 if the observed person enters employment according to the SHDL database or exits unemployment to employment according to the unemployment registry within 6 months after the programme; and 0 otherwise.

> Outcome 3: Exit to employment any time during the observation period. It takes on the value 1 if the observed person enters employment according to the SHDL database or exits from unemployment to employment according to the unemployment registry anytime after entering the programme; and 0 otherwise.

As programme entry date and duration of participation is different for each individual, observation period is also individual-specific. We follow participants and their control pairs for exactly the same time period. These periods start in the beginning of 2008 and end during the second half of 2012. Most participants (70%) spent less than 1 year in the programme; average participation duration is 9.5 months.

Figure 5. Timing of outcome variables of employment
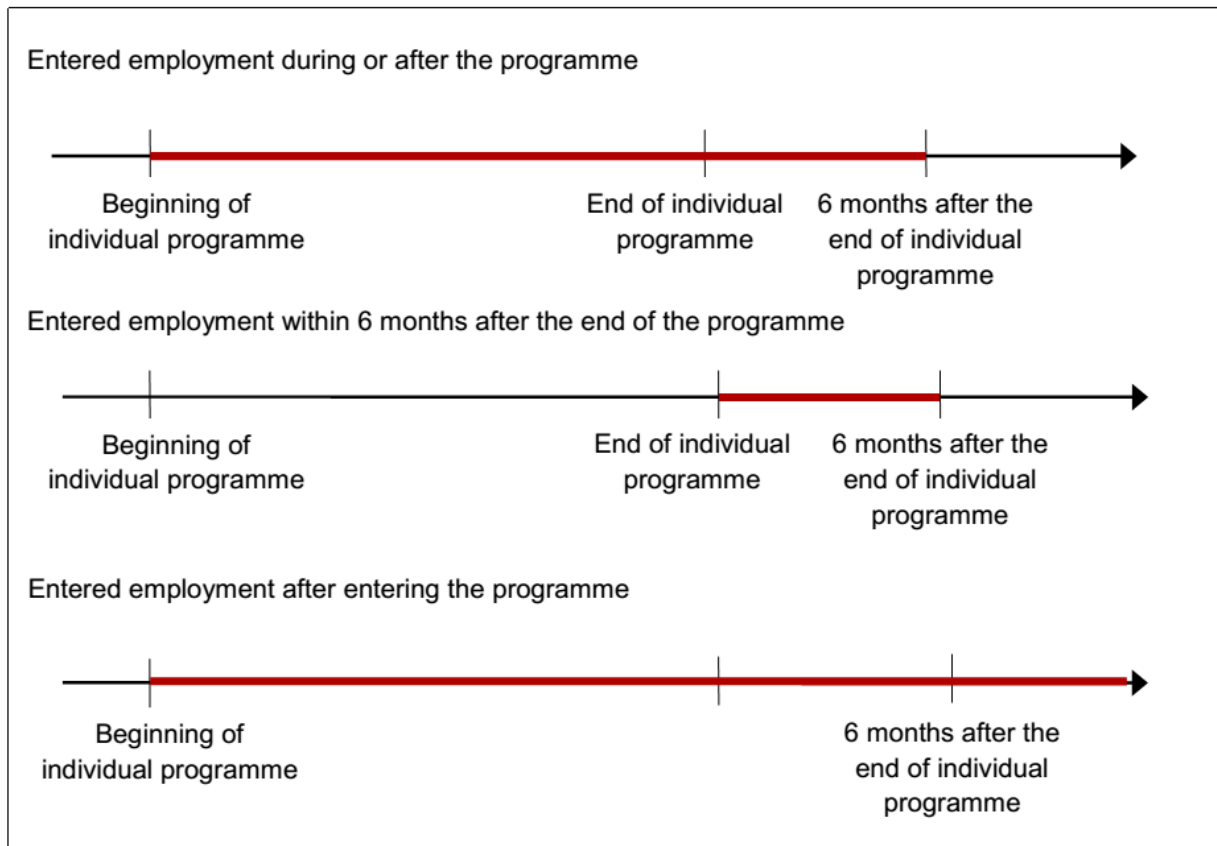
In case of 20% of participants (1,382 people) our database does not contain programme exit date. For these observations we estimate an exit date based on average programme length (285 days). To further refine our analysis we separate those participants from our sample who received wage or wage-cost subsidies under the programme as part of their individual package.

The other two outcome variables that we use are based solely on the unemployment register:

> Outcome 4: Not re-entering unemployment after the programme for 6 months. Its value is equal to 1 if the observed person exits the last unemployment spell he entered before exiting the programme and does not register again for 6 month; and 0 otherwise.

> Outcome 5: Not re-entering unemployment after the programme. It takes on the value 1 if the observed person exits the last unemployment spell he entered before

exiting the programme and does not register again during the whole observation period; and 0 otherwise.

Figure 6. Timing of outcome variables of non-reentering unemployment
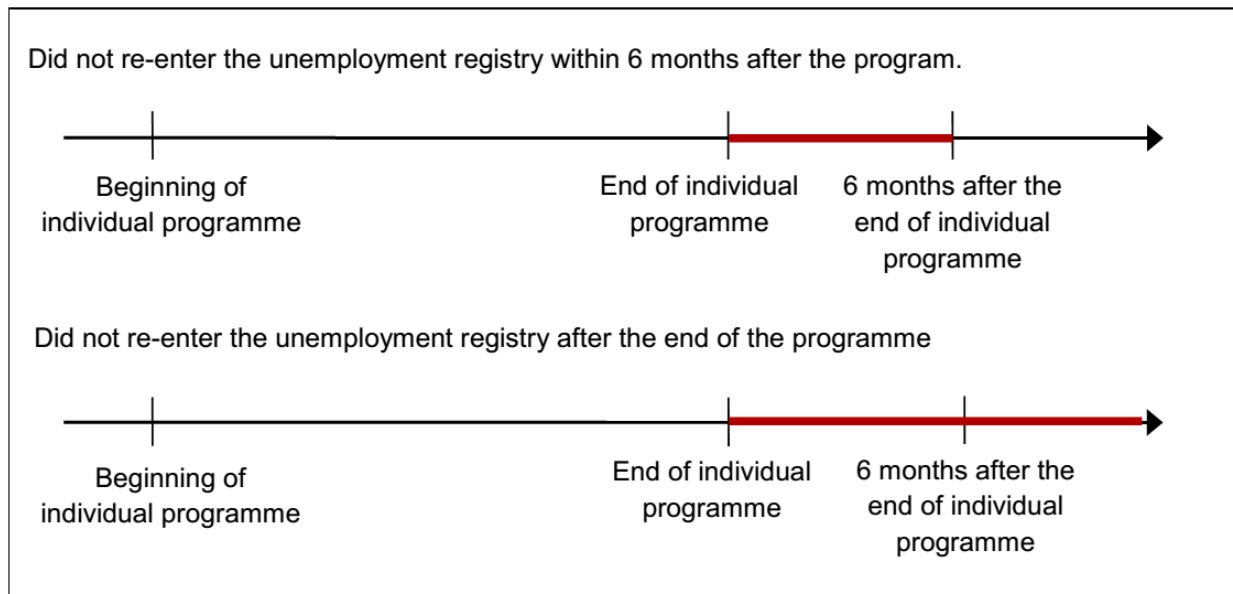
Results are presented in Table 17. Depending on which outcome variable is used, 75-91% of participants found employment, while these ratios in the matched control group are only 11-42%. Most participants found a job during the programme. If we do not consider jobs with wage or wage-cost subsidies, employment rates are still as high as 43-89% among participants while only 10-42% among controls. Consequently, the programme increased the probability of finding a job in the medium term by 49%points. Programme effect excluding wage subsidy recipients is still 39%points.

Outcome variables based on non-re-entering unemployment show a smaller effect. Data suggest that 49-75% of participants left unemployment registry and did not register again, comparing to 43-53% of controls. However, while over 90% of participants entered a job at least once, only 75% of them did not return to the unemployment register. Among controls, 42% percent found a job while 53% of left the register with no return.

Table 17. Employment effect of SROP 1.1.2

MEN	All participants				Without wage subsidy			
	Control		Treated		Control		Treated	
	capita	%	capita	%	capita	%	capita	%
Entered employment during the programme or within 6 months after the end of the programme.	2197	32%	5245	76%	1644	31%	3628	68%
Entered employment after the end of the programme within 6 months	740	11%	2808	40%	559	10%	2294	43%
Entered employment at some point of the whole observed period after entering the programme.	2939	42%	6336	91%	2214	42%	4719	89%
Exited unemployment and did not enter again within 6 months after the end of the programme.	2998	43%	3418	49%	2271	43%	2550	48%
Exited unemployment and did not enter again during the whole observed period.	3652	53%	5181	75%	2791	52%	4036	76%
Number of observations		6946		6946		5329		5329

Note: uneducated men participants who were not included in any other programmes.
Source: BI estimation using NLO and 2011 Census data.

Measuring programme effect using employment-based outcome variables gives a 29-44%points impact while using non-re-entering-unemployment-based outcome variables results in a 6-22%points estimated effect. Those not re-entering unemployment without finding a job might either have become inactive or have worked unofficially in the shadow economy. It might be reasonable to assume that that the controls are more likely to work in unregistered jobs as this can actually be a reason why they have not participated in any programmes in spite of being eligible. Using this assumption, the employment-based measure overestimates while unemployment-based measure underestimates the effect of the programme. The real impact could be somewhere between 6-22 and 29-44%points. Effects are of gross nature: we cannot correct for substitution effect and deadweight loss.[30]

Outcome measures do not differentiate between entering employment once and several times. As they are binary, they take value 1 in case of one and several exits to employment as well. On Figure 7 and Figure 8 we plot each entry to employment and unemployment of the matched

---

[30] Substitution effect refers to the situation when a programme participant is employed instead of someone who did not participate; deadweight loss emerges if a programme participant who found employment would have found employment without participating in the programme as well.

sample between 2006 and 2012. The result is similar than before: impact based on re-entry to unemployment is much smaller then based on entry to employment.

Figure 7. Unemployment entry rates in the matched evaluation sample of SROP 1.1.2
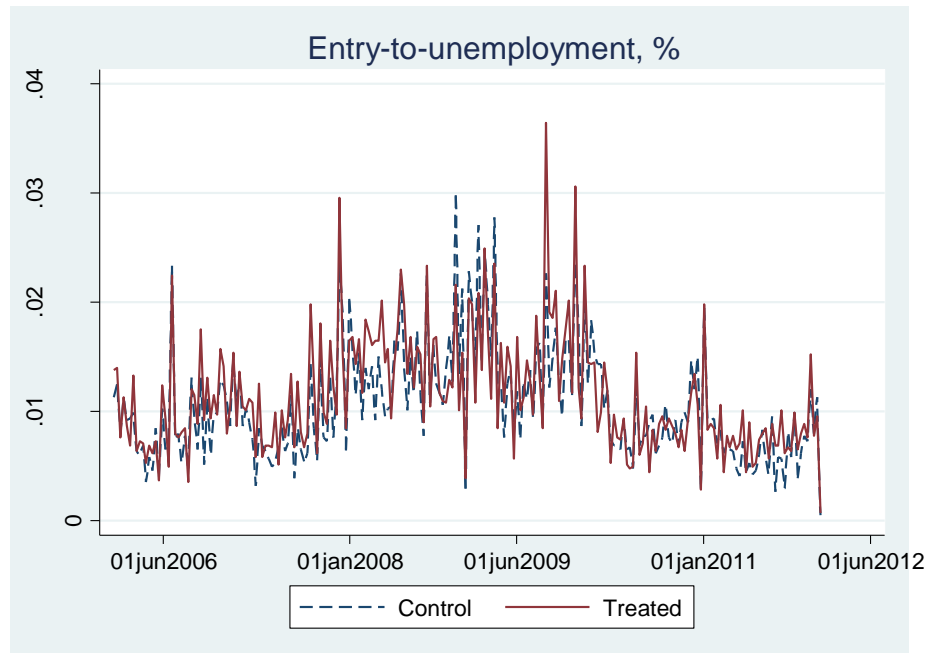
Source: BI estimation using NLO and 2011 Census data.
N=6 946 treated, 6 946 control

Figure 8. Employment entry rates in the matched evaluation sample of SROP 1.1.2
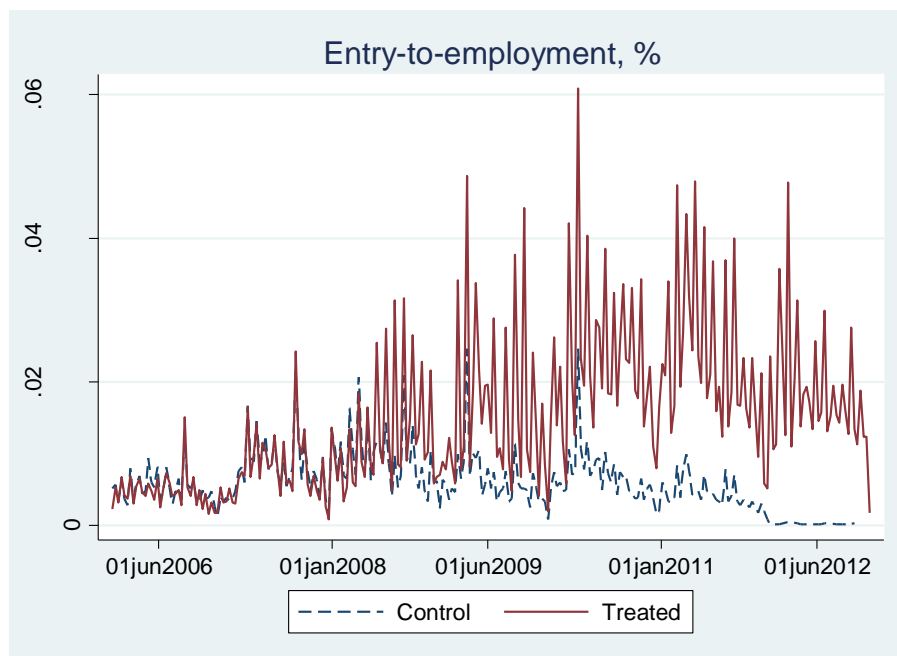
Source: BI calculation using NLO and 2011 Census data.
N=6 946 treated, 6 946 control

Restricting sample to those being long-term unemployed on day of entering programme, thus had been unemployed for more than 365 consecutive days, programme still have a significant positive effect on finding employment (see Table 18). Long-term unemployed participants are 48%points more likely to find employment during the programme or within 6 months afterwards. Similarly to previous results, probability of exit to employment among long-term unemployed is somewhat lower but still 39%points without wage or wage-cost subsidies.

Table 18. Employment effect of SROP 1.1.2 – long-term unemployed

MEN	All participants				Without wage subsidy			
	Control		Treated		Control		Treated	
	capita	%	capita	%	capita	%	capita	%
Entered employment during the programme or within 6 months after the end of the programme.	258	23%	774	71%	193	23%	519	62%
Entered employment after the end of the programme within 6 months	98	8.6%	388	35%	74	8.6%	309	37%
Entered employment at some point of the whole observed period after entering the programme.	388	34%	983	90%	293	34%	728	86%
Exited unemployment and did not enter again within 6 months after the end of the programme.	367	32%	475	43%	263	31%	349	41%
Exited unemployment and did not enter again during the whole observed period.	472	42%	812	74%	352	41%	632	75%
Number of observations		1 133		1 097		856		842

Source: BI estimation using NLO and 2011 Census data.

5. THE *ONE STEP AHEAD!* PROGRAMME

The *One Step Ahead!* programme offered general or vocational training to participants having primary education or less, and in exceptional cases, vocation retraining to those with a vocation considered outdated. The programme had two phases. The first phase, *One Step Ahead! I.* (HRDOP 3.5.3) was launched at the beginning of 2006. It was organised within the framework of Human Resources Development Operational Programme (HRDOP)[31], and co-financed by the Hungarian government and the European Union. The second phase, *One Step Ahead! II.* (SROP

---

[31] In Hungarian: *Humánerőforrás-fejlesztési Operatív Program, HEFOP*

2.1.1) programme started on the 1ˢᵗ of September, 2007, this time financed by the Social Renewal Operational Programme (SROP) of the New Hungary Development Plan. The second phase ended on the 4ᵗʰ of November, 2009. This phase was augmented with a sub-phase, SROP 2.1.1/B, for the beneficiaries of a social benefit, Nursing Aid[32], between December 15, 2009 and December 30, 2010. Those taking care of their disabled family member at home, and thus not being able to work full-time, are eligible for Nursing Aid.

The main goal of the programme was to target the uneducated and unskilled, and improve their job market potential by offering an opportunity to finish primary education and/or completing free vocational training. Those who already had vocational qualification but wished to get another one in a different field with more demand were also allowed to take part. Besides training, participants also received cash transfers (the amount of one-month minimum wage after every successful 150 hours of training) during the programme. Those who took part in the first phase but dropped out without completing their training plan were not allowed to participate in the second phase.

According to our database, 23,088 persons participated in at least one phase of the programme (Table 19). 6,953 participated in the first phase (HRDOP 3.5.3) and 16,224 took part in the second phase (SROP 2.1.1 and 2.1.1-B.).

Table 19: Inflows to the One step ahead! programme

	Entering periods	No. of participants according to the NLO database[33]
HRDOP 3.5.3	9 Jan 2006 – 31 Dec 2007	6 953
SROP 2.1.1	1 Sept 2007 – 14 Nov 2009	13 790
SROP 2.1.1-B	15 Dec 2009 - 31 Dec 2010	2 530

Source: BI calculations from NLO data.

Participants are identified in one of three ways in the programme participant's database:

---

[32] Nursing Aid: Ápolási díj.

[33] The number of those participating in the given phase at least once. Total number of participants with no overlaps: 23,088.

1. identification number found in the database can be linked to the unemployment registry data of the given person in case of 16,360 participants;

2. identification number cannot be linked to the unemployment registry data of the given person in case of 4,581 participants;

3. identification number is missing in case of 2,147 participants.

Therefore, in those steps of the analysis where we need additional data outside from the participation database itself we have to restrict the analysis to the sample of those participants whose data can be linked to the registry.[34]

Also, in case of those participants who can be linked to the unemployment registry based on their identification number, there are some controversies between the educational attainment data of the participation database and the unemployment registry. Eligibility criteria of the programme included:

- the uneducated, who completed primary education at most, and

- those with out-of-date professions.

Based on the educational attainment data of the participation database, 88% of the 23,088 participants entered the programme completing primary education at most. Based on the educational attainment data of the unemployment registry, in case of those participants who can be linked to their unemployment registry data (16,360 participants, see above), this ratio is 65%.[35] Neither we nor the data supplier NLO can give an explanation for the discrepancy of the two data sources. To handle the problem, in the following analysis we use the educational attainment data from the unemployment registry for those participants who can be linked to the unemployment registry (16,360 participants), and we use the educational attainment data from the original participation database for those who cannot be linked (4,581+2,147=6,728

---

[34] The comparison of subsamples can be found in Appendix 3.

[35] Restricting the sample to those 16 360 participants who can be linked to the unemployment registry based on their identification number, but using the original educational attainment data from the participation database, the share of the uneducated is 86%. The difference between the shares of the uneducated in the two data sources does not come from the different sample size.

participants) to that. Using this data handling method, 16,719 participants (72%) entered the programme with no higher education than a primary degree.

As before, we analyse selection by comparing programme participants to those who would have been eligible but did not participate (potential participants). We select potential participants from the unemployment registry. However, unemployment registry does not contain information about professions of the unemployed, so we cannot select potential participants with out-of-date professions. Therefore, we proceed as follows:

- At first, we investigate selection on a sample restricted to uneducated participants, and we also select uneducated potential participants from the unemployment registry as the base of comparison (Chapter 5.1).

- Then, we broaden the sample of the programme participants to everybody, and we also select a larger sample of potential participants from the unemployment registry, including those completing vocational training.[36] In this case we assume that all types of professions can be seen as out-of-date, depending on the local labour market conditions and the general skills of the unemployed (Chapter 5.2).

- Finally, we conduct an impact evaluation of the programme in case of uneducated participants only, as this is the subsample of participants that most of the Roma could have belonged to. In this case, control group is selected from uneducated potential participants as well. (Chapter 5.3)

5.1. Selection and take up of the uneducated

Age distribution of uneducated participants is slightly different from that of potential participants (Table 20). The share of those under age 50 is 92% among programme participants while 78% in the comparison group. Programme participants are thus younger on average than potential participants.

---

[36] Szakiskola, szakmunkásképző.

Table 20: Age distribution of uneducated participants of the One step ahead! programme

Age	Programme participants		Potential participants	
	Capita	%	Capita	%
Under 25	4 258	25.5	116 222	16.6
25-35	5 257	31.4	179 510	25.6
35-50	5 925	35.4	251 666	35.8
Above 50	1 279	7.6	154 725	22.0
Total	16 719	100.0	702 123	100.0

Source: BI calculations from NLO data. Remark: age as on 01 Sept 2007. Potential participants: 100% sample of the uneducated unemployed taken from the unemployment registry. It contains everyone who was registered for at least 1 day in the unemployment registry during the inflow period of the One step ahead! programme.

The following two tables show take up among the uneducated by region and settlement size, as the function of share of Roma in the population of home settlement. Average take up is 2.3%: from the 702,123-person group of uneducated potential participants 16,719 participated in the programme.[37]

Just as before, we highlighted in both tables those categories where more than 2% of the Roma population belongs to. This way we indicated the cells covering altogether half of the entire Roma population in Hungary. Highlighted cells are the most informative regarding the coverage of the Roma: if targeting of the Roma was successful, take-up rates should be higher most of these categories.

[37] The take-up rate is calculated as the ratio of participants over the sum of the potential and actual participants: take-up rate=No. of participants/(No. of participants + No. of potential participants).

Table 21: Take up in the One step ahead! programme by regions (%)

Share Roma in the population	Regions							Total
	Northern Great Plain	Southern Great Plain	Central Hungary	Southern Trans-danubia	Northern Hungary	Central Transdanubia	Western Trans-danubia	
0%	3.8	1.8	1.7	1.3	1.9	0.8	1.2	1.6
0-1%	2.5	1.9	1.5	2.0	1.7	1.2	2.0	1.8
1-2%	2.8	1.8	1.5	2.9	2.5	1.2	1.7	1.8
2-3%	2.4	2.7	2.2	1.7	2.6	1.5	1.4	2.1
3-4%	3.0	2.6	2.0	1.8	3.0	0.7	0.6	2.5
4-5%	5.1	5.6	2.4	2.7	2.4	0.5	2.5	3.8
5-6%	3.4	3.5	2.0	2.5	3.6	1.1	1.4	3.2
6-7%	2.2	2.8	2.9	2.4	1.4	2.1	1.2	2.2
7-8%	3.2	8.6	0.6	1.5	2.6	1.2	2.2	2.7
8-9%	3.3	3.0	1.6	4.1	2.0	0.7	2.4	2.9
9-10%	4.9	4.6	2.8	1.8	1.4	1.5	4.2	3.2
10-20%	3.6	2.8	7.2	3.1	2.9	5.4	4.6	3.5
20-30%	3.1		16.3	2.3	1.4	0.2	2.5	2.6
30-40%	4.6	1.8		3.3	2.1		27.3	3.0
40-50%	12.9			7.2	1.5	0.0	4.5	2.7
50-100%	3.8			6.0	1.8			3.9
Total	3.2	2.4	1.9	2.4	2.4	1.3	1.9	2.3
The regional distribution of the Roma population (%)	*26.1*	*8.2*	*13.0*	*13.8*	*29.6*	*5.1*	*4.2*	*100.0*

Source: BI calculations from NLO data.

N=16,719 uneducated programme participants and 702,123 uneducated potential participants.

Note: the take-up rate is No. of participants divided by the sum of the number of potential and actual participants. 0 indicates no participants in that category; an empty cell indicates either no settlement or no actual and potential participants in that category. The highlighted cells indicate the categories where at least 2% of the Roma population lives. All highlighted cells together cover 48% of the Hungarian Roma population.

On average, take up of the uneducated is rising with a higher rate of Roma in the population (Figure 9). In terms of regional variation, take up is the highest in the Northern Great Plain region (3.2%), and it is also above average in the Southern Great Plain (2.4%), Southern Transdanubia (2.4%) and Northern Hungary (2.4%) regions. 78% of the Roma population in Hungary live in these four regions. (see Table 21 and Table 6).

Table 22: Take up of the One step ahead! programme by settlement size (%)

The share of the Roma in the settlements	Settlement categories								Total
	Budapest and county centres	Cities, above 10 000 inhabitants	Cities, 6 750-10 000 inhabitants	Villages, 5 000-6 750 inhabitants	Villages, 2 000-5 000 inhabitants	Villages, 1 000-2 000 inhabitants	Villages, 601-1 000 inhabitants	Villages, below 600 inhabitants	
0%			1.1	0.7	1.4	1.5	1.2	2.1	1.6
0-1%	2.0	1.8	2.2	2.0	1.2	1.7	1.4	2.7	1.8
1-2%	1.7	1.9	1.7	1.6	1.6	2.1	1.7	1.4	1.8
2-3%	1.9	2.4	1.4	1.8	2.5	1.9	1.7	2.7	2.1
3-4%	2.9	1.8	3.1	4.2	2.3	1.4	2.0	1.5	2.5
4-5%		1.4	8.4	3.1	6.2	1.9	6.4	3.5	3.8
5-6%	4.8	2.8	3.5	2.1	3.2	2.4	2.0	4.4	3.2
6-7%		2.7	2.0	1.6	2.1	2.6	0.7	2.4	2.2
7-8%		2.9	0.5	1.0	3.2	2.6	4.7	1.9	2.7
8-9%		3.7	1.4	3.1	2.7	4.1	1.9	1.3	2.9
9-10%		3.5	0.5	3.0	5.7	1.5	1.9	3.2	3.2
10-20%		2.9	6.3	2.2	3.9	3.3	3.2	4.0	3.5
20-30%				0.9	3.6	2.2	2.3	2.5	2.6
30-40%					3.1	2.3	3.1	3.5	3.0
40-50%						1.5	2.5	4.2	2.7
50-100%					0.3	3.4	5.5	4.6	3.9
Total	2.0	2.1	2.7	2.1	2.7	2.3	2.2	2.8	2.3
The distribution of the Roma population by settlement size (%)	*15.0*	*19.0*	*5.0*	*5.4*	*20.7*	*18.7*	*7.8*	*8.2*	*100.0*

Source: BI calculations from NLO and 2011 Census data.

N=16,719 uneducated programme participants and 702,123 uneducated potential participants.

Note: the take up is No. of participants divided by the sum of the number of potential and actual participants. 0 indicates no participants in that category; an empty cell indicates either no settlements or no actual and potential participants in that category. Highlighted cells indicate the categories where at least 2% of the Roma population lives. All highlighted cells together cover 48% of the Hungarian Roma population.

According to settlement-level analysis, take up of the uneducated is the highest in less-than-600-inhabitant villages (2.8%), 2,000-5,000-inhabitant villages (2.7%) and 6,750-10,000-inhabitant cities (2.7%). However, these three categories cover only 34% of the entire Roma population in Hungary (see Table 22 and Table 7).

Figure 9: Take up of the uneducated in the One step ahead! programme
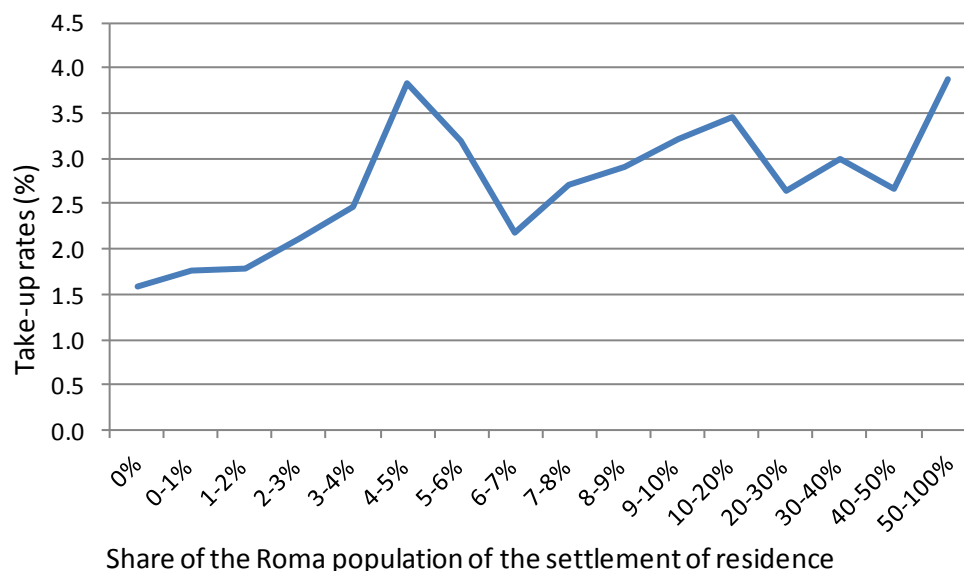
Share of the Roma population of the settlement of residence

Figure 10: Take-up rates of the uneducated participants of the One step ahead! I. and II. programmes by regions and by settlements
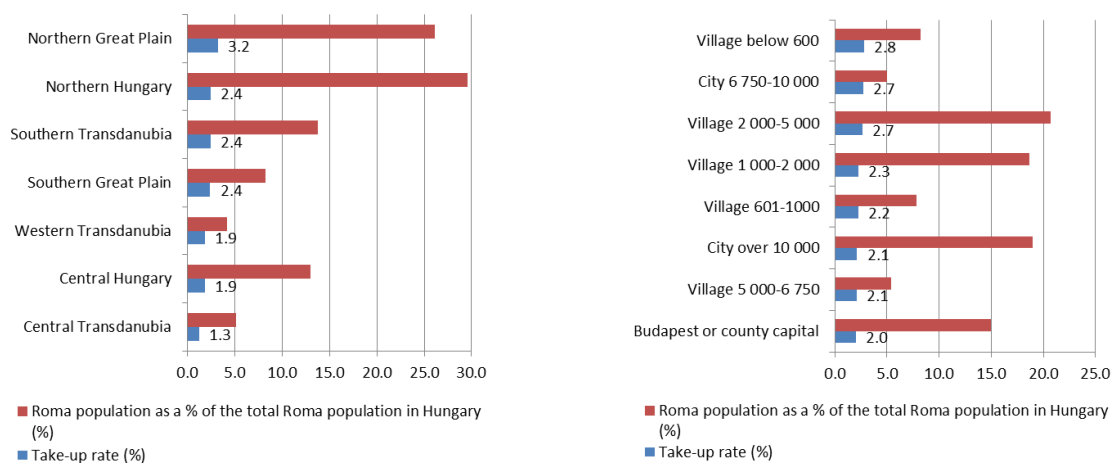
5.2.Selection and Roma inclusion based on NLO and Census data

As we indicated above, we have access to individual-level participation data covering 23,088 participants taking part in at least one phase of the programme. Just as in case of the SROP 1.1.2 programme, we do not have information about ethnicities neither of participants nor of potential participants, so we extended the database with publicly available settlement-level ethnicity data of the 2011 Census. 79% of participants lived in settlements where share of Roma in the population is not higher than 10%, while the same ratio in case of the potential participants is 87%. Share of participants living in settlements where share of the Roma is high is thus higher than that of potential participants, that is, probability of entering the programme is higher in these types of settlements. This is exactly what we would like to see if the programme covered the Roma effectively: as the share of the unemployed and uneducated is higher among the Roma, more participants should come from those settlements where the share of the Roma is higher. However, it is not clear how much the difference should exactly be.

Table 23. The distribution of participants versus potential participants of the One step ahead! programme as a function of share of Roma in the population

Share of the Roma in the settlement	All participants		Women		Men	
	Potential participants (%)	Partici-pants (%)	Potential participants (%)	Partici-pants (%)	Potential participants (%)	Partici-pants (%)
0%	6.8	4.4	6.8	4.0	6.9	4.8
0-1%	23.8	16.7	23.8	16.1	23.8	17.5
1-2%	20.3	15.2	20.5	14.6	20.1	15.9
2-3%	12.8	11.9	12.9	12.7	12.6	10.9
3-4%	7.4	8.0	7.3	8.3	7.5	7.6
4-5%	3.9	6.7	3.9	7.1	3.9	6.1
5-6%	3.6	5.4	3.6	5.2	3.6	5.6
6-7%	2.3	2.2	2.3	2.0	2.4	2.5
7-8%	3.0	3.9	3.0	4.0	3.1	3.8
8-9%	2.2	2.9	2.1	3.1	2.2	2.5
9-10%	1.2	1.9	1.2	2.0	1.2	1.9
Below 10%	87.2	79.1	87.3	79.0	87.2	79.1
10-20%	8.3	13.9	8.2	13.9	8.4	14.0
20-30%	2.6	3.8	2.6	3.8	2.6	3.9
30-40%	1.0	1.7	1.0	1.8	1.0	1.5
40-50%	0.4	0.6	0.4	0.6	0.4	0.5
50-100%	0.4	0.9	0.4	0.9	0.4	1.0
Total	100.0	100.0	100.0	100.0	100.0	100.0

Source: BI calculation based on NLO and 2011 Census data.
N=23,088 participants and 1.499,741 potential participants.

The difference between the distribution of participants and potential participants with respect to the share of the Roma population is about the same for men and women. Unlike in the case of the SROP 1.1.2 programme, here we do not observe gender-based differences.

5.2.1 Selection on settlement level

5.2.1.1. Correlation between the share of the Roma people and the number of programme participants on settlement level

In this section we analyse the correlation between the share of participants measured as a per cent of the non-employed active-age population and the share of Roma in the population. As Table 24. suggests, we do find the positive correlation we are looking for: a 1 %point higher share of Roma in the population increases share of participants with 0.07-0.11%points, depending on model specification. The correlation is heterogeneous through regions: it is positive in baseline Central Hungary, however, it turns to negative in Northern Hungary, the Northern Great Plain and Central Transdanubia.

Correlation is even more heterogeneous in smaller than 600-inhabitant villages. Share of women participants tends to be higher with a higher share of Roma in the population in Central-Hungary (+0.234%points with every %point increase), however, it turns to negative in Northern Great Plain (-0.078%points), Southern Great Plain (-0.326%points), Northern Hungary (-0.236%points) and Central Transdanubia (-0.226%points). Similarly to SROP 1.1.2, coverage in the most disadvantaged areas (small villages with a high share of Roma in poor regions) seems to be inefficient.

Table 24. Correlation between number of participants of the One step ahead! programme and share of Roma in the population – settlements over 600 inhabitants

	Women	Women	Men	Men
Proportion of Roma	0.0782***	0.159**	0.0760***	0.114**
	(0.0190)	(0.0641)	(0.0171)	(0.0473)
Square of the proportion of Roma	-0.0854	-0.0442	-0.0675	-0.0233
	(0.0553)	(0.0366)	(0.0543)	(0.0614)
Northern Great Plain (NGP)		0.0147***		0.00765***
		(0.00220)		(0.00271)
Southern Great Plain (SGP)		0.00721***		0.00320***
		(0.00145)		(0.00120)
Southern Transdanubia (ST)		0.00379**		0.00161
		(0.00161)		(0.00189)
Northern Hungary (NH)		0.00704***		0.00233
		(0.00147)		(0.00162)
Central Transdanubia (CT)		0.00174		0.000155
		(0.00129)		(0.00154)
Western Transdanubia (WT)		0.00133		0.000318
		(0.00125)		(0.00120)
Capital or county capital		-0.00164		-0.00143
		(0.00143)		(0.00157)
City with over 10,000 inhabitants		-0.00187		-0.00160
		(0.00145)		(0.00159)
Village, 5 000-6 750 inhabitants		-0.000354		-0.00331*
		(0.00192)		(0.00176)
Village, 2 000-3 000 inhabitants		0.000543		0.000374
		(0.00155)		(0.00167)
Village, 1 000-2 000 inhabitants		-0.000211		-0.000140
		(0.00149)		(0.00167)
Village, 601-1 000 inhabitants		0.00130		0.00161
		(0.00176)		(0.00189)
Share of Roma and NGP interaction		-0.117*		-0.0139
		(0.0669)		(0.0577)
Share of Roma and SGP interaction		-0.0531		-0.0634
		(0.0771)		(0.0524)
Share of Roma and ST interaction		-0.0879		-0.0826
		(0.0681)		(0.0560)
Share of Roma and NH interaction		-0.125*		-0.0855*
		(0.0657)		(0.0507)
Share of Roma and CT interaction		-0.121*		0.0290
		(0.0718)		(0.120)
Share of Roma and WT interaction		-0.0543		-0.00965
		(0.0747)		(0.0670)
Constant	0.00677***	0.00173	0.00609***	0.00386**
	(0.000511)	(0.00155)	(0.000423)	(0.00156)
Observations	1,929	1,927	1,929	1,928
R-squared	0.054	0.117	0.053	0.116

Robust standard errors in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$
Source: BI calculation based on NLO and 2011 Census data.
Note: The baseline categories are *Central Hungary* (regions), *Cities with 6,750 – 10,000 inhabitants* (settlement type), and the *interaction of the Share of the Roma and Central Hungary*.

Table 25. Correlation between number of participants of the One step ahead! programme and the share of Roma in the population – villages below 600 inhabitants

VARIABLES	Women	Women	Men	Men
Proportion of Roma	0.0804***	0.234**	0.0743***	-1.317
	(0.0240)	(0.102)	(0.0219)	(1.107)
Square of the proportion of Roma	-0.0898**	-0.0223	-0.100***	-0.0549
	(0.0352)	(0.0407)	(0.0330)	(0.0349)
Northern Great Plain (NGP)		0.0579***		0.0286
		(0.0158)		(0.0282)
Southern Great Plain (SGP)		0.0199**		0.00482
		(0.0101)		(0.0324)
Southern Transdanubia (ST)		0.0115***		-0.0190
		(0.00301)		(0.0229)
Northern Hungary (NH)		0.0200***		-0.00941
		(0.00725)		(0.0236)
Central Transdanubia (CT)		0.00342***		-0.0241
		(0.000957)		(0.0228)
Western Transdanubia (WT)		0.00897***		-0.0221
		(0.00188)		(0.0228)
Share of Roma and NGP interaction		-0.312**		1.265
		(0.126)		(1.109)
Share of Roma and SGP interaction		-0.326*		1.011
		(0.195)		(1.167)
Share of Roma and ST interaction		-0.174*		1.392
		(0.106)		(1.107)
Share of Roma and NH interaction		-0.236**		1.321
		(0.109)		(1.107)
Share of Roma and CT interaction		-0.226**		1.337
		(0.103)		(1.107)
Share of Roma and WT interaction		-0.0474		1.399
		(0.133)		(1.107)
Constant	0.0120***	-0.000743	0.0106***	0.0276
	(0.00179)	(0.000567)	(0.00167)	(0.0228)
Observations	1,249	1,249	1,249	1,249
R-squared	0.010	0.051	0.009	0.060

Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1
Source: BI calculation based on NLO and 2011 Census data.
Note: The baseline categories are *Central Hungary* (regions), *Cities with 6 750 - 10 000 inhabitants* (settlement type), and the *interaction of the Share of the Roma and Central Hungary* (interaction of the share of the Roma and the region).

5.2.1.2. Settlements with and without participants in the programme

From the 3,176 Hungarian settlements, at least one woman participated in the programme in 1,229 settlements and at least one man in 1,107 settlements (see Table 26). Comparing to SROP 1.1.2, number of settlements reached by the programme is only half as much. Obviously, the number of participants was about half as much in the *One step ahead!* programme as well. Those settlements, from where no one participated in the programme, are on average smaller and share

of the Roma in the population is higher. Again, comparing to the coverage of SROP 1.1.2, while in the 1.1.2 the average size of non-reached settlements was around 300, in this case it is about 1,000 inhabitants. As the data suggests, if a programme is bigger in size, it can reach not simply more people but smaller settlements as well. This is important, because 16% of the Hungarian Roma population lives in villages under 1,000 inhabitants (see Table 22).

Total population of those settlements, where no one participated in the programme, is 1,131,791 persons in case of women and 1,246,438 persons in case of men. Number of Roma population in these settlements is 49,552 and 53,400, respectively. Comparing to 2011 Census ethnicity data, this is about 15.9-16.9% of the entire Roma population.

Table 26. Comparison of settlements with and without participants

	Settlements with participants			Settlements without participants		
	No. of settlements	Mean	Standard deviation	No. of settlements	Mean	Standard deviation
Women						
Population No.	1706	5162	15 247	1470	770	947
Share of Roma people (%)	1706	3.0	5.1	1470	4.4	9.1
Men						
Population No.	1578	5508	15799	1598	780	979
Share of Roma people (%)	1578	3.0	5.2	1598	4.3	8.4

Source: BI calculations using NLO and 2011 Census data.

Figure 11 compares the distribution of share of Roma in the population of settlements covered and not-covered by the programme. We examined settlements above and below 600 inhabitants separately. As we have already seen in the analysis of SROP 1.1.2 (see Figure 2 and Figure 1), the 45-degree line is the theoretical situation in which the distribution of the Roma is the same in settlements reached and not reached by the programme. Points below the line represent that the share of Roma population is higher in non-participant settlements than in participant settlements, and vice versa.

In settlements with population less than 600 the proportion of Roma in the population is slightly higher in participating settlements (see Figure 11). As pointed out earlier, efficient targeting of the Roma would imply the pattern we observe here: due to overrepresentation of Roma people in

the target group we expect participating settlements having a higher share of Roma in their population.

Figure 11: Distribution of the Roma in the population of settlements with and without participants in the One step ahead! programme – villages under 600 inhabitants

Source: BI calculations using NLO and 2011 Census data.

In case of settlements with more than 600 inhabitants, the difference of distributions is smaller. Again, Roma women might have been more likely to be left out of while men seem to be more likely to be involved in the programme in settlements with a very high share of Roma in the population (Figure 12).

Figure 12: Distribution of the Roma in the population of settlements with and without participants –
settlements over 600 inhabitants
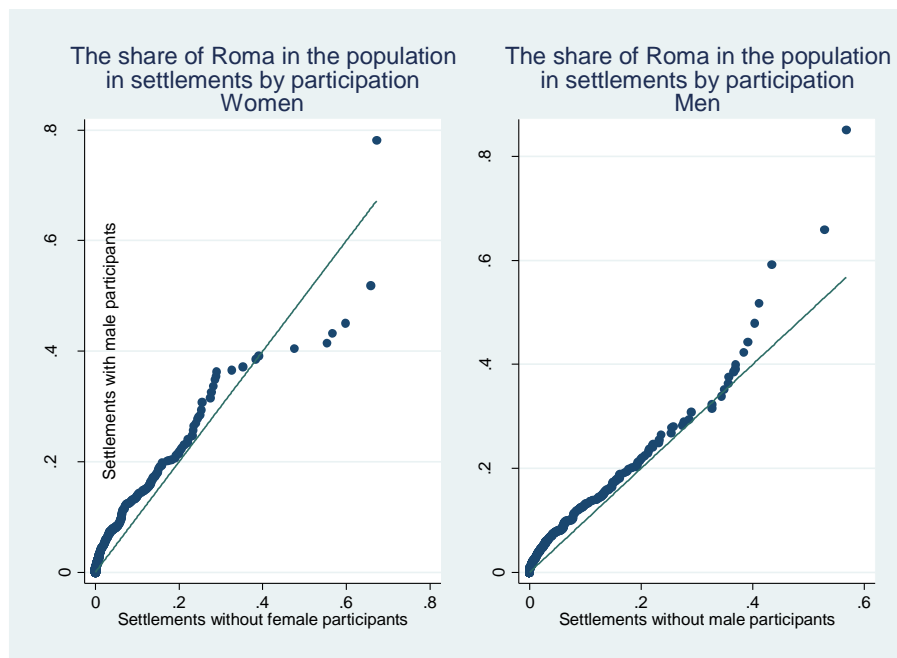
5.2.1.3. Correlation between the share of the Roma and number of participants using alternative measures

Just as before, correlation between share of Roma in the population and share of programme participants has been investigated by two additional measures as well. Again, we construct the following two binary variables:

Variable 1:

> It takes value 1, if share of Roma in the population is higher than weighted country average (3.2 %) and 0 otherwise;

Variable 2:

> It takes value 1, if share of Roma in the population is higher than weighted country median (1.3 %) and 0 otherwise.

Then we estimate the same models as in Section 5.2.1.1 using one of these new binary variables as explanatory variables instead of the continuous measure of share of Roma in the population. As usually, we run estimations separately for settlements with fewer and more than 600 inhabitants.

Results are summarized Table 54 and Table 55 of Appendix 2.5. In case of above-600 settlements the share of participants is higher when share of Roma is above mean or median. However, although the difference is significant, it is small: share of participants in settlements with above-mean or above-median share of Roma in the population is 0.5-0.8%points higher.

5.2.2. Selection on individual-level data

In this chapter we examine whether probability of participation is correlated with share of Roma in the population of home settlement using individual-level data. We compare all participants that can be linked to the unemployment registry (16,360) to those eligible for but left out from the programme 1,499,741. As before, those being eligible are selected from the unemployment registry, and they are either uneducated, or have vocational training.

The following selection model was estimated:

$$P(Participation = 1|X) = X'\beta$$

The dependent variable of the model[38] is a binary variable taking value of 1 if the individual participated in the programme, and 0 if not. X refers to the following explanatory variables:

- share of Roma in the population of home settlement;
- settlement type;
- region;
- interaction term of share of Roma in the population and region;
- education level;
- whether the individual is disabled; and
- age.

---

[38]Linear probability model using standard errors clustered at sub-regional level to control for regionally different random shocks.

Table 27. Correlation between the probability of programme participation and the share of Roma in the population – settlements over 600 inhabitants

	Women	Women	Men	Men
Proportion of Roma	0.139***	0.445***	0.0651***	0.203***
	(0.0392)	(0.132)	(0.0192)	(0.0699)
Square of proportion of Roma	-0.176	-0.0395	-0.0676	-0.000644
	(0.108)	(0.0602)	(0.0455)	(0.0552)
Northern Great Plain (NGP)		0.0204***		0.00586**
		(0.00372)		(0.00238)
Southern Great Plain (SGP)		0.0131***		0.00162
		(0.00320)		(0.00185)
Southern Transdanubia (ST)		0.0117***		0.00279
		(0.00351)		(0.00236)
Northern Hungary (NH)		0.0193***		0.00453*
		(0.00360)		(0.00237)
Central Transdanubia (CT)		0.00412*		-7.05e-05
		(0.00245)		(0.00187)
Western Transdanubia (WT)		0.00746***		0.00254
		(0.00250)		(0.00213)
Secondary education		-0.0198***		-0.0151***
		(0.00140)		(0.000929)
Higher education		0.968***		0.978***
		(0.00241)		(0.00159)
Capital or county capital		-0.00216		-0.000646
		(0.00288)		(0.00212)
City with over 10,000 inhabitants		-0.00308		-0.00152
		(0.00294)		(0.00198)
Village, 5 000-6 750 inhabitants		-0.00229		-0.00477**
		(0.00357)		(0.00215)
Village, 2 000-3 000 inhabitants		-0.000603		-9.64e-05
		(0.00298)		(0.00182)
Village, 1 000-2 000 inhabitants		-0.00399		-0.00145
		(0.00299)		(0.00213)
Village, 601-1 000 inhabitants		-0.00331		1.51e-05
		(0.00364)		(0.00241)
Disability		-0.0164***		-0.00638***
		(0.00134)		(0.00116)
Age		-0.000496***		-0.000315***
		(4.62e-05)		(2.02e-05)
Share of Roma and NGP interaction		-0.414***		-0.166**
		(0.137)		(0.0742)
Share of Roma and SGP interaction		-0.214		-0.139*
		(0.151)		(0.0767)
Share of Roma and ST interaction		-0.362**		-0.198***
		(0.141)		(0.0755)
Share of Roma and NH interaction		-0.443***		-0.207***
		(0.135)		(0.0717)
Share of Roma and CT interaction		-0.395***		-0.130
		(0.148)		(0.0991)
Share of Roma and WT interaction		-0.289**		-0.106
		(0.146)		(0.0838)
Constant	0.0153***	0.0352***	0.00869***	0.0279***
	(0.00138)	(0.00367)	(0.000613)	(0.00303)

Observations	602,520	602,518	824,283	824,281
R-squared	0.002	0.131	0.001	0.099

Clustered robust standard errors at the sub-regional level are in parentheses.[39] *** p<0.01, ** p<0.05, * p<0.1

Source: BI estimation from NLO and 2011 Census data

Results are similar than those from the earlier settlement-level analysis. In settlements over 600 participants a 1%point increase of share of Roma in the population raises probability of participation by 0.14 (women)-0.07(men) %points. (Table 27) Controlling for additional individual characteristics and allowing for correlation to be heterogeneous through regions reveals that estimated coefficients are much lower and may also turn to negative in the most disadvantaged regions, in Northern Hungary and Northern Great Plain. Effects are similar in case of small villages as well (Table 28).

Table 28: Correlation between the probability of programme participation and the share of the Roma population – villages under 600 inhabitants

	Women	Women	Men	Men
Proportion of Roma	0.113***	1.385**	0.0730***	-0.766*
	(0.0420)	(0.627)	(0.0171)	(0.388)
Square of proportion of Roma	-0.102**	-0.00227	-0.0873***	-0.0488**
	(0.0501)	(0.0536)	(0.0263)	(0.0241)
Northern Great Plain (NGP)		0.0901**		0.0334
		(0.0345)		(0.0228)
Southern Great Plain (SGP)		0.0325**		-2.63e-06
		(0.0130)		(0.0118)
Southern Transdanubia (ST)		0.0231**		-0.00596
		(0.00912)		(0.00797)
Northern Hungary (NH)		0.0320***		-0.00271
		(0.00901)		(0.00801)
Central Transdanubia (CT)		0.0147**		-0.00976
		(0.00693)		(0.00775)
Western Transdanubia (WT)		0.0202***		-0.00645
		(0.00704)		(0.00777)
Secondary education		-0.0256***		-0.0175***
		(0.00351)		(0.00246)
Higher education		0.950***		0.968***
		(0.0181)		(0.00761)
Disability		-0.0140**		-0.000360
		(0.00620)		(0.00544)
Age on the day when entered the programme (Aug 1, 2007)		-0.000656***		-0.000521***
		(0.000119)		(6.16e-05)
Share of Roma and NGP interaction		-1.482**		0.721*
		(0.635)		(0.391)

[39] When using clustered robust standard errors at the ZIP-code level we get an insignificant coefficient estimation for the share of Roma population variable in the model for women with controls.

Share of Roma and SGP interaction		-1.431**		0.716*
		(0.639)		(0.403)
Share of Roma and ST interaction		-1.307**		0.830**
		(0.630)		(0.388)
Share of Roma and NH interaction		-1.408**		0.775**
		(0.628)		(0.388)
Share of Roma and CT interaction		-1.418**		0.792**
		(0.627)		(0.388)
Share of Roma and WT interaction		-1.185*		0.820**
		(0.631)		(0.388)
Constant	0.0181***	0.0289***	0.00928***	0.0428***
	(0.00387)	(0.00867)	(0.00193)	(0.00855)
Observations	40,476	40,476	55,506	55,506
R-squared	0.003	0.089	0.002	0.076

Clustered robust standard errors at the sub-regional level are in parentheses.[40] *** $p<0.01$, ** $p<0.05$, * $p<0.1$
Source: BI estimation from NLO and 2011 Census data.

5.3. Employment impact of the programme

In this chapter we conduct a counterfactual impact evaluation of the programme.[41] As we indicated earlier, from the 23 088 participants we can link 16 360 to the unemployment registry and to the SHLD database. Comparing raw employment rates of uneducated participants (10 584 persons) and potential participants suggests that starting from early 2008 programme participants are more likely to find employment (Figure 13), but there is no difference in the share of those entering unemployment registry (

Figure 14). However, comparing raw outcomes cannot be identified as the causal effect of the programme as samples of actual and potential participants are not balanced with respect to distribution of individual characteristics.

---

[40] When using clustered robust standard errors at the ZIP-code level we get an insignificant coefficient estimation for the share of Roma population variable in the model for women with controls.

[41] We evaluate the impact of the One step ahead! II. (SROP 2.1.1) programme only.

Figure 13: Employment entry rates of uneducated actual and potential participants of SROP 2.1.1
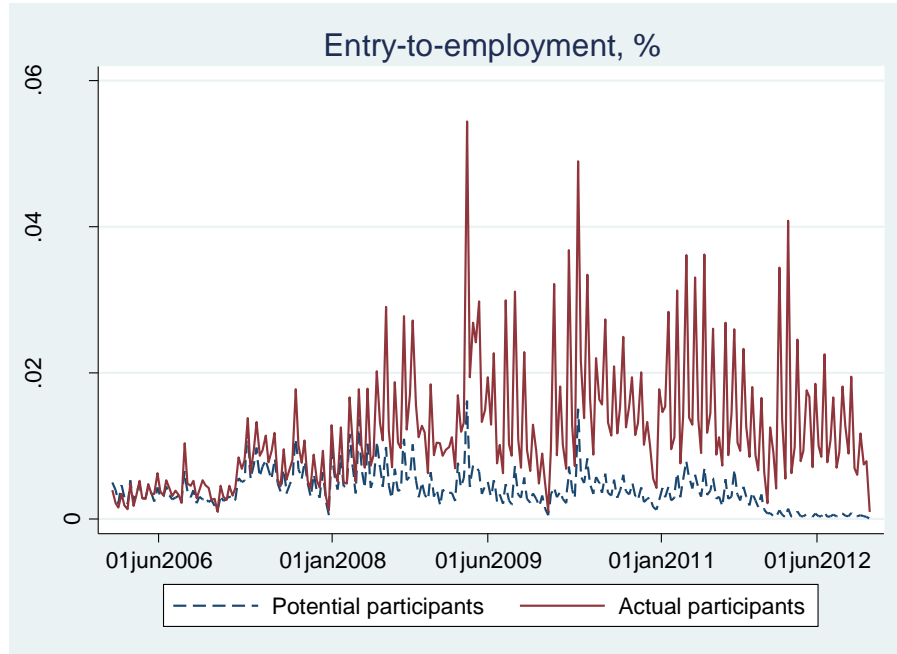
Source: BI calculations using NLO data.
N=10 584 uneducated programme participants and 702 123 uneducated potential participants. Measure: the number of those exiting to employment in 10-day intervals, as a % of the sample size.

To identify the impact of the programme we construct a comparable control group using a combination of propensity score and nearest neighbour matching, just as before.[42] Then, we identify the causal effect of the programme as the difference of factual and counterfactual outcomes. To construct the control group, we narrow down the group of uneducated participants to those who

- participated in the second phase of the programme, in SROP 2.1.1, and

- have not participated in any other programmes.

---

[42] See the exact methodology in Appendix 1

Figure 14: Unemployment entry rates of uneducated actual and potential participants of SROP 2.1.1

Then, using the same procedure as detailed in Appendix 1, we match exactly one control pair to all participants in the final sample. We set up pairs such that:

- they entered unemployment registry within the same +/- 60-day period;

- control pair is still unemployed when treat enters the programme;

- they are in the same age range;

- they have the same sex;

- they have the same educational attainment: they both are drop-outs or they both finished elementary school at most;

- they live in the same region and in a same type of settlement;

- their estimated probability of participation (propensity score) is as close to each other within the intersection of all categories above as possible.

Using this method we matched 6 720 controls to 6 720 programme participants. Due to matching procedure, sample is balanced in terms of all observable characteristics. (see Table 62 in Appendix 2.6) Therefore, we define the causal effect of the programme as the difference of outcomes in treated and control group. We set up the same 5 outcome variables as before (see details in chapter 4.4).

Table 29: The employment effect of the SROP 2.1.1 programme

MEN	Control		Treated		Effect
	No.	%	No.	%	%ps
Exit to employment during the programme or within 6 months afterwards	892	31.0	2050	71.4	40.3
Exit to employment within 6 months after participating in the programme	44	1.5	920	32.0	30.5
Exit to employment anytime during the observation period	980	34.1	2471	86.0	51.9
No re-entering into unemployment within 6 months after the programme	794	27.6	1686	58.7	31.0
No re-entering into unemployment anytime after the programme during the observation period	1173	40.8	2049	71.3	30.5
No. of observations	2873		2873		
WOMEN	Control		Treated		Effect
	No.	%	No.	%	%ps
Exit to employment during the programme of within 6 months afterwards	894	23.2	2213	57.5	34.3
Exit to employment within 6 months after participating in the programme	56	1.5	926	24.1	22.6
Exit to employment anytime during the observation period	1016	26.4	2820	73.3	46.9
No re-entering into unemployment within 6 months after the programme	1231	32.0	2054	53.4	21.4
No re-entering into unemployment anytime after the programme during the observation period	1780	46.3	2773	72.1	25.8
No. of observations	3847		3847		

Source: BI estimation using NLO data.

Results are presented in Table 29. Depending on which outcome variable is used, 24-86% of participants found employment, while these ratios in the matched control group are only 2-34%.

Most participants found a job during the programme. The programme increased the probability of finding a job in the medium term by 47-52%points.

Table 30: The effect of the programme – the long term unemployed

MEN	Control		Treated		Effect
	No.	%	No.	%	%ps
Exit to employment during the programme of within 6 months afterwards	285	25.7	849	68.1	42.3
Exit to employment within 6 months after participating in the programme	32	2.9	402	32.2	29.3
Exit to employment anytime during the observation period	323	29.2	1060	85.0	55.8
No re-entering into unemployment within 6 months after the programme	244	22.0	751	60.2	38.2
No re-entering into unemployment anytime after the programme during the observation period	406	36.7	932	74.7	38.1
No. of observations	1107		1247		
WOMEN	Control		Treated		Effect
	No.	%	No.	%	%ps
Exit to employment during the programme of within 6 months afterwards	308	19.0	978	56.9	37.9
Exit to employment within 6 months after participating in the programme	19	1.2	414	24.1	22.9
Exit to employment anytime during the observation period	359	22.1	1250	72.7	50.6
No re-entering into unemployment within 6 months after the programme	409	25.2	934	54.3	29.1
No re-entering into unemployment anytime after the programme during the observation period	680	41.9	1262	73.4	31.5
No. of observations	1624		1720		

Source: BI estimations using NLO data.

Outcome variables based on non-re-entering unemployment again show a smaller effect. Data suggest that 53-72% of participants left unemployment registry and did not register again, comparing to 28-46% of controls. Measuring programme effect using employment-based outcome variables gives a 23-52-%point impact while using non-re-entering-unemployment-based outcome variables results in a 21-31-%point estimated effect. As we argued in case of the SROP 1.1.2 programmes, employment-based measures probably overestimate while

unemployment-based measures probably underestimate the effect of the programme because controls are more likely to work in the black economy. The real impact of the programme could be somewhere between, and they are of gross nature as we cannot correct for substitution effect and deadweight loss.

The magnitude of the effects is similar even if we restrict the sample to long term unemployed participants. The share of those who had been unemployed for at least a year when they entered the programme is 43-44%. Among them 68% of men and 57% of women found a job during the programme or in the following 6 months. The same ratios in the control group were 25-19%, respectively, so the programme increased probability of entering employment with 42-38%. Just as before, effects are lower if unemployment-entry is used as outcome variable.

6. REFERENCES

ÁSZ (2008): A magyarországi cigányság helyzetének javítására és felemelkedésére a rendszerváltás óta fordított támogatások mértéke és hatékonysága. Állami Számvevőszék, April, 2008.
http://www.asz.hu/ASZ/tanulmanyok.nsf/0/79ED5720B293EBC9C12574F30031B5C7/$File/t206.pdf

Csite et al (2013): Foglalkoztathatóság javítását szolgáló intézkedések értékelése. http://www.nfu.hu/foglalkoztathatosag_javitasat_szolgalo_intezkedesek_ertekelese

Havas, G. (2008): Esélyegyenlőség, deszegregáció. In: Zöld könyv a magyar közoktatás megújításáért, ed.: K. Fazekas, J. Köllő and J. Varga. Ecostat, Budapest.

Janky, B. (2004): A cigány családok jövedelmi helyzete. In: Társadalmi riport 2004, ed.: T. Kolosi, I. Gy. Tóth, Gy. Vukovich. Budapest: TÁRKI, pp. 400–413

Kemény, I. and B. Janky (2003): A 2003. évi cigány felmérésről – Népesedési, nyelvhasználati és nemzetiségi adatok. Beszélő, October 2003, pp. 64–76

Kemény, I., B. Janky and G. Lengyel (2004): A magyarországi cigányság 1971–2003. Budapest: Gondolat.

http://www.mtaki.hu/kiadvanyok/kemeny_janky_lengyel_moi_ciganysag_main.html

Kertesi G. (2005): A társadalom peremén – Romák a munkaerőpiacon és az iskolában. Budapest, 2005, Osiris Kiadó. pp. 533

Kertesi, G. (2005): Roma foglalkoztatás az ezredfordulón. Szociológiai Szemle, 2005/2, pp. 57-87

Ladányi, J. and I. Szelényi (2002): Cigányok és szegények Magyarországon, Romániában és Bulgáriában. Szociológiai Szemle, 2002/4. pp. 72–94

Messing, V. (2011): Megjegyzések egy „roma" projekt margójára. Online Journal of HAS Sociology Institute.

Scharle, Á. (2011): Áttekintés a roma népesség helyzetéről és annak okairól. Budapest Intézet, February 2011

Váradi, L. (2012): Előtanulmány a roma családnevek diszkriminációteszteléshez való kiválasztásához, In: A diszkrimináció mérése, ed.: E. Sik and A. Simonovits. E-tankönyv, ELTE TTK, Budapest
 http://www.tarki.hu/hu/about/staff/sb/Diszkriminacio_merese.pdf

Vokó Z., Zs. Kósa, Gy. Széles, L. Kardos, K. Kósa, R. Németh, S. Országh, G. Fésüs and és R. Ádány (2006): A roma telepeken élők egészségének felmérése. Presented on the 15th meeting of NETT, Siófok

APPENDIX 1   PROGRAMME EVALUATION METHODOLOGY

Our aim in this research is to measure the causal effect of the two programmes on selected employment/unemployment variables. As the base of comparison, we would like to answer the following question: what would have happened to the outcome of programme participants if they had not participated in the programme? Theoretically, in a thought experiment, we would like to compare the outcome of the same individuals under two different scenarios: participating and not participating in the programme. Obviously, this is impossible, because an individual either participates or not. Thus, to estimate the effect of the programme we construct a theoretical outcome, which shows us what would have happened to the programme participants without participating in the programme. This theoretical outcome is the *counterfactual outcome*, and the procedure is sometimes referred as *counterfactual impact (programme) evaluation*.

The credibility of programme evaluation lies in the construction of the counterfactual. We need to create a control group that consists of individuals who are not significantly different in their observable and unobservable characteristics from programme participants, most importantly in those aspects that might have a direct effect on the outcome variable of interest. Difference in unobservable characteristics cannot be measured; we have to assume that they do not differ in those.

We created control groups for the evaluation using a combination of 1:1 direct and propensity score matching. In particular, we matched exactly one control pair to each examined programme participant based on all their observable characteristics. Thus, we created "statistical twins" in the sense that the members of these pairs are as similar to each other as possible, given the information in the dataset we had access to. In the next two sections we explain the matching procedure used in the evaluation of the two programmes.

1. The matching strategy of the SROP 1.1.2 programme

In this procedure we match exactly one control pair to each participant in the sample. As it is indicated in section 4.4, we evaluated the impact of the SROP 1.1.2 programme on those male, uneducated participants who did not participated in any other programmes only, and we constructed a pool of potential controls the same way from the unemployment registry as a first

step. Then, we created 9 Strata of the jobseekers. In Stratum 1, we separated those not completing elementary education. To Strata 2-9, we assigned the rest of the sample, those who completed elementary school at most, based on the size of their home settlement. Within each Strata we estimated a propensity score (programme participation probability) variable based on the following probit model:

$$P(\text{Participation} = 1|X) = \Phi(X'\beta),$$

where matrix X includes the following explanatory variables, their squares and interactions:

- share of Roma in the population of home settlement;

- sub-regional unemployment rate in 2008;

- region[43];

- days spent as being registered unemployed until the first day of the last unemployment spell before entering the programme;

- number of register entries until the first day of the last unemployment spell;

- number of days spent in employment until the first day of the last unemployment spell, if the employment spell was at least 30-day long;

- alternative work history indicator: number of days spent being employed in the percentage of the number of days between the 16th birthday and the beginning of the registry period;

- age on the day of the last unemployed register entry before the programme.

We specified the propensity score model in each Stratum (Strata 1-9) separately in a way to make sure that balancing conditions are fulfilled, thus, the distributions of explanatory variables in the treatment and control group does not differ significantly. Our aim was to have the same distribution of observable attributes of participant and non-participant individuals who have the

---

[43] In Group 1 also settlement size category.

same propensity score. We tested the balance of the samples within each Stratum 1-9 by dividing the treatment and the control groups to sub-strata, in which the mean of propensity scores were the same across the two groups. Then, we compared the distribution of all explanatory variables from the treatment model in the treatment and control groups in all sub-strata using a t-test. In case of each Stratum (Strata 1-9) we accepted the propensity score model if and only if we could not reject the null hypothesis of equal averages within each sub-strata on a 5 percent significance level. If a model did not pass this balance test we expanded the model with powers and interactions of the explanatory variables. At the end of the estimation procedure we kept observations from the common support only.

In the second step we matched a control pair to each participant directly 1:1 based both on their estimated propensity scores and other observable characteristics. In particular, we matched pairs within Strata 1-9 one-by-one such that

- both members of the pairs entered unemployment registry within the same +/- 60-day period;

- the control pair is still unemployed when treat pair enters the programme;

- they are in the same age range;

- they have the same educational attainment: they both are drop-outs or they both finished elementary school at most.

Then, we chose that particular control pair from the set of possible controls fulfilling all these criteria above, whose estimated propensity score was the closest to the propensity score of the participant. We applied no replacement, so one control pair could have been the pair of one participant only, and vice versa. However, as a robustness check, we repeated the procedure allowing replacement as well, and we received the same results.

Further detail of the procedure that the NLO database includes several unemployment spells for each observed individual, and the date of unemployment entry is one of our matching variables. In the case of programme participants we used the last unemployment entry date before entering the programme in the matching procedure, thus, that we had one row for one person in the data.

However, in the case of potential controls, we did not know in advance the programme entry date of their pairs, so when we chose among controls we chose not only from persons but also from different unemployment spells of the same individuals. Practically, among potential controls, one observation meant one unemployment spell and not one person, as it did in the case of participants. This procedure ensures that all controls were chosen with that particular unemployment spell of his which was compatible with one of the participants.

If we could not find an appropriate pair for a participant using this procedure, we left out that participant form the evaluation sample. We could match a control pair to 6,946 participants out of the total 7,049, so only 103 participants were left out because of having no proper pair.

Finally, we received a sample of 6,946 uneducated programme participants and the same amount of non-participant controls. The distributions of before-program observable characteristics in the two groups are not significantly different, so the outcomes of the two groups are comparable (see Appendix 2.3, Table 46).

2. The matching strategy of the SROP 2.1.1 programme

We used a similar matching strategy in the evaluation of the SROP 2.1.1 programme as in the case of the SROP 1.1.2. Again, we matched exactly one control pair to each programme participant. The evaluation sample of the programmes consists of those who participated in the second phase of the *One step ahead!* programme, namely the SROP 2.1.1, and has not participated in any other programmes during the observation period (7 226 persons).

As a first step we created 2*3*8*7=336 Strata in the sample, based on the following individual characteristics of the observations: gender (2 categories), age group (3 categories: below 25, 25-50, above 50), type of settlement (8 categories), and region (7 categories). Then, within each Stratum, we estimated a propensity score variable the same way as detailed in the previous section based on the following information:

- share of Roma in the population of home settlement;

- sub-regional unemployment rate in 2008;

- days spent as being registered unemployed until the first day of the last unemployment spell before entering the programme;

- number of register entries until the first day of the last unemployment spell;

- number of days spent in employment until the first day of the last unemployment spell, if the employment spell was at least 30-day long;

- alternative work history indicator: number of days spent being employed in the percentage of the number of days between the 16th birthday and the beginning of the registry period;

- age on the day of the last unemployed register entry before the programme.

Again, we accepted estimation results in case of completely fulfilled balancing criteria and re-specified the model if not, and we kept observations from the common support only. As a second step, we matched exactly 1 control pair to each participant such a way that:

- both members of the pairs entered the unemployment registry within the same +/- 60-day period;

- the control pair was still unemployed when the treated entered the programme;

- they had the same educational attainment: they both were drop-outs or they both finished elementary school at most.

Just as before, from all the potential pairs fulfilling these criteria, we chose the one with the closest propensity score to that of the participant.

Using this method we matched 6 720 controls to 6 720 programme participants. Due to the matching procedure, the sample is balanced in terms of all observable pre-program characteristics (see Table 62 in Appendix 2.6).

APPENDIX 2   TABLES

2.1. The distribution of the participants and potential participants of SROP 1.1.2 programme by the Share of Roma in the population at the settlement of residence, tables by regions

Table 31. The distribution of the participants and potential participants of SROP 1.1.2 programme by the Share of Roma in the population at the settlement of residence, Northern Great Plain

	Full Sample		Women		Men	
Proportion of Roma people in the settlement	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	2.9	3.2	3.0	3.3	2.9	3.2
0-1%	16.6	19.8	17.0	20.2	16.3	19.4
1-2%	9.2	15.7	9.3	16.5	9.0	14.9
2-3%	14.9	15.3	15.2	15.5	14.6	15.1
3-4%	7.6	6.9	7.7	7.5	7.6	6.3
4-5%	6.6	6.1	6.4	6.1	6.8	6.2
5-6%	4.4	4.1	4.4	4.0	4.4	4.2
6-7%	2.0	2.0	2.0	2.0	2.0	1.9
7-8%	7.1	6.0	7.0	5.9	7.2	6.2
8-9%	2.2	1.8	2.1	1.5	2.2	2.1
9-10%	1.6	1.3	1.6	1.2	1.6	1.4
Total Below 10%	75.1	82.3	75.7	83.6	74.5	80.9
10-20%	16.8	11.3	16.4	10.8	17.2	11.9
20-30%	5.8	5.4	5.6	4.9	5.9	6.0
30-40%	1.5	0.8	1.4	0.7	1.5	1.0
40-50%	0.1	0.1	0.1	0.1	0.1	0.0
Above 50 %	0.8	0.1	0.8	0.1	0.9	0.2
Total	100	100	100	100	100	100

Table 32. The distribution of the participants and potential participants of SROP 1.1.2 programme by the Share of Roma in the population at the settlement of residence, Southern Great Plain

Proportion of Roma people in the settlement	Full Sample		Women		Men	
	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	6.8	5.2	6.7	5.2	7.0	5.1
0-1%	32.2	30.5	32.1	30.7	32.2	30.3
1-2%	26.5	29.3	26.9	29.3	26.2	29.4
2-3%	9.8	8.1	9.8	8.7	9.7	7.4
3-4%	6.2	7.2	6.1	7.3	6.2	7.1
4-5%	3.6	4.5	3.6	4.5	3.6	4.6
5-6%	7.0	6.3	7.2	6.7	6.9	5.7
6-7%	4.0	5.0	3.9	3.9	4.0	6.3
7-8%	0.5	0.3	0.4	0.4	0.5	0.3
8-9%	1.2	0.9	1.1	0.9	1.2	0.9
9-10%	1.5	2.3	1.4	2.1	1.5	2.6
Total Below 10%	99.1	99.6	99.1	99.6	99.0	99.6
10-20%	0.8	0.3	0.8	0.3	0.8	0.3
20-30%	0.1	0.1	0.1	0.1	0.1	0.1
30-40%						
40-50%						
Above 50 %						
Total	100	100	100	100	100	100

Table 33. The distribution of the participants and potential participants of SROP 1.1.2 programme by the Share of Roma in the population at the settlement of residence, Central Hungary

Proportion of Roma people in the settlement	Full Sample		Women		Men	
	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	1.8	2.0	1.7	2.0	1.9	2.0
0-1%	32.8	34.0	33.2	35.8	32.5	31.7
1-2%	39.7	39.1	39.4	38.6	40.0	39.8
2-3%	12.1	12.0	12.3	11.4	12.0	12.7
3-4%	7.2	6.8	7.1	6.6	7.3	7.1
4-5%	2.5	2.9	2.6	2.8	2.5	3.0
5-6%	0.8	0.7	0.8	0.5	0.8	0.9
6-7%	0.4	0.3	0.4	0.2	0.4	0.4
7-8%	0.4	0.5	0.4	0.4	0.5	0.6
8-9%	0.1	0.2	0.1	0.2	0.1	0.2
9-10%	0.3	0.2	0.2	0.2	0.3	0.1
Total Below 10%	98.1	98.6	98.1	98.8	98.1	98.3
10-20%	1.5	1.1	1.4	0.9	1.5	1.3
20-30%	0.4	0.3	0.4	0.3	0.4	0.4
30-40%						
40-50%						
Above 50 %						
Total	100	100	100	100	100	100

Table 34. The distribution of the participants and potential participants of SROP 1.1.2 programme by the Share of Roma in the population at the settlement of residence, Southern Transdanubia

Proportion of Roma people in the settlement	Full Sample		Women		Men	
	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	5.0	4.3	4.9	4.5	5.0	4.1
0-1%	5.3	4.9	5.3	5.6	5.2	4.2
1-2%	12.8	13.2	13.1	14.4	12.6	11.9
2-3%	27.7	29.5	28.3	30.6	27.2	28.5
3-4%	8.9	12.1	8.8	11.9	9.0	12.4
4-5%	6.7	7.7	6.9	7.2	6.5	8.1
5-6%	3.7	3.6	3.7	3.4	3.7	3.8
6-7%	3.4	3.4	3.3	3.4	3.5	3.3
7-8%	3.8	3.2	3.6	2.8	3.9	3.6
8-9%	5.4	5.1	5.2	5.0	5.5	5.2
9-10%	0.7	0.6	0.7	0.5	0.7	0.6
Total Below 10%	83.3	87.3	83.9	89.1	82.7	85.6
10-20%	9.2	7.5	9.0	6.6	9.5	8.4
20-30%	4.4	3.5	4.2	3.0	4.6	4.0
30-40%	1.2	0.8	1.1	0.7	1.3	0.9
40-50%	0.4	0.2	0.4	0.1	0.5	0.4
Above 50 %	1.5	0.6	1.5	0.5	1.6	0.8
Total	100	100	100	100	100	100

Table 35. The distribution of the participants and potential participants of SROP 1.1.2 programme by the Share of Roma in the population at the settlement of residence, Northern Hungary

	Full Sample		Women		Men	
Proportion of Roma people in the settlement	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	4.8	5.2	4.9	5.7	4.8	4.7
0-1%	5.2	6.0	5.3	6.8	5.0	5.3
1-2%	6.6	7.7	6.7	7.7	6.5	7.6
2-3%	7.4	9.5	7.7	9.4	7.2	9.7
3-4%	14.9	16.9	15.0	19.0	14.7	14.8
4-5%	2.4	2.7	2.4	2.7	2.4	2.8
5-6%	6.2	6.8	6.3	6.9	6.2	6.7
6-7%	3.7	4.7	3.7	4.7	3.8	4.8
7-8%	5.8	4.8	5.8	4.7	5.8	4.8
8-9%	5.1	5.1	5.1	4.6	5.0	5.7
9-10%	1.6	1.5	1.6	1.5	1.6	1.5
Total Below 10%	63.6	71.1	64.4	73.7	62.9	68.4
10-20%	22.2	19.6	21.8	18.3	22.5	20.9
20-30%	6.5	4.9	6.4	4.2	6.7	5.6
30-40%	4.5	2.7	4.4	2.4	4.6	3.0
40-50%	2.1	1.2	2.0	1.1	2.2	1.4
Above 50 %	1.1	0.5	1.0	0.4	1.1	0.7
Total	100	100	100	100	100	100

Table 36. The distribution of the participants and potential participants of SROP 1.1.2 programme by the Share of Roma in the population at the settlement of residence, Central Transdanubia

Proportion of Roma people in the settlement	Full Sample		Women		Men	
	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	10.7	10.3	10.6	9.9	10.7	10.8
0-1%	42.7	45.1	42.5	46.1	43.0	43.7
1-2%	24.8	25.6	25.5	25.7	24.2	25.3
2-3%	11.5	11.5	11.5	11.2	11.6	12.0
3-4%	1.9	1.6	1.8	1.2	1.9	2.1
4-5%	0.9	0.9	0.9	0.8	0.9	1.0
5-6%	0.7	0.7	0.7	0.9	0.7	0.4
6-7%	1.9	1.4	1.9	1.5	2.0	1.4
7-8%	0.6	0.3	0.6	0.3	0.7	0.4
8-9%	0.6	0.4	0.6	0.2	0.6	0.6
9-10%	1.9	1.3	1.8	1.3	2.0	1.3
Total Below 10%	98.2	99.0	98.2	98.9	98.1	99.0
10-20%	1.3	1.0	1.3	1.0	1.4	1.0
20-30%	0.5	0.1	0.4	0.1	0.6	0.0
30-40%	0.0	0.0	0.0	0.0	0.0	0.0
40-50%						
Above 50 %						
Total	100	100	100	100	100	100

Table 37. The distribution of the participants and potential participants of SROP 1.1.2 programme by the Share of Roma in the population at the settlement of residence, Western Transdanubia

Proportion of Roma people in the settlement	Full Sample		Women		Men	
	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	19.2	17.5	18.5	17.3	19.8	17.7
0-1%	39.6	39.7	40.6	40.6	38.7	38.6
1-2%	23.1	29.3	23.4	29.7	22.9	28.7
2-3%	5.0	4.2	4.8	4.2	5.2	4.3
3-4%	2.1	1.4	2.1	1.3	2.2	1.5
4-5%	2.5	1.8	2.5	1.7	2.5	1.9
5-6%	1.2	1.1	1.2	1.0	1.3	1.2
6-7%	0.7	0.6	0.7	0.5	0.7	0.8
7-8%	1.1	0.9	1.1	0.7	1.2	1.1
8-9%	0.6	0.2	0.5	0.3	0.6	0.2
9-10%	1.3	0.7	1.2	0.7	1.3	0.8
Total Below 10%	96.5	97.4	96.6	97.9	96.3	96.8
10-20%	2.7	2.1	2.6	1.7	2.8	2.5
20-30%	0.6	0.5	0.5	0.4	0.6	0.7
30-40%	0.1	0.0	0.1	0.0	0.1	0.0
40-50%	0.2	0.0	0.2	0.1	0.2	0.0
Above 50 %						
Total	100	100	100	100	100	100

2.2. Alternative indicators measuring correlation between settlement-level participation rate of SROP 1.1.2 and share of Roma in the population

Table 38. Connection between settlement-level participation rate of SROP 1.1.2 and share of Roma in the population – settlements with more than 600 inhabitants

	Women	Women	Men	Men
Share of Roma is above average	0.00274*		0.00279	
	(0.00152)		(0.00212)	
Share of Roma is above median		0.00305***		0.00418***
		(0.00113)		(0.00153)
Northern Great Plain	0.0140***	0.0145***	0.0158***	0.0171***
	(0.00209)	(0.00254)	(0.00265)	(0.00311)
Southern Great Plain	0.00854***	0.00815***	0.00833***	0.00858***
	(0.00112)	(0.00126)	(0.00126)	(0.00144)
Southern Transdanubia	0.0107***	0.00957***	0.0114***	0.00783***
	(0.00155)	(0.00205)	(0.00177)	(0.00209)
Northern Hungary	0.0121***	0.0113***	0.0132***	0.0125***
	(0.00159)	(0.00189)	(0.00209)	(0.00263)
Central Transdanubia	0.00851***	0.00867***	0.00687***	0.00641***
	(0.00106)	(0.00120)	(0.00142)	(0.00142)
Western Transdanubia	0.0117***	0.0112***	0.0103***	0.0106***
	(0.00132)	(0.00144)	(0.00161)	(0.00173)
Capital or county capital	-0.00345**	-0.00350**	-0.00459**	-0.00505**
	(0.00165)	(0.00165)	(0.00206)	(0.00209)
City with above 10,000 inhabitants	0.00103	0.000990	-0.000270	-0.000608
	(0.00157)	(0.00156)	(0.00209)	(0.00210)
Village, 5 000-6 750 inhabitants	0.00158	0.00125	-0.000585	-0.000955
	(0.00212)	(0.00211)	(0.00246)	(0.00244)
Village, 2 000-3 000 inhabitants	-0.00199	-0.00205	-0.00181	-0.00177
	(0.00152)	(0.00152)	(0.00205)	(0.00206)
Village, 1 000-2 000 inhabitants	-0.00174	-0.00176	-0.000856	-0.000658
	(0.00153)	(0.00153)	(0.00206)	(0.00207)
Village, 601-1 000 inhabitants	0.000639	0.000630	0.00184	0.00190
	(0.00170)	(0.00169)	(0.00226)	(0.00225)
Share of Roma and NGP interaction	-0.00797***	-0.00713**	-0.00712**	-0.00806**
	(0.00275)	(0.00289)	(0.00361)	(0.00365)
Share of Roma and SGP interaction	0.00127	0.00135	0.00320	0.000494
	(0.00233)	(0.00189)	(0.00321)	(0.00230)
Share of Roma and ST interaction	-0.00175	4.29e-06	0.00191	0.00511*
	(0.00255)	(0.00249)	(0.00321)	(0.00279)
Share of Roma and NH interaction	-0.00371	-0.00204	-0.000425	4.59e-05
	(0.00230)	(0.00221)	(0.00309)	(0.00308)
Share of Roma and CT interaction	-0.00238	-0.00110	-0.00256	0.000811
	(0.00266)	(0.00195)	(0.00358)	(0.00303)
Share of Roma and WT interaction	-0.00310	0.000673	0.00330	0.00104
	(0.00312)	(0.00253)	(0.00463)	(0.00330)
Constant	0.0127***	0.0120***	0.0145***	0.0134***
	(0.00145)	(0.00154)	(0.00194)	(0.00203)
Observations	1,928	1,928	1,928	1,928
R-squared	0.063	0.065	0.071	0.078

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 39. Connection between settlement-level participation rate of SROP 1.1.2 and share of Roma in the population – settlements with less than 600 inhabitants

	Women	Women	Men	Men
Share of Roma is above average	-0.00102		-0.00491	
	(0.00744)		(0.00333)	
Share of Roma is above median		-0.0161		-0.00735
		(0.0179)		(0.00640)
Northern Great Plain	0.00218	-0.00239	0.0248***	0.0226**
	(0.0126)	(0.0175)	(0.00867)	(0.00976)
Southern Great Plain	-0.000515	-0.00431	0.0119**	0.00791
	(0.0121)	(0.0170)	(0.00544)	(0.00742)
Southern Transdanubia	0.00428	-0.000973	0.0271***	0.0233***
	(0.0120)	(0.0170)	(0.00556)	(0.00773)
Northern Hungary	0.0132	0.00910	0.0254***	0.0255***
	(0.0127)	(0.0176)	(0.00615)	(0.00802)
Central Transdanubia	0.00530	-0.000187	0.0189***	0.0171**
	(0.0121)	(0.0170)	(0.00604)	(0.00773)
Western Transdanubia	0.00711	0.00124	0.0231***	0.0209***
	(0.0120)	(0.0169)	(0.00592)	(0.00763)
Share of Roma and NGP interaction	-3.68e-05	0.0140	0.0165	0.0176
	(0.00929)	(0.0188)	(0.0133)	(0.0139)
Share of Roma and SGP interaction		0.0103		0.0116
		(0.0185)		(0.00819)
Share of Roma and ST interaction	0.000292	0.0154	0.0126**	0.0154*
	(0.00807)	(0.0182)	(0.00558)	(0.00817)
Share of Roma and NH interaction	-0.00152	0.0116	0.0160**	0.0120
	(0.0101)	(0.0192)	(0.00770)	(0.00935)
Share of Roma and CT interaction	-0.00570	0.0113	-0.00324	-0.00128
	(0.00901)	(0.0185)	(0.00694)	(0.00857)
Share of Roma and WT interaction	-0.000892	0.0165	-0.00151	0.00143
	(0.00864)	(0.0183)	(0.00590)	(0.00809)
Constant	0.0195*	0.0249	0.00490	0.00735
	(0.0118)	(0.0167)	(0.00450)	(0.00640)
Observations	1,249	1,249	1,249	1,249
R-squared	0.010	0.010	0.020	0.018

2.3. Comparison of some characteristics of actual and potential participants of programme 1.1.2 at different stages of the impact analysis

Table 40. The number of participants and potential participants in the samples of the different stages of evaluation

	SROP 1.1.2	
	Participant	Potential participant
All participants (P)	57 894	1 301 419
Impact assessment sample (IAS)	7 049*	371 789*
Matched sample (MS)	6 946*	6 946*

*Men only.

At the three levels of the samples (all participants; participants in the sample we used for the impact evaluation; matched sample) we compared the group of participants and controls based on some variables. Also we made a comparison of the same group at the different levels. In case of continuous variables we compared the averages by a two-sided t-test at the 5 percent significance level. We rejected the null hypotheses that the two means are the same if the p-value was less than 0.05. The means are in the columns "Treated" and "Control" starting from the second line (the number of participants is in the first line). In case of categorical variables we used Pearson's chi-squared test to determine the difference between the means of the samples. We rejected the null hypotheses that there is no difference in the means of the two groups if the p-value was smaller than 0.05.

Table 41. Comparison of the full set of participants (P) and the group of the participants that were included in the impact assessment sample (IAS)

	P participant	IAS participant	Test	P-value	Is there a difference?
Number of people	57 894	7 049			
Men	0.470	1.000	t-test	0.000	yes
Age	32.130	29.850	t-test	0.000	yes
Subregional unemployment level in 2008	0.110	0.120	t-test	0.000	yes
Share of Roma in the population category at the settlement	0.040	0.060	t-test	0.000	yes
Entered employment in 2008?	0.210	0.230	t-test	0.000	yes
Entered employment in 2009?	0.340	0.380	t-test	0.000	yes
Entered employment in 2010?	0.490	0.530	t-test	0.000	yes
Entered employment in 2011?	0.460	0.520	t-test	0.000	yes
Entered unemployment in 2008?	0.390	0.480	t-test	0.000	yes
Entered unemployment in 2009?	0.550	0.520	t-test	0.000	yes
Entered unemployment in 2010?	0.300	0.310	t-test	0.028	yes
Entered unemployment in 2011?	0.230	0.260	t-test	0.000	yes
Region			chi2-test	0.000	yes
Settlement size			chi2-test	0.000	yes
Education level on 9 level scale			chi2-test	0.000	yes

Table 42. Comparison of the full set of participants (P) and the group of the participants that were included in the matched sample (MS)

	P participant	MS participant	Test	P-value	Is there a difference?
Number of people	57 894	6 946			
Men	0.470	1.000	t-test	0.000	yes
Age	32.130	29.860	t-test	0.000	yes
Subregional unemployment level in 2008	0.110	0.120	t-test	0.000	yes
Share of Roma in the population category at the settlement	0.040	0.060	t-test	0.000	yes
Entered employment in 2008?	0.210	0.240	t-test	0.000	yes
Entered employment in 2009?	0.340	0.390	t-test	0.000	yes
Entered employment in 2010?	0.490	0.530	t-test	0.000	yes
Entered employment in 2011?	0.460	0.520	t-test	0.000	yes
Entered unemployment in 2008?	0.390	0.480	t-test	0.000	yes
Entered unemployment in 2009?	0.550	0.520	t-test	0.000	yes
Entered unemployment in 2010?	0.300	0.310	t-test	0.088	no
Entered unemployment in 2011?	0.230	0.260	t-test	0.000	yes
Region			chi2-test	0.000	yes
Settlement size			chi2-test	0.000	yes
Education level on 9 level scale			chi2-test	0.000	yes

Table 43. Comparison of the participants from the impact assessment sample (IAS) and the group of the participants that were included in the matched sample (MS)

	IAS participant	MS participant	Test	P-value	Is there a difference?
Number of people	7049	6946			
Men	1	1	t-test		no
Age	29.85	29.86	t-test	0.957	no
Subregional unemployment level in 2008	0.12	0.12	t-test	0.796	no
Share of Roma in the population category at the settlement	0.06	0.06	t-test	0.896	no
Entered employment in 2008?	0.23	0.24	t-test	0.824	no
Entered employment in 2009?	0.38	0.39	t-test	0.8	no
Entered employment in 2010?	0.53	0.53	t-test	0.995	no
Entered employment in 2011?	0.52	0.52	t-test	0.877	no
Entered unemployment in 2008?	0.48	0.48	t-test	0.983	no
Entered unemployment in 2009?	0.52	0.52	t-test	0.876	no
Entered unemployment in 2010?	0.31	0.31	t-test	0.717	no
Entered unemployment in 2011?	0.26	0.26	t-test	0.766	no
Region			chi2-test	1	no
Settlement size			chi2-test	1	no
Education level on 9 level scale			chi2-test	0.996	no

Table 44. Comparison of the participants and the potential participants of programme 1.1.2 (P sample)

	Treat	Control	Test	P-value	Is there a difference?
Number of people	57 894	1 301 419			
Men	0.47	0.53	t-test	0	yes
Age	32.13	38.03	t-test	0	yes
Subregional unemployment level in 2008	0.11	0.1	t-test	0	yes
Share of Roma in the population category at the settlement	0.04	0.05	t-test	0	yes
Entered employment in 2008?	0.21	0.07	t-test	0	yes
Entered employment in 2009?	0.34	0.04	t-test	0	yes
Entered employment in 2010?	0.49	0.04	t-test	0	yes
Entered employment in 2011?	0.46	0.02	t-test	0	yes
Entered unemployment in 2008?	0.39	0.23	t-test	0	yes
Entered unemployment in 2009?	0.55	0.26	t-test	0	yes
Entered unemployment in 2010?	0.3	0.22	t-test	0	yes
Entered unemployment in 2011?	0.23	0.21	t-test	0	yes
Region			chi2-test	0	yes
Settlement size			chi2-test	0	yes
Education level on 9 level scale			chi2-test	0	yes

Table 45. Comparison of the participants and the potential participants of programme 1.1.2 who were included in the impact assessment analysis (IAS sample)

	Treat	Control	Test	P-value	Is there a difference?
Number of people	7 049	371 789			
Men	1	1	t-test		no
Age	29.85	36.09	t-test	0	yes
Sub-regional unemployment level in 2008	0.12	0.11	t-test	0	yes
Share of Roma in the population category at the settlement	0.06	0.06	t-test	0.027	yes
Entered employment in 2008?	0.23	0.06	t-test	0	yes
Entered employment in 2009?	0.38	0.03	t-test	0	yes
Entered employment in 2010?	0.53	0.02	t-test	0	yes
Entered employment in 2011?	0.52	0.01	t-test	0	yes
Entered unemployment in 2008?	0.48	0.26	t-test	0	yes
Entered unemployment in 2009?	0.52	0.28	t-test	0	yes
Entered unemployment in 2010?	0.31	0.22	t-test	0	yes
Entered unemployment in 2011?	0.26	0.21	t-test	0	yes
Region			chi2-test	0	yes
Settlement size			chi2-test	0.214	no
Education level on 9 level scale			chi2-test	0	yes

Table 46. Comparison of the sample participants and the potential participants of programme 1.1.2 after the matching (MS sample)

	Treat	Control	Test	P-value	Is there a difference?
Number of people	6 946	6 946			
Men	1	1	t-test		no
Age	29.86	29.95	t-test	0.61	no
Sub-regional unemployment level in 2008	0.12	0.11	t-test	0.001	yes
Share of Roma in the population category at the settlement	0.06	0.06	t-test	0.29	no
Entered employment in 2008?	0.24	0.09	t-test	0	yes
Entered employment in 2009?	0.39	0.03	t-test	0	yes
Entered employment in 2010?	0.53	0	t-test	0	yes
Entered employment in 2011?	0.52	0	t-test	0	yes
Entered unemployment in 2008?	0.48	0.45	t-test	0	yes
Entered unemployment in 2009?	0.52	0.5	t-test	0.018	yes
Entered unemployment in 2010?	0.31	0.32	t-test	0.13	no
Entered unemployment in 2011?	0.26	0.24	t-test	0.005	yes
Region			chi2-test	1	no
Settlement size			chi2-test	1	no
Education level on 9 level scale			chi2-test	0.607	no

2.4. The distribution of the participants versus potential participants of the One step ahead! I-II. programme as a function of the share of the Roma population, by regions

Table 47: The distribution of the participants versus potential participants of the One step ahead! I-II. programme as a function of the share of the Roma population, Northern Great Plain

Share of the Roma in the settlement	All participants		Women		Men	
	Potential participants (%)	Par-ticipants (%)	Potential participants (%)	Par-ticipants (%)	Potential participants (%)	Par-ticipants (%)
0%	3.1	3.9	3.1	3.3	3.1	4.6
0-1%	17.4	12.2	17.3	13.8	17.5	10.2
1-2%	9.6	7.2	9.5	6.6	9.6	7.9
2-3%	15.1	11.1	15.3	11.5	15.0	10.4
3-4%	7.7	7.3	7.7	7.6	7.7	7.0
4-5%	6.7	11.0	6.6	11.9	6.8	9.8
5-6%	4.3	4.5	4.3	4.5	4.3	4.5
6-7%	2.1	1.3	2.1	0.8	2.1	2.0
7-8%	7.2	7.2	7.2	7.5	7.1	6.9
8-9%	2.2	2.3	2.2	2.4	2.2	2.3
9-10%	1.5	2.8	1.5	2.1	1.5	3.8
Below 10%	76.8	70.8	76.8	71.9	76.8	69.3
10-20%	16.0	19.6	15.9	19.5	16.1	19.7
20-30%	5.3	6.3	5.3	6.0	5.2	6.8
30-40%	1.2	2.1	1.3	1.9	1.2	2.4
40-50%	0.1	0.3	0.1	0.4	0.1	0.1
50-100%	0.7	1.0	0.7	0.4	0.7	1.7
Total (%)	100	100	100	100	100	100

Remark: the equality sign is on the upper limits of the intervals.

Table 48: The distribution of the participants versus potential participants of the One step ahead! I-II. programme as a function of the share of the Roma population, Southern Great Plain

Share of the Roma in the settlement	All participants		Women		Men	
	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	7.2	5.2	7.2	4.9	7.2	5.5
0-1%	31.9	25.3	31.4	25.8	32.3	24.4
1-2%	26.3	18.8	26.4	18.1	26.2	20.1
2-3%	9.9	12.4	10.1	13.4	9.7	10.5
3-4%	6.3	7.1	6.3	5.8	6.2	9.5
4-5%	3.7	8.5	3.8	8.5	3.7	8.4
5-6%	7.2	11.5	7.5	10.7	7.0	12.9
6-7%	3.9	4.3	3.8	5.0	3.9	3.1
7-8%	0.5	1.5	0.4	1.4	0.5	1.6
8-9%	1.1	1.5	1.1	1.5	1.1	1.4
9-10%	1.3	3.0	1.2	3.7	1.4	1.7
Below 10%	99.2	98.9	99.2	98.8	99.2	99.1
10-20%	0.7	1.0	0.7	1.1	0.7	0.8
30-40%	0.1	0.1	0.1	0.1	0.1	0.1
Total	100	100	100	100	100	100

Remark: the equality sign is on the upper limits of the intervals.

Table 49: The distribution of the participants versus potential participants of the One step ahead! I-II. programme as a function of the share of the Roma population, Central Hungary

Share of the Roma in the settlement	All participants		Women		Men	
	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	2.1	1.9	2.0	0.8	2.2	2.9
0-1%	31.2	24.2	30.9	21.9	31.3	26.5
1-2%	40.4	32.2	40.2	32.2	40.6	32.2
2-3%	12.3	15.4	12.7	16.1	12.0	14.6
3-4%	7.2	8.2	7.3	8.1	7.1	8.4
4-5%	2.8	4.4	2.9	5.4	2.8	3.4
5-6%	0.9	1.2	0.9	1.4	0.9	1.1
6-7%	0.4	0.7	0.4	0.3	0.4	1.1
7-8%	0.5	0.2	0.4	0.1	0.5	0.3
8-9%	0.1	0.1	0.1	0.1	0.1	0.1
9-10%	0.3	0.4	0.3	0.6	0.3	0.3
Below 10%	98.1	88.9	98.1	87.0	98.1	90.8
10-20%	1.5	6.8	1.5	7.4	1.5	6.2
20-30%	0.4	4.3	0.4	5.6	0.4	3.0
Total	100	100	100	100	100	100

Remark: the equality sign is on the upper limits of the intervals.

Table 50: The distribution of the participants versus potential participants of the One step ahead! I-II. programme as a function of the share of the Roma population, Southern Transdanubia

Share of the Roma in the settlement	All participants		Women		Men	
	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	5.5	3.2	5.4	3.3	5.5	3.0
0-1%	5.5	4.3	5.5	4.1	5.5	4.6
1-2%	13.0	13.0	13.1	13.2	13.0	12.8
2-3%	27.5	19.4	27.7	20.8	27.3	17.2
3-4%	9.1	7.4	8.9	7.8	9.2	6.8
4-5%	6.9	7.3	7.1	7.1	6.7	7.5
5-6%	3.8	4.2	3.7	3.9	3.8	4.7
6-7%	3.5	3.8	3.4	2.6	3.5	5.6
7-8%	3.7	3.0	3.7	2.8	3.8	3.2
8-9%	5.3	9.1	5.1	9.4	5.4	8.7
9-10%	0.7	0.6	0.7	0.6	0.7	0.7
Below 10%	84.4	75.2	84.3	75.6	84.4	74.7
10-20%	9.0	12.5	9.1	11.3	8.9	14.3
20-30%	4.1	4.7	4.0	5.3	4.1	3.9
30-40%	1.1	1.9	1.1	1.4	1.1	2.6
40-50%	0.4	1.5	0.4	1.7	0.4	1.2
50-100%	1.2	4.2	1.2	4.8	1.2	3.4
Total	100	100	100	100	100	100

Remark: the equality sign is on the upper limits of the intervals.

Table 51: The distribution of the participants versus potential participants of the One step ahead! I-II. programme as a function of the share of the Roma population, Northern Hungary

Share of the Roma in the settlement	All participants		Women		Men	
	Potential participants (%)	Par- ticipants (%)	Potential participants (%)	Par- ticipants (%)	Potential participants (%)	Par- ticipants (%)
0%	5.5	3.1	5.4	3.2	5.5	3.0
0-1%	5.9	3.6	5.8	4.1	5.9	2.7
1-2%	7.0	7.4	6.8	7.2	7.1	7.8
2-3%	7.6	7.7	7.5	9.1	7.6	5.3
3-4%	14.5	14.7	14.1	15.8	14.8	12.9
4-5%	2.6	2.8	2.5	2.5	2.7	3.5
5-6%	5.8	8.7	5.9	6.9	5.6	11.6
6-7%	3.8	2.4	3.9	2.2	3.8	2.6
7-8%	5.7	6.1	5.8	5.7	5.7	6.7
8-9%	5.2	4.6	5.3	4.9	5.1	4.1
9-10%	1.6	1.2	1.6	1.5	1.6	0.6
Below 10%	65.0	62.2	64.5	63.0	65.4	60.9
10-20%	21.8	26.7	22.0	26.3	21.8	27.4
20-30%	6.2	4.3	6.4	3.5	6.1	5.6
30-40%	4.1	4.4	4.3	5.1	4.0	3.2
40-50%	1.9	1.6	1.9	1.2	1.9	2.3
50-100%	0.9	0.9	1.0	1.0	0.9	0.7
Total	100	100	100	100	100	100

Remark: the equality sign is on the upper limits of the intervals.

Table 52: The distribution of the participants versus potential participants of the One step ahead! I-II. programme as a function of the share of the Roma population, Central Transdanubia

Share of the Roma in the settlement	All participants		Women		Men	
	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)	Potential participants (%)	Participants (%)
0%	11.2	7.1	11.1	7.6	11.2	6.7
0-1%	42.5	37.5	41.9	39.3	43.0	35.8
1-2%	24.7	25.6	25.4	24.7	24.1	26.3
2-3%	11.6	14.3	11.7	14.2	11.5	14.5
3-4%	1.9	1.4	1.8	0.9	1.9	1.9
4-5%	1.0	0.5	0.9	0.5	1.0	0.6
5-6%	0.7	0.6	0.7	0.5	0.7	0.7
6-7%	1.9	3.1	1.9	2.7	1.9	3.5
7-8%	0.6	0.6	0.6	1.1	0.6	0.1
8-9%	0.6	0.4	0.6	0.5	0.5	0.3
9-10%	1.8	2.3	1.8	2.8	1.8	1.7
Below 10%	98.3	93.3	98.3	94.7	98.3	92.0
10-20%	1.3	6.6	1.3	5.4	1.3	7.7
20-30%	0.4	0.2	0.4	0.0	0.4	0.3
40-50%	0.0	0.0	0.0	0.0	0.0	0.0
Total	100	100	100	100	100	100

Remark: the equality sign is on the upper limits of the intervals.

Table 53: The distribution of the participants versus potential participants of the One step ahead! I-II. programme as a function of the share of the Roma population, Western Transdanubia

Share of the Roma in the settlement	All participants		Women		Men	
	Potential participants (%)	Par-ticipants (%)	Potential participants (%)	Par-ticipants (%)	Potential participants (%)	Par-ticipants (%)
0%	20.8	12.6	20.3	12.7	21.1	12.5
0-1%	39.6	39.8	40.2	33.5	39.1	46.1
1-2%	22.1	22.2	22.2	26.7	22.0	17.8
2-3%	5.1	4.6	4.9	4.1	5.3	5.1
3-4%	2.2	1.2	2.2	1.5	2.2	1.0
4-5%	2.5	3.2	2.5	3.9	2.5	2.5
5-6%	1.2	1.2	1.2	1.6	1.2	0.7
6-7%	0.7	0.9	0.7	0.6	0.7	1.1
7-8%	1.1	1.1	1.1	1.1	1.1	1.1
8-9%	0.5	0.9	0.5	0.9	0.6	0.9
9-10%	1.1	2.7	1.1	2.7	1.2	2.7
Below 10%	96.8	90.3	96.9	89.2	96.8	91.4
10-20%	2.5	7.5	2.5	7.5	2.5	7.5
20-30%	0.5	0.9	0.5	1.2	0.6	0.6
30-40%	0.0	0.9	0.1	1.5	0.0	0.4
40-50%	0.1	0.4	0.1	0.6	0.1	0.1
Total	100	100	100	100	100	100

Remark: the equality sign is on the upper limits of the intervals.

2.5. Alternative indicators for measuring the relationship between the settlement-level participation rate of programme 2.1.1 and the settlement-level share of Roma in the population

Table 54. Measuring the connection between the settlement-level participation rate of SROP 2.1.1 and the town level Share of Roma in the population – settlements with more than 600 inhabitants

VARIABLES	Women	Women	Men	Men
Share of Roma is above average	0.00797**		0.00704**	
	(0.00336)		(0.00288)	
Share of Roma is above median		0.00610***		0.00484***
		(0.00159)		(0.00152)
Northern Great Plain	0.00935***	0.00894***	0.00575***	0.00596**
	(0.00175)	(0.00200)	(0.00189)	(0.00279)
Southern Great Plain	0.00578***	0.00482***	0.00165*	0.000893
	(0.00101)	(0.00105)	(0.000915)	(0.00108)
Southern Transdanubia	0.00258**	0.00153	-0.00109	-0.00109
	(0.00104)	(0.00129)	(0.000968)	(0.00123)
Northern Hungary	0.00410***	0.00307***	-0.000275	-0.00163
	(0.00106)	(0.00117)	(0.00110)	(0.00129)
Central Transdanubia	0.000179	0.000678	0.000271	0.000134
	(0.000812)	(0.000951)	(0.000858)	(0.000950)
Western Transdanubia	-1.79e-05	-0.000688	-8.24e-05	-0.000422
	(0.000796)	(0.000742)	(0.000922)	(0.00103)
Capital or county capital	-0.00102	-0.00159	-0.000207	-0.000906
	(0.00144)	(0.00169)	(0.00156)	(0.00178)
City with above 10,000 inhabitants	-0.00136	-0.00162	-0.000733	-0.00110
	(0.00152)	(0.00165)	(0.00162)	(0.00173)
Village, 5 000-6 750 inhabitants	-0.000239	0.000222	-0.00270	-0.00216
	(0.00197)	(0.00204)	(0.00175)	(0.00176)
Village, 2 000-3 000 inhabitants	0.000844	0.00184	0.00126	0.00231
	(0.00160)	(0.00172)	(0.00165)	(0.00175)
Village, 1 000-2 000 inhabitants	0.000362	0.00167	0.000958	0.00231
	(0.00153)	(0.00165)	(0.00164)	(0.00172)
Village, 601-1 000 inhabitants	0.00218	0.00334*	0.00292	0.00407*
	(0.00179)	(0.00194)	(0.00202)	(0.00218)
Share of Roma and NGP interaction	0.00170	0.00325	0.00649	0.00649
	(0.00444)	(0.00327)	(0.00448)	(0.00402)
Share of Roma and SGP interaction	-6.79e-05	0.00133	-0.000539	0.000658
	(0.00418)	(0.00244)	(0.00359)	(0.00211)
Share of Roma and ST interaction	-0.00201	0.000179	-0.00132	-0.000578
	(0.00376)	(0.00225)	(0.00321)	(0.00201)
Share of Roma and NH interaction	-0.00275	0.000350	-0.00143	0.00189
	(0.00372)	(0.00221)	(0.00323)	(0.00210)
Share of Roma and CT interaction	-0.00742**	-0.00570***	-0.00196	-0.00172
	(0.00373)	(0.00200)	(0.00539)	(0.00261)
Share of Roma and WT interaction	-0.00186	0.000651	-0.00217	-0.000939
	(0.00410)	(0.00231)	(0.00400)	(0.00220)
Constant	0.00318**	0.00161	0.00409***	0.00286*
	(0.00135)	(0.00150)	(0.00146)	(0.00163)
Observations	1,927	1,927	1,928	1,928

R-squared	0.113	0.112	0.096	0.086

Table 55. Measuring the connection between the settlement-level participation rate of *One step ahead!* programme and the town level Share of Roma in the population – settlements with fewer than 600 inhabitants

VARIABLES	Women	Women	Men	Men
Share of Roma is above average	-0.0105		-0.0179	
	(0.0107)		(0.0202)	
Share of Roma is above median		0.00439		-0.0227
		(0.00312)		(0.0198)
Northern Great Plain	0.0601***	0.0714***	0.0386	0.0376
	(0.0211)	(0.0247)	(0.0266)	(0.0334)
Southern Great Plain	0.0173*	0.0198*	0.0146	0.0114
	(0.00902)	(0.0114)	(0.0245)	(0.0328)
Southern Transdanubia	0.00933**	0.00719***	-0.00750	-0.0159
	(0.00407)	(0.00206)	(0.0140)	(0.0199)
Northern Hungary	0.0174**	0.0194**	0.00313	-0.00167
	(0.00827)	(0.00952)	(0.0158)	(0.0216)
Central Transdanubia	0.000635	0.00226***	-0.0118	-0.0193
	(0.00151)	(0.000745)	(0.0140)	(0.0198)
Western Transdanubia	0.00684***	0.00821***	-0.00982	-0.0172
	(0.00195)	(0.00145)	(0.0139)	(0.0198)
Share of Roma and NGP interaction	-0.0112	-0.0383	0.0100	0.00618
	(0.0263)	(0.0270)	(0.0326)	(0.0351)
Share of Roma and SGP interaction		-0.0107		0.00329
		(0.0135)		(0.0336)
Share of Roma and ST interaction	0.0199*	0.00819*	0.0300	0.0337*
	(0.0116)	(0.00476)	(0.0205)	(0.0200)
Share of Roma and NH interaction	0.00913	-0.00653	0.0138	0.0141
	(0.0143)	(0.0110)	(0.0218)	(0.0218)
Share of Roma and CT interaction	0.0139	-0.00242	0.0197	0.0239
	(0.0110)	(0.00383)	(0.0204)	(0.0199)
Share of Roma and WT interaction	0.0319**	0.0121*	0.0273	0.0291
	(0.0130)	(0.00672)	(0.0205)	(0.0200)
Constant	0.00146	0	0.0152	0.0227
	(0.00134)		(0.0139)	(0.0198)
Observations	1,249	1,249	1,249	1,249
R-squared	0.043	0.044	0.054	0.054

2.6. Comparison of some characteristics of actual and potential participants of programme 2.1.1 at different stages of the impact analysis

Table 56: The number of participants and potential participants at the samples of the different stages

	SROP 1.1.2	
	Participant	Potential participant
All participants (P)	16 320	1 499 741
Impact assessment sample (IAS)	7 226	446 940
Matched sample (MS)	6 720	6 720

Table 57. Comparison of the full set of programme 2.1.1 participants (P) and the group of the participants that were included in the impact assessment sample (IAS)

	P participant	IAS participant	Test	P-value	Is there a difference?
Number of people	16 320	7 226			
Men	0.42	0.43	t-test	0.12	no
Age	33.29	32.2	t-test	0	yes
Subregional unemployment level in 2008	0.13	0.13	t-test	0	yes
Share of Roma in the population category at the settlement	0.07	0.09	t-test	0	yes
Entered employment in 2008?	0.26	0.24	t-test	0	yes
Entered employment in 2009?	0.37	0.42	t-test	0	yes
Entered employment in 2010?	0.4	0.45	t-test	0	yes
Entered employment in 2011?	0.4	0.41	t-test	0.022	yes
Entered unemployment in 2008?	0.31	0.32	t-test	0.183	no
Entered unemployment in 2009?	0.29	0.29	t-test	0.485	no
Entered unemployment in 2010?	0.22	0.21	t-test	0.008	yes
Entered unemployment in 2011?	0.21	0.19	t-test	0.002	yes
Region			chi2-test	0	yes
Settlement size			chi2-test	0	yes
Education level on 9 level scale			chi2-test	0	yes

Table 58. Comparison of the full set of programme 2.1.1 participants (P) and the group of the participants that were included in the matched sample (MS)

	P participant	MS participant	Test	P-value	Is there a difference?
Number of people	16 320	6720			
Men	0.42	0.43	t-test	0.133	no
Age	33.29	32.28	t-test	0	yes
Subregional unemployment level in 2008	0.13	0.13	t-test	0	yes
Share of Roma in the population category at the settlement	0.07	0.09	t-test	0	yes
Entered employment in 2008?	0.26	0.24	t-test	0	yes
Entered employment in 2009?	0.37	0.43	t-test	0	yes
Entered employment in 2010?	0.4	0.46	t-test	0	yes
Entered employment in 2011?	0.4	0.42	t-test	0.002	yes
Entered unemployment in 2008?	0.31	0.33	t-test	0.017	yes
Entered unemployment in 2009?	0.29	0.29	t-test	0.895	no
Entered unemployment in 2010?	0.22	0.2	t-test	0	yes
Entered unemployment in 2011?	0.21	0.19	t-test	0	yes
Region			chi2-test	0	yes
Settlement size			chi2-test	0	yes
Education level on 9 level scale			chi2-test	0	yes

Table 59. Comparison of programme 2.1.1 participants from the impact assessment sample (IAS) and the group of the participants that were included in the matched sample (MS)

	IAS participant	MS participant	Test	P-value	Is there a difference?
Number of people	7226	6720			
Men	0.43	0.43	t-test	0.991	no
Age	32.2	32.28	t-test	0.63	no
Subregional unemployment level in 2008	0.13	0.13	t-test	0.067	no
Share of Roma in the population category at the settlement	0.09	0.09	t-test	0.415	no
Entered employment in 2008?	0.24	0.24	t-test	0.867	no
Entered employment in 2009?	0.42	0.43	t-test	0.103	no
Entered employment in 2010?	0.45	0.46	t-test	0.352	no
Entered employment in 2011?	0.41	0.42	t-test	0.43	no
Entered unemployment in 2008?	0.32	0.33	t-test	0.355	no
Entered unemployment in 2009?	0.29	0.29	t-test	0.487	no
Entered unemployment in 2010?	0.21	0.2	t-test	0.325	no
Entered unemployment in 2011?	0.19	0.19	t-test	0.305	no
Region			chi2-test	0.691	no
Settlement size			chi2-test	0.995	no
Education level on 9 level scale			chi2-test	0.523	no

Table 60. Comparison of the participants and the potential participants of programme 2.1.1. (P sample)

	Treat	Control	Test	P-value	Is there a difference?
Number of people	16 320	1356743			
Men	0.42	0.58	t-test	0	yes
Age	33.29	36.13	t-test	0	yes
Subregional unemployment level in 2008	0.13	0.1	t-test	0	yes
Share of Roma in the population category at the settlement	0.07	0.05	t-test	0	yes
Entered employment in 2008?	0.26	0.06	t-test	0	yes
Entered employment in 2009?	0.37	0.04	t-test	0	yes
Entered employment in 2010?	0.4	0.03	t-test	0	yes
Entered employment in 2011?	0.4	0.02	t-test	0	yes
Entered unemployment in 2008?	0.31	0.24	t-test	0	yes
Entered unemployment in 2009?	0.29	0.29	t-test	0.883	no
Entered unemployment in 2010?	0.22	0.22	t-test	0.231	no
Entered unemployment in 2011?	0.21	0.2	t-test	0.078	no
Region			chi2-test	0	yes
Settlement size			chi2-test	0	yes
Education level on 9 level scale			chi2-test	0	yes

Table 61. Comparison of the participants and the potential participants of programme 2.1.1 who were included in the impact assessment analysis (IAS sample)

	Treat	Control	Test	P-value	Is there a difference?
Number of people	7226	446940			
Men	0.43	0.49	t-test	0	yes
Age	32.2	39.9	t-test	0	yes
Subregional unemployment level in 2008	0.13	0.11	t-test	0	yes
Share of Roma in the population category at the settlement	0.09	0.06	t-test	0	yes
Entered employment in 2008?	0.24	0.03	t-test	0	yes
Entered employment in 2009?	0.42	0	t-test	0	yes
Entered employment in 2010?	0.45	0	t-test	0	yes
Entered employment in 2011?	0.41	0	t-test	0	yes
Entered unemployment in 2008?	0.32	0.24	t-test	0	yes
Entered unemployment in 2009?	0.29	0.24	t-test	0	yes
Entered unemployment in 2010?	0.21	0	t-test	0	yes
Entered unemployment in 2011?	0.19	0	t-test	0	yes
Region			chi2-test	0	yes
Settlement size			chi2-test	0	yes
Education level on 9 level scale			chi2-test	0	yes

Table 62. Comparison of the sample participants and the potential participants of programme 2.1.1 after the matching (MS sample)

	Treat	Control	Test	P-value	Is there a difference?
Number of people	6720	6720			
Men	0.43	0.43	t-test	1	no
Age	32.28	33.49	t-test	0	yes
Subregional unemployment level in 2008	0.13	0.13	t-test	0.393	no
Share of Roma in the population category at the settlement	0.09	0.09	t-test	0.979	no
Entered employment in 2008?	0.24	0	t-test	0	yes
Entered employment in 2009?	0.43	0	t-test	0	yes
Entered employment in 2010?	0.46	0	t-test	0	yes
Entered employment in 2011?	0.42	0	t-test	0	yes
Entered unemployment in 2008?	0.33	0.34	t-test	0.096	no
Entered unemployment in 2009?	0.29	0.19	t-test	0	yes
Entered unemployment in 2010?	0.2	0	t-test	0	yes
Entered unemployment in 2011?	0.19	0	t-test	0	yes
Region			chi2-test	1	no
Settlement size			chi2-test	1	no
Education level on 9 level scale			chi2-test	0	yes