



Funded by  
the European Union



MINISTERSTVO PRÁCE,  
SOCIÁLNYCH VECÍ A RODINY  
SLOVENSKEJ REPUBLIKY

# The net effects of graduate work experience and the promotion of self-employment

## Technical report



2015

*This Pilot Counterfactual Impact Evaluation of Self-employment and Graduate Practice VS/2014/0072 has been produced with support from the European Social Fund.*

*Authors:*

Borik Vladimir, PhD.

Durica Marek, PhD.

Molnarova Miboslava, PhD.

Svabova Lucia, PhD.

In the evaluation report are presented conclusions, recommendations and opinions which do not necessarily represent the views of the Ministry of Labour and Social Affairs of the Slovak Republic; further, the evaluation report is a collection of the interviewed opinions of the participants of the programs.

## Table of Contents

Executive summary.....	4
Introduction.....	7
1 Slovak labour market at a glance during the period under focus.....	9
2 Data source .....	12
2.1 Data preparation.....	12
2.1.1 Input and output data (treatment variables).....	13
2.1.2 Outcome data (treatment characteristics) .....	20
2.1.3 Context data (control variables).....	21
2.2 Selection bias.....	22
3 Methodology .....	26
3.1 Hierarchical cluster analysis .....	26
3.2 Parametric and non-parametric tests .....	26
3.3 Correlation .....	26
3.4 Post-only non-equivalent comparison design .....	27
3.5 Propensity score matching .....	27
3.6 Cost-benefit analysis .....	28
3.7 IT tools applied.....	29
4 Traineeship.....	30
4.1 Treatment effects of traineeship .....	30
4.2 Reference periods.....	32
4.3 Target groups.....	32
4.4 Test of representativeness of samples .....	33
4.4.1 Treated group excluded from the sample.....	33
4.4.2 Distributions of frequencies of treated individuals included and excluded from the sample .....	34
4.4.3 Non-treated group excluded from the sample.....	36
4.4.4 Distributions of frequencies of non-treated individuals included and excluded from the sample .....	36
4.5 Description of samples .....	38
4.5.1 Permanent residence .....	38
4.5.2 Gender .....	40
4.5.3 Marital status.....	41
4.5.4 Types of disadvantages.....	42
4.5.5 Age.....	42
4.5.6 Level of education .....	46
4.5.7 Registered before 2007 .....	48
4.5.8 Category of driving licence.....	49
4.6 Analysis of variance.....	49

4.6.1	1 <sup>st</sup> reference period.....	50
4.6.2	2 <sup>nd</sup> reference period.....	53
4.6.3	3 <sup>rd</sup> reference period.....	55
4.6.4	4 <sup>th</sup> reference period.....	58
4.7	Qualitative survey of graduate work experience.....	62
4.7.1	Conclusions from the interviews.....	63
4.8	Net effects of traineeship.....	65
4.8.1	Analysis of influences on employability.....	65
4.8.2	“Post-only non-equivalent comparison design” Method.....	68
4.8.3	Exact matching with the application of Post-only non-equivalent comparison design.....	72
4.8.4	Propensity score exact matching.....	77
4.8.5	Propensity score nearest neighbour matching.....	83
4.8.6	Comparisons of the method results.....	86
4.8.7	Identification of the successful target group for traineeship.....	89
4.8.8	Impact of the traineeship.....	93
5	Self-employment.....	96
5.1	Treatment effects of self-employment.....	96
5.2	Reference periods.....	98
5.3	Target group of self-employment promotion.....	99
5.4	Test of representativeness of the samples.....	100
5.4.1	Treated group excluded from the sample.....	100
5.4.2	Distributions of frequencies of treated individuals included and excluded from the sample.....	101
5.4.3	Non-treated group excluded from the sample.....	102
5.4.4	Distributions of frequencies of non-treated individuals included and excluded from the sample.....	102
5.5	Description of samples.....	104
5.5.1	Permanent residence.....	104
5.5.2	Gender.....	105
5.5.3	Marital status.....	105
5.5.4	Types of disadvantage.....	106
5.5.5	Age.....	106
5.5.6	Level of education.....	108
5.5.7	Registered before 2007.....	109
5.5.8	Category of driving licence.....	109
5.6	Analysis of variance.....	110
5.6.1	1st reference period.....	111
5.6.2	2nd reference period.....	113
5.7	Qualitative survey of self-employment.....	115
5.7.1	Conclusions from the interviews.....	116
5.8	Net effects of self-employment.....	120
5.8.1	Analysis of influences on self-employability.....	120
5.8.2	“Post-only non-equivalent comparison design” method.....	122
5.8.3	Exact matching with the application of Post-only non-equivalent	

comparison design.....	125
5.8.4 Propensity score exact matching .....	130
5.8.5 Propensity score nearest neighbour matching .....	134
5.8.6 Comparison of the method results .....	136
5.8.7 Identification of the successful target group for self-employment .....	139
5.8.8 Impact of the self-employment promotion .....	142
6 Strengths and weaknesses of the evaluation .....	144
6.1 Strengths .....	144
6.2 Weaknesses .....	145
6.3 Lessons learned.....	146
7 Conclusions and recommendations .....	148
7.1 Conclusions.....	148
7.1.1 Traineeship .....	148
7.1.2 Self-employment.....	150
7.2 Recommendations.....	153
7.2.1 Traineeship .....	153
7.2.2 Self-employment.....	154
7.2.3 Monitoring of the relevant data.....	155
References .....	157
List of abbreviations.....	158

## Executive summary

This summary relates to an evaluation aimed at the net effects of graduate work experience and the promotion of self-employment as two intervened measures supported by the Operational Programme Employment and Social Inclusion 2007 – 2013 (co-financed by the ESF). The evaluation has been carried out under the Pilot Counterfactual Impact Evaluation of Self-employment and Graduate Practice that was granted by the European Commission within the grant agreement No. VS/2014/0072. The grant was of a maximum amount of 124 417.90 €.

The activities were realised by internal evaluation team of the Ministry of Labour, Social Affairs and Family of the Slovak Republic and external experts in the field of statistics and counterfactual methods evaluations. The principal role of the Pilot counterfactual impact evaluation (hereinafter only “CIE”) was to provide four quasi-experimental approaches to counter-factual impact evaluation methods, of which the main message was to estimate the net effect of the graduate work experience and self-employment. The performed methods showed results very similar to the net effects based on the employability of the jobseekers due to the intervention. A supportive objective of this evaluation was the estimation of the net and gross financial effect of the interventions on the national budget in the impact period, estimated in respect to paid/saved unemployment allowance, taxes, increase of consumption, etc.

The analysis, statistical interpretation and evaluation of interviewed respondents for the two above-mentioned measures resulted in the main findings. For the measure of graduate work experience, it was stated that graduates were mostly placed in full-time jobs; they were very rarely interested in self-employment, which was considered as another type of labour market placement. Part-time jobs registered in the Slovak Insurance Agency were considered as jobseekers that were not fully placed on the labour market. According to the results, in most cases and methods, the participants of the graduate work experience were more strenuous and, on average, they were able to find part-time jobs for a longer period compared to their peers. In the last three reference periods, the independence tests confirmed a significant positive treatment effect on participants’ placement in part-time jobs due to the intervention. Jobseekers who had attended graduate work experience were earning, on average, from 430 up to 500 euros per month, depending on the particular year, during the 2-year-long period after the intervention was ended. The evaluation generally uncovered significant negative differences between participants and non-participants of the programme. Just to simplify, the unemployed and registered graduates who had attended the graduate work experience were earning on average from 30 to 80 euros per month.

As for the measure aimed at self-employment, the most desired effect of this active labour policy measure is sustainable self-employment of the participants on the open market or their placement on the open labour market as full-time job employees, i.e. being out of the jobseekers evidence. This effect is represented in the variable “Placement on the labour market.” According to the counterfactual impact evaluation methods carried out, it could be estimated that, on average, participants managed to stay out of the jobseekers evidence approximately less than 20% of the impact period (2 years after sustainability of self-employment). In other words, one programme participant would have been employed approximately 50 days less if the financial intervention had not been granted. The most rigorous methods performed for the estimation of the net effects show that the

programme had a negative effect on the self-employment sustainability of participants. Participants prefer full-time jobs. Non-participants of the programme remained self-employed approximately one month longer than participants in the 2-year-long impact period. Generally, participants as well as non-participants of the programme do not prefer to be placed in part-time jobs. This is probably due to the higher average age of both samples. The previous intervention, focusing on graduate work experience, proved to be interesting for young jobseekers, more or less in the same way as part-time jobs. Part of the evaluation was dedicated to the estimation of the average financial effect that occurred as a result of the distribution of grants to promote self-employment. In the cost-benefit analysis, the financial flows of one participant and one non-participant were compared according to the average time of their employment and unemployment. According to the final outcomes, the intervention had a negative effect on the national budget. Both reference periods pointed to a very similar net financial impact on public finance.

Based on the provided values it is estimated that one programme participant can generate almost 3500 euros less than a non-participant for the national budget. On the other hand, the provided grant was also calculated in the cost-benefit analysis. If the grant was not counted, the net impact would be significantly lower (assigned grants were on average more than 2900 euros). In the first reference period it was estimated that one participant earned approximately 80 euros per month more than jobseekers that started self-employment without a grant from COLSaF. In the second reference period one non-participant of the programme earned 20 euros more than a participant of the same type, but this difference was tested as being insignificant. The overall estimated financial impact of the intervention refers to the negative influence on public finance reaching almost -140 million euros over the evaluated reference period as the difference between participants and non-participants, i.e. net financial impact.

A number of recommendations for both measures are made in the CIE report depending on their economic and administrative circumstances. The main recommendations are as follows:

- COLSaF should actively search for companies and organizations that would better fit the participant's profession. Graduates should have experience in the field in which they studied and graduated. This could be ensured through transparent and clear categorization. The COLSaF should be encouraged to create an electronic system that would identify the economic nomenclature of the organization for a particular group of professions.
- Four-hour working time appears to be insufficient according to the multiple opinions of the programme participants. They claim that the working time was insufficient to manifest their capabilities. The policy makers could start a pilot with a prolongation of working days.
- During the traineeship, some recommendation/certificate could enforce the participants' positions in job interviews as active jobseekers and would upgrade the intervention to a more serious level.
- Self-employment is a rather wide topic exposed to a number of influences determining its success. There are some aspects from the open market that decide whether the established business gets across "the death valley", which is one of the most important initiative stages of the business cycle of any start-up. The relevant information

provided to participants of the intervention would ensure a healthy start and sustainability of the self-employment business.

- It would be helpful to gather and analyze problems of self-employed persons by means of FAQ or an account on a social network site that would represent a place to publish some information concerning the support for the self-employed, the start of cooperation with the Slovak Business Agency or with the National Business Centre (expert counseling, legal counseling, marketing counseling, market experts, accounting counseling, graphics ensuring transmission information about additional funding of the business plans through grants or non-grant schemes, etc. are highly supported by interviewed participants of the intervention).
- Policy makers could pilot an introduction of the selective intervention for jobseekers that have not had any experience with self-employment or with other form of entrepreneurship (by using a limitation of the retrospective assessment of the distinguished criterion). The treatment should be much more complex, especially for the first-time participants of the programme.
- It would be useful to ensure reliable databases to analyze the effects that occurred as a result of the distributed intervention (i.e. collection of data logically complementing each other on different levels, such as level of education of jobseekers, types of schools and fields of specialization; ensuring control mechanisms; using unique official nomenclature to unify data recording, fulfilling all records on jobseekers).
- It is desirable to create direct linkages between COLSaF and SIA to supply data already recorded in SIA. It could simplify the work of regional PES offices and prevent overlapping tasks of managers and officers. These data should be unified by a common methodological procedure.
- SIA should register the identification number of the organization (IČO) of self-employed persons, which is important for the identification of the jobseeker in other official databases of the Financial Directorate of SR, which could provide exhaustive information on financial and economic conditions of businesses.



## Introduction

Increasing the rate of employment and decreasing unemployment were some of the general objectives applied in the Operational Programme Employment and Social Inclusion for the programming period 2007 – 2013 in the Slovak Republic. This objective was set up due to the situation in the country regarding the critically high rate of the unemployed economically active population (13.4% in the year 2006). In this respect, specific measures of ALMP (hereinafter “ALMP”) were proposed to be carried out with the aim of assisting in the improvement of the population's employability.

Traineeship and self-employment are frequently used within active labour market policy measures. Traineeship is an intervention focused on young unemployed jobseekers which occurred as a phenomenon of the financial crisis; the so called “lost generation” according to their weak ability to be placed on the labour market due to their lack of skills. This factor is significant and it is desired that it be eliminated in the Slovak Republic.

On the other hand, it was identified as being necessary to evaluate self-employment according to the previously carried out Pilot assessment of the impact of selected measures of active labour market policy which stated a potential positive net effect of the intervention. The promoting of self-employment is also an actual topic currently taken into account as a trustworthy tool for dealing with the high unemployment rate and lack of free jobs on the open labour market. There are some individuals among jobseekers that need just an initial impulse to start with self-employment. Additionally, this active labour market policy measure is a supplement that contributes to the “Small Business Act” for Europe.

The existence of relevant and credible data was another crucial determining point of the undertaken evaluation. Primarily, we used data from selected interventions provided by the implementation body which is the Central Office of Labour and Social Affairs (here in after “COLSaF”), and the second important data source was evidence from the Social Insurance Agency (here in after “SIA”), which enables the measurement of performance of the individual jobseekers. This administrative evidence would ensure the highest level of validity of conclusions arising from the impact evaluation.

The evaluation used as large a sample as was possible according to available individual data from COLSaF and SIA. The evaluation of the traineeship was applied to 130 thousand participants and non-participants of the intervention, while self-employment was evaluated in the assistance of more than 30 thousand participants and non-participants with comprehensive records.

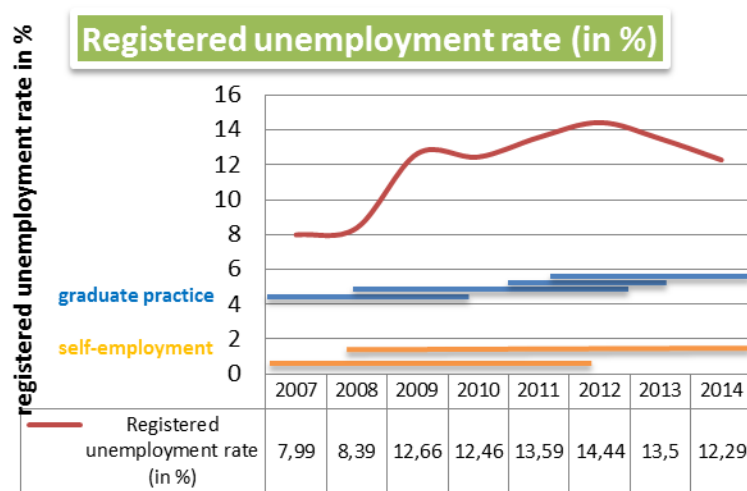
This monitoring report describes the results of the provided four quasi-experimental approaches to the counter-factual impact evaluation methods, of which the main message was to estimate the net effect of the interventions. In other words, this report should find an answer to the fundamental counter-factual question: what would have happened if the intervention had not been provided or promoted? Quite simply, it is possible to say that the methods subtract the individual performance of participants and non-participants in the impact period 2 years after the activities of intervention had finished, or the sustainability period had been complied with. The performed methods established very similar results to the net effects based on the employability of the jobseekers due to the intervention.

Another dimension which has been presented in the evaluation is the net and gross financial effect of the intervention on the national budget in the impact period, estimated

in respect to paid/saved unemployment allowance, taxes, increase of consumption, etc. Last but not least, the evaluation report provides the aggregated opinions of the interviewed respondents that were intervened. The survey has brought forth valuable information about the undertaken activities, which has confirmed the designed theory of change of traineeship and self-employment.

## 1 Slovak labour market at a glance during the period under focus

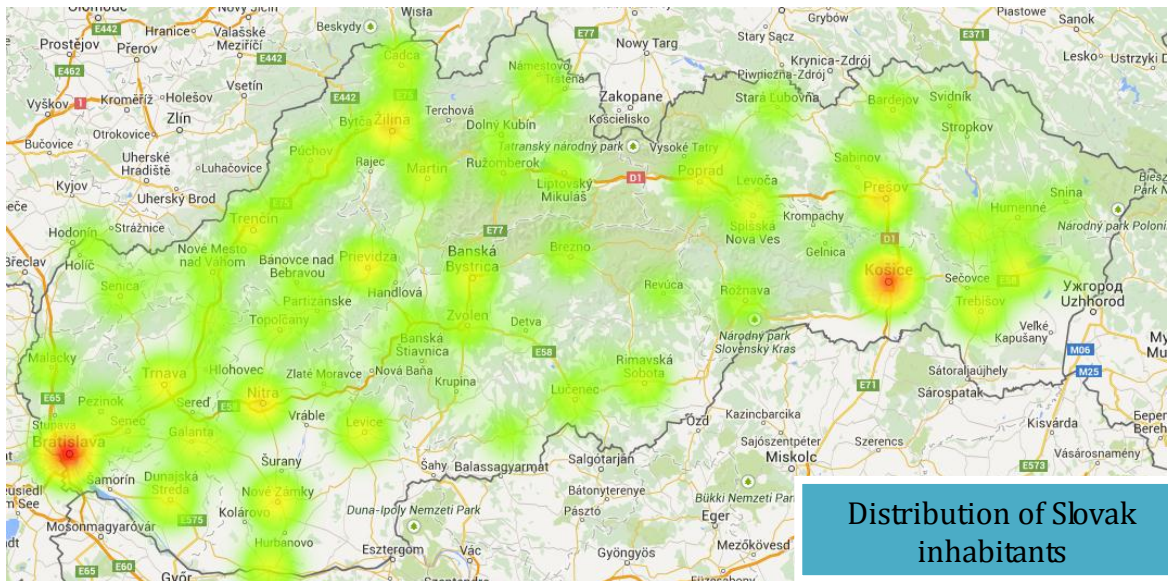
It is an undeniable fact that the Slovak labour market is still suffering from the world economic crisis, as are many European economies. As can be seen in the graph of the total Slovak registered unemployment rate, the lowest rate was measured during the first two years of the period that is covered by this counterfactual impact evaluation. After that, the unemployment rate rapidly increased by almost half and then merely increased till 2011. In 2012 another local extreme appeared where the unemployment rate again started its increasing tendency, which at the end of the year started falling to the level of when the economic crisis started in 2009, which is a signal of the economy and labour market's recovery process. At the bottom part of the chart, miniatures of the Gant charts are presented which describe different reference periods which were designed to ensure homogeneity of the evaluated interventions according to the novelization of the Act on Employment Services under the relevant paragraphs. As can be seen in the first Gant chart, the self-employment promotion has two reference periods. The lines represent the treatment period of the active labour measure (intervention) as well a two-year long sustaining period of self-employment and another two-year long impact period together. The second blue Gant chart describes four reference periods of the traineeship. The blue line represents the treatment period as well a two-year long impact period.



The evaluation period of the self-employment promotion ended in the spring of 2010, when the unemployment rate was at a level higher than 12 %. That is the period when the last financial grants for founding a self-employment licence were distributed, and this was taken into account for the evaluation. The first two years of the self-employment reference periods were years of conjuncture of the Slovak economy. The other reference period of self-employment covered the treatment period of the first wave of the economic crisis.

The first treatment period of the traineeship was also implemented in the period of economic boom, when the lowest level of the unemployment rate was registered. However, the impact was estimated partially in the initiative stage of the world economic crisis. The other three reference periods were implemented mostly during the recession of the Slovak economy and labour market, which is why the first evaluated period achieved on average better results than the rest of the reference periods.

The picture below the text describes the distribution of population density in Slovakia. The red points on the map represent places with the highest number of population (the capital Bratislava and the metropolis of East Slovakia – Košice). The strongest population districts are situated mostly in the south-west and western part of the country; other more populated regions are concentrated in eastern Slovakia.



Distribution of Slovak inhabitants

Source: SO SR, Google fusion tables

The table below the text describes the regional development of three basic labour market indicators: registered unemployment rate, average gross nominal monthly earnings, and employed with workplace outside the SR, which was measured by the Labour Force Survey<sup>1</sup>.

As can be seen, Bratislava region has the lowest unemployment rate in Slovakia and, on the contrary, the highest gross nominal month earnings and, of course, the lowest level of employed outside of the SR. The highest unemployment rate is in the south-central and eastern parts of Slovakia, where are also the highest number of persons employed abroad. The extreme average gross income is in Bratislava region and in other parts there are averages distributed almost equally in the regions of the SR.

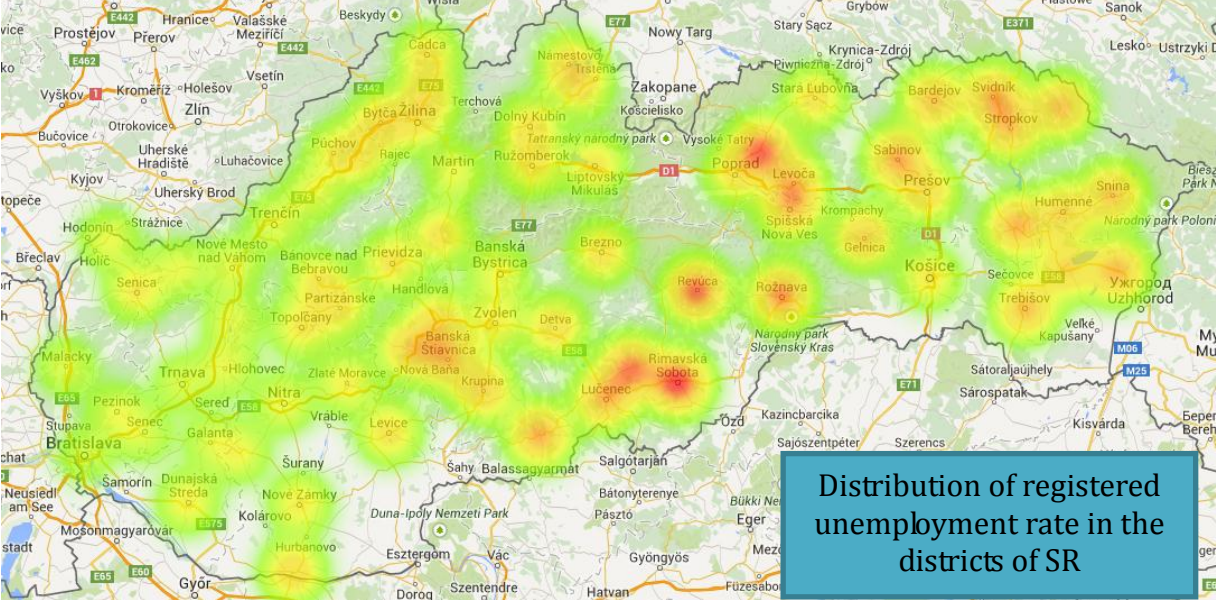
Extremes of people that find a job abroad are visible in the Prešov, Žilina and Nitra regions, where are also the highest share of jobseekers with occupations in construction, unskilled occupation or auxiliary occupations. These are very frequent and traditional kinds of occupation characteristic mainly in Kysuce region, Orava and Prešov regions.

Region	Registered unemployment rate (in %)								Average gross nominal monthly earnings (EUR)				Employed with workplace outside the SR (LFS)							
	2007	2008	2009	2010	2011	2012	2013	2014	2010	2011	2012	2013	2007	2008	2009	2010	2011	2012	2013	2014
Region of Bratislava	1,98	2,27	4,36	4,63	5,41	5,72	6,17	6,13	1116	1157	1184	1205	5,1	4,6	4,1	3,1	4,1	4,7	7,6	5,9
Region of Trnava	4,3	4,29	8,37	8,17	8,88	9,43	9,16	8,03	789	819	848	860	10,7	8	5,4	5,2	4,1	5,5	6,6	4,8
Region of Trenčín	4,5	4,95	10,13	9,51	9,95	10,89	10,74	9,56	739	766	798	821	13,7	14,2	10,6	11,1	11,1	10,9	11,4	8,6
Region of Nitra	7,1	7,41	11,72	11,76	13,27	14,08	12,52	11,21	738	742	776	789	33,1	31,2	27,1	28,2	23,1	18,8	21,9	23,1
Region of Žilina	5,55	6,2	10,89	10,86	11,91	12,79	12,51	10,91	756	783	816	839	27,1	24,2	19,6	20,8	16,3	18,8	20,3	23,6
Region of Banská Bystrica	14,1	14,25	19,19	18,86	19,83	20,81	18,26	17,22	719	740	772	798	17,3	17	11,9	10,4	11,7	11,2	14,8	16,5
Region of Prešov	12,05	12,86	18,29	17,75	18,95	20,66	19,35	17,45	672	680	718	736	52,1	47,7	33,7	32	32,9	35,4	38,7	36,1
Region of Košice	13,02	13,5	17,3	16,78	18,76	19,58	17,23	15,92	799	814	853	883	18,3	20,9	16,7	16,1	11,8	15,3	15,1	15,5

Source: Statistical Office of the Slovak Republic

1- LFS is the continuous monitoring of labour based on direct surveys in selected households. The basis for the Labour Force Survey consists of a stratified selection of apartments, which evenly covers the whole territory of the Slovak Republic. To sample the quarter included 10,250 dwellings, which represents 0.6% of the total number of permanently occupied dwellings in the Slovak Republic.

The final map additionally presents the distribution of the registered unemployment rates across the Slovak districts. To compare with the previous heat map, it is obvious that the highest unemployment rates occur mostly in the less populated parts of Slovakia. On the maps it can be seen that districts exposed to the highest levels of the unemployment rate are located in the central south of Slovakia and in the east of the country. The unemployment rate is reduced in the districts closest to the capital Bratislava.



Source: SO SR, Google fusion tables

## 2 Data source

The intention of the evaluators was to use all relevant and available data sources about all treated and eligible controls. That is the reason why the evaluators applied for the data census of all eligible treated and non-treated jobseekers for § 49 and 51 registered in the database during the time period covered by the evaluation.

A number of data sources were identified. Firstly, the most important database was the database of treated and non-treated jobseekers maintained by COLSaF and supported by regional Public Employment Services offices. This database has the main purpose of providing us with identification of treated and non-treated individuals, information about verifying the eligibility of jobseekers, time period of treatment, amount of grant, etc.

The other most important data source was the database of the Slovak Insurance Agency, which provided mostly dependent variables helping to verify the employability of the treated and controls, the amount of wages earned during individual impact periods, types of employer, or data which could partially uncover the reasons why jobseekers could not find a placement on the open labour market through type of registrations. The other effect of the data is verification and addition of some incorrect or missing variables, such as gender, date of birth, or permanent residence.

COLSaF provided a database of personal identification numbers of all jobseekers who were registered during the focused period of evaluation to the Social Insurance Agency. The Social Insurance Agency extracted all records of jobseekers and prepared all necessary data for evaluators in accordance with Act No. 122/2013 Coll. on Protection of Personal Data and on Changing and Amending of other acts, resulting from amendments and additions executed by Act. No. 84/2014 Coll. Any selected jobseekers in the treated and control groups were not treated by any other intervention, except intervention by § 46 - Education and training for the labour market of jobseekers which was complementarily realized to the intervention by § 49 - Self-employment in the preparatory process for business.

COLSaF is a government entity, ensuring the execution of state administration in the field of social affairs and employment services. The institution was established in January 2004, via Act No.453/2003 on state administration bodies in the field of social affairs and employment services, as amended. The headquarters ensures management, control, coordination and methodological guidance performance through 46 offices of Labour, Social Affairs and Family.

### 2.1 Data preparation

This chapter describes the process of modification, categorization and coding of variables from individuals in treated and non-treated groups which we obtained from COLSaF and the Social Insurance Agency, Statistical Office of the Slovak Republic, Slovak Information and Marketing Company and the University of Žilina.

We decided to group data into four fundamental categories according to the type of information they provide in the context of log frame intervention.

The first type of data is **inputs** – there belongs data as sources which were used for the treatment effect. The basic data source for this kind of data was the database of COLSaF.

The second sort of data is **outputs**, which monitor identification of treated and non-treated groups, time periods of treatment, and places where active labour market

measures were carried out. The main data source for this information is the database of COLSaF about jobseekers and, partially, the database of the Social Insurance Agency. The third kind of data is **outcomes**, which monitor the employability of jobseekers and the success of placement on the open labour market through wages. The data source for this kind of information is the database of registrations of the Social Insurance Agency. The fourth sort of data informs us about conditions (**context data**) on the local labour market in the regions where the unemployed seek their jobs. There is some other information about population in the regions, number of municipalities and cities, etc.

**2.1.1 Input and output data (treatment variables)**

In general, the data extracted from COLSaF refers to inputs and outputs of both interventions. There is data about the identification of individuals that were treated and jobseekers that are potentially incorporated into our controls. There is also some information about the direct outputs of interventions from the end of registration or SK NACE of an employer where graduates carry out their work experience, amount of grants, etc.

In total, we obtained 2,886,510 records from COLSaF. In the dataset, one jobseeker has multiple records about different registration periods. The data contains only jobseekers

that were not exposed to multiple interventions, i.e. jobseekers who were supported by other than the evaluated intervention were excluded. The tables below present frequency statistics about the dataset from COLSaF.

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Age	2886266	100,0%	244	0,0%	2886510	100,0%

**1) Independent variable: Gender**

The total of values is 2,886,510 cases. Less than 0.1 % of population filled in the incorrect value “1” in the dataset and 0.1 % of presented cases referred no value. More than 53 % of the treated and non-treated records are men and less than 47 % are women.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid				
Incorrect value	1146	,0	,0	,0
Men	1538344	53,3	53,3	53,3
not identified	3173	,1	,1	53,4
Women	1343847	46,6	46,6	100,0
Total	2886510	100,0	100,0	

## 2) Independent variable: Age

Individual records present participant's age at the first registration in the database of jobseekers. It was the recorded age at first registration in the case of multiple registrations in the database of jobseekers. Therefore, the values of the variable are shifted by the difference between the two dates of the beginning of the records into the database of jobseekers. (i.e.

the difference between the beginning of the 2<sup>nd</sup> time and the beginning of the 1<sup>st</sup> time were added to the first age and thus gradually further for all records of the jobseeker). The values of age are rounded up to two decimal places to eliminate rounding up errors. We excluded jobseekers whose records did not meet the eligibility criteria at the reference time for traineeship (less than 25/26 years of age).

		Statistic	Std. Error	
Age	Mean	39,3130	,00755	
	95% Confidence Interval for Mean	Lower Bound	39,2982	
		Upper Bound	39,3278	
	5% Trimmed Mean	39,0148		
	Median	37,0000		
	Variance	164,596		
	Std. Deviation	12,82951		
	Minimum	,00		
	Maximum	85,60		
	Range	85,60		
	Interquartile Range	21,67		
	Skewness	,342	,001	
	Kurtosis	-1,017	,003	

The dataset from COLSaF contains just 244 cases without referring to years of age, but as is presented in the table above, the minimum value is zero years, which indicates some incorrect records. These records must be merged with data from the Social Insurance Agency, otherwise these records (jobseekers) must be excluded from our sample.

## 3) Independence variable: Marital status

Marital status is information based on the time of the registration of the jobseeker before the intervention was granted.

Almost every second registration of jobseekers is single and about 40 % of jobseekers' registrations are married. More than 9 % of jobseekers registrations are divorced and more than 1.5 % of registrations of jobseekers are widowers. The minority of the registrations subscribes to registered partners, only about 0.04 %. More than 7 thousand jobseekers' registrations do not specify their marital status and they will probably be excluded from our dataset.

		Frequency	Percent
Valid	not identified	7115	,25
	registered partners	1158	,04
	divorced	267095	9,25
	single	1425824	49,40
	widow/er	45434	1,57
	married	1139884	39,49
	Total	2886510	100,00

## 4) Independent variable: Permanent residence

Nomenclature of Units for Territorial Statistics (here in after "NUTS") - the code was reduced from 5 digits to just 3 digits (regional permanent address) and to 4 digits representing the district of permanent residence of the jobseeker. Those digits are sufficient for the matching and statistical description of individuals in treated and non-treated groups. The permanent residence is missing in 0.5 % of records. That information



must be obtained from the dataset of the Social Insurance Agency otherwise the jobseekers must be excluded from the sample.

		Permanent residence_region			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Missing	13913	,5	,5	,5
	Bratislavský kraj	206659	7,2	7,2	7,6
	Trnavský kraj	273546	9,5	9,5	17,1
	Trenčiansky kraj	292784	10,1	10,1	27,3
	Nitriansky kraj	380836	13,2	13,2	40,5
	Žilinský kraj	343911	11,9	11,9	52,4
	Banskobystrický kraj	410572	14,2	14,2	66,6
	Prešovský kraj	505232	17,5	17,5	84,1
	Košický kraj	459057	15,9	15,9	100,0
	Total	2886510	100,0	100,0	

#### 5) Independent variable: Temporary residence

This variable has been excluded from the data set. Only a limited number of cases indicated information about temporary residence. The information was not significant from a statistical point of view and experience from previous examination of its significance in the process of modelling dependence.

#### 6) Independent variable: Level of education

This variable represents the highest achieved level of education of the jobseeker according to the International Standard Classification of Education (hereinafter "ISCED"). In our dataset, almost 18 % of the records are without this information. This variable will not be excluded at the moment. We will decide on it during the next evaluation process.

		Level of education			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not finished education	15991	,6	,7	,7
	Primary education	229596	8,0	9,7	10,4
	Lower secondary professional education	32742	1,1	1,4	11,7
	Secondary vocational education	804982	27,9	33,9	45,7
	Full secondary vocational education	839439	29,1	35,4	81,0
	Full secondary comprehensive education	117690	4,1	5,0	86,0
	Upper vocational education	5093	,2	,2	86,2
	Bachelor	29984	1,0	1,3	87,5
	Master	293629	10,2	12,4	99,8
	Doctoral	3601	,1	,2	100,0
	Total	2372747	82,2	100,0	
Missing	System	513763	17,8		
Total		2886510	100,0		

### **7) Independent variable: School specialisation field**

This variable was recorded into the system in two ways. The first was based on the individual description of jobseekers about the field of specialisation at their highest level of education. The second way of recording the field of specialisation was carried out via the 7 digits of the school specialisation field code. Those different approaches of reporting the field of specialisation caused an enormous number of different specialisation categories. The variable was used as a starting point for the creation of the next variables representing the education of the jobseekers.

The independent variable is connected to the previous variable – level of education, which is the reason the dataset contains almost 18 % of missing values.

### **8) Independent variable: Type of school**

This variable represents the last attended school of the jobseeker. The codes of types of schools were categorized into several categories of schools. The types of schools varied mainly at the level of secondary and tertiary education. For example, universities were sorted into categories such as technical, social, economic, police, health, art, etc. Secondary schools were sorted into comprehensive school, girls secondary school, business academy, conservatory, etc.

This independent variable is connected to the previous variable – level of education, which is the reason the dataset contains almost 18 % of missing values. One third of jobseekers have, as their highest level of education, secondary vocational school, or vocational school.

### **9) Independent variable: Code of degree program**

Another variable which was deduced from the School specialisation field is "Code of degree program", which originally contained a 7 digit code that was reduced to a 4 digit code. That is why the code represents just a degree program. Seven digits were used in a small number of records, which is another reason why we decided to reduce the number of digits in the code. Additionally, we assumed that through this reduction we would ensure easier matching of treated and controls if the variable was significant in our model. The independent variable is connected to the previous variable – level of education, which is the reason the dataset contains almost 18 % of missing values.

### **10) Independent variable: Driving license**

This variable represents the type of driving licence of registered jobseekers, composed of treated and non-treated individuals. We deduced from this variable another 16 dummy variables of driving licence categories because we assumed that there would be a significant difference between a jobseeker that has a driving licence for lorries and a jobseeker that has a driving licence just for cars. It could be a significant ability which excludes the jobseeker with a driving licence just for cars from free working positions in the transport industry. About 30 % of jobseekers had a driving licence for cars and about 6 % of jobseekers had a permit to drive lorries.

No.	Type of driving license	Frequency		Percent		Total
		no	yes	no	yes	
1	Drivin license: group DE	2 883 212	3 298	99,9	,1	2 886 510
2	Drivin license: group D	2 865 513	20 997	99,3	,7	2 886 510
3	Drivin license: group D1E	2 883 029	3 481	99,9	,1	2 886 510
4	Drivin license: group D1	2 865 513	20 997	99,3	,7	2 886 510
5	Drivin license: group CE	2 821 364	65 146	97,7	2,3	2 886 510
6	Drivin license: group C	2 715 572	170 938	94,1	5,9	2 886 510
7	Drivin license: group C1E	2 821 364	65 146	97,7	2,3	2 886 510
8	Drivin license: group C1	2 715 572	170 938	94,1	5,9	2 886 510
9	Drivin license: group BE	2 821 364	65 146	97,7	2,3	2 886 510
10	Drivin license: group B	2 021 902	864 608	70,0	30,0	2 886 510
11	Drivin license: group B1	2 021 902	864 608	70,0	30,0	2 886 510
12	Drivin license: group A	2 633 956	252 554	91,3	8,7	2 886 510
13	Drivin license: group A2	2 886 453	57	100,0	,0	2 886 510
14	Drivin license: group A1	2 633 956	252 554	91,3	8,7	2 886 510
15	Drivin license: group AM	2 009 864	876 646	69,6	30,4	2 886 510
16	Drivin license: group T	2 695 510	191 000	93,4	6,6	2 886 510

### 11) Independent variable: disadvantages

This variable represents categories of disadvantages stated in Act No. 5/2004 Coll. on Employment Services, § 8 Disadvantaged jobseekers. These are categories such as jobseekers of more than 50 years of age, graduates, long-term unemployed, disabled etc. As presented in the table, most of the records have no attribute of a disadvantage. Just about 10 % of the records had a symptom of disadvantage. These were the long-term unemployed, graduates and jobseekers of more than 50 years of age in most of the cases.

		Disadvantages		
		Frequency	Percent	Cumulative Percent
Valid	no disadvantage	2599151	90,0	90,0
	graduate	65400	2,3	92,3
	unemployed	17	,0	92,3
	foreigner	1	,0	92,3
	long-term unemployed	180783	6,3	98,6
	not-finished	307	,0	98,6
	low education	494	,0	98,6
	organizational	3508	,1	98,7
	drop of capability	29	,0	98,7
	termination	297	,0	98,7
	finished	13	,0	98,7
	migration	1	,0	98,7
	care	2464	,1	98,8
	hardship	419	,0	98,8
	age above 50	31054	1,1	99,9
	health	110	,0	99,9
	disability	2462	,1	100,0
Total	2886510	100,0	.	

### 12) Independent variable: occupation

This variable represents the International Standard Classification of Occupations (hereinafter "ISCO") of the jobseeker. The code was reduced from 7 digits (which was available just for a limited number of cases) to a 2 digit code. In the table below are presented categories of occupations. We eliminated the difference in the monitoring of this variable. We reduced the code to 2 digits because there were less than a thousand records which had records just with 1 digit. Finally, we grouped the jobseekers into 45 categories which should be appropriate for the matching. Most of the records tell us that jobseekers are support staff in mining, construction, manufacturing and transport, or sales assistants or administrative staff.

There are slightly more than 30 % of records without values for occupation.

Name of occupation	Frequency	Percent	Cumulative Percent
Office workers	12	,0	,0
Workers in services and trade	5	,0	,0
Skilled workers and artisans	2	,0	,0
Operators, and assemblers of machinery and equipment	3	,0	,0
Elementary occupations	232	,0	,0
Legislators, senior government officials and senior representatives of enterprises and organizations	3273	,1	,2
Managers (managers) administrative, support and business activities	9480	,3	,6
Managers (managers) Production and specialized services	8077	,3	1,0
Managers (managers) in accommodation, dining, business and other services	23046	,8	2,2
Specialists in the field of science and technology	19212	,7	3,1
Health professionals	5298	,2	3,4
Teachers and professionals in education	34033	1,2	5,1
Specialists administrative, support and business activities	16643	,6	5,9
Specialists in the field of information and communication technologies	5063	,2	6,2
Legal professionals, social and cultural	11645	,4	6,8
Technicians and associate professionals in the field of science and technology	47016	1,6	9,1
Health professionals	14689	,5	9,8
Professors administrative, support and business activities	134681	4,7	16,5
Professionals in the legal, social and cultural and related workers	8598	,3	16,9
Technicians in the field of information and communication technologies	8812	,3	17,4
General office clerks and registrars	53938	1,9	20,0
Clerks Customer services	19817	,7	21,0
Clerks to record the number and store data	49524	1,7	23,5
Other office staff	12674	,4	24,1
Personal service workers	114137	4,0	29,8
vendors	168013	5,8	38,1
Workers in custody	27088	,9	39,5
Employees of public safety and security services	30279	1,0	41,0
Skilled workers in agriculture (market-oriented)	13951	,5	41,7
Skilled forestry, fishing and hunting (market-oriented)	13913	,5	42,4
The farmers, fish farmers, hunters and gatherers	573	,0	42,4
Skilled craftsmen and construction workers, excluding electricians	121407	4,2	48,4
Skilled workers in metallurgy, engineering, and related workers	112819	3,9	54,0
Art and handmade artisans and printers	13427	,5	54,7
Electronics engineers and electricians	26509	,9	56,0
Processors and producers of food products, wood products and clothing	95798	3,3	60,8
Operators of stationary machinery and equipment	73017	2,5	64,4
assemblers	87835	3,0	68,7
Drivers and mobile plant operators	94906	3,3	73,5
Cleaners and helpers	52159	1,8	76,0
Laborers in agriculture, forestry and fisheries	33124	1,1	77,7
Labourers in mining, construction, manufacturing and transport	366123	12,7	95,9
Labourers in food preparation	2620	,1	96,0
Street vendors and auxiliaries similar services	1314	,0	96,1
Workers in waste disposal and other unskilled workers	79178	2,7	100,0
Total	2013963	69,8	
Missing	872547	30,2	
Total	2886510	100,0	

### 13) Independent variable: Period of registration

This variable tells us how long a jobseeker was unemployed before the starting date of the reference period of this impact evaluation, i.e. 1.1.2007. All the values have been recoded into four simple variables because the variable measured the days of registration in the register of jobseekers. The values categorize jobseekers into these categories:

- non-registered jobseekers before 1.1.2007 (non-unemployed),
- jobseekers registered less than 1 year before the reference period of the evaluation,
- jobseekers registered more than 1 year and less than 3 years in the PES register, and jobseekers registered more than 3 years before the reference period.

registration before			
	Frequency	Percent	Cumulative Percent
non-unemployed	1044571	36,2	36,2
< 1 year	521185	18,1	54,2
1-3 years	554376	19,2	73,4
> 3 years	766378	26,6	100,0
System	1842810	63,8	
System	1043700	36,2	
Total	2886510	100,0	

### 14) Independent variable: SK NACE

This variable represents the structure of the Slovak classification of economic activities of the last employers of registered jobseekers. The code was reduced from 5 digits to 2 digits because of the infrequency of the full 5 digit code. Through reduction the cases are equal. More than 96 % of records do not contain a value for SK NACE. This is because the records represent controls that have not been supported by any measure of ALMP (SK NACE is a figure only for self-employed persons supported by ALMP).

### 15) Independent variable: NUTS of measure performance

This variable represents Nomenclature of Units for Territorial Statistics of the region, or districts where jobseeker performance was measured by ALMP. The code was reduced and equalized to a 3 digit code representing regions of Slovakia and a 4 digit code representing districts of Slovakia. The table next to the text shows that in our dataset there are more than 96 % of the records without values for regions where the ALMP measure was performed.

	Frequency	Percent	Cumulative Percent
missing	2786494	96,5	96,5
Bratislava region	3804	,1	96,7
Trnava region	9634	,3	97,0
Trencin region	10304	,4	97,4
Nitra region	12415	,4	97,8
Zilina region	14575	,5	98,3
Banska Bystrica region	15186	,5	98,8
Prešov region	20880	,7	99,5
Kosice region	13218	,5	100,0
Total	2886510	100,0	

### 16) Dependent variable: Date of entry

This variable represents date of entrance into the database of jobseekers at local Public Employment Services offices between 1st January 2007 and 31st July 2014.

### 17) Dependent variable: Date of departure

This dependent variable represents the date of departure from the database of jobseekers. In the cases without a value we added the date 30.6.2014 as the last day of the reference period of the impact evaluation. The cases without values are still registered in the

database of jobseekers.

**18) Dependent variable: Time period of registration according to dates**

This is an additional, deduced variable, which represents the time period of registration in the database of jobseekers as the difference between the disposal date and the registration date of jobseeker in months. The variable was used as the control value for checking the eligibility criteria of the self-employment intervention, i.e. a minimum three months registration of jobseekers in the database.

**19) Dependent variable: Decommissioning due to departure abroad**

This dummy variable reports the reason of decommissioning from the jobseekers register due to the departure of jobseekers abroad. If a jobseeker departed abroad, he is likely placed on the open labour market abroad otherwise the jobseeker would return after some period of time and again return to register in the database of jobseekers.

		Frequency	Cumulative Percent
Valid	yes	414	,01
	no	2886096	100,0
	Total	2886510	

There are slightly more than 400 registrations that indicate the departure of jobseekers abroad. This variable was voluntarily reported. Exclusion will be considered.

**2.1.2 Outcome data (treatment characteristics)**

Data from the Social Insurance Agency is mostly output data in the context of an intervention log. In the database was found also output data. This data set contains inputs such as gender, permanent residence and date of birth.

Data from this institution was rather comprehensive because there were almost 210 mil. registrations for more than 3 mil. individuals. The process of data preparation was accompanied by a number of problems in scripting and removing errors which occurred during the extraction process from the data storage of the Social Insurance Agency.

Finally, we selected more than 28 mil. registrations of individuals that were identified in the COLSaF database.

This data contained also some independent variables which were used in the COLSaF database, which is why we could test the accuracy of data and add missing data in variables: date of birth, gender and permanent residence. Through that process we eliminated deleting some cases which would be excluded from the dataset of the treated and non-treated.

However, mostly the data monitored dependent variables based on employability. Through categorisation of registrations in the Social Insurance Agency, we could estimate and eliminate cases when individuals have an objective barrier to employment on the open labour market. We distinguish these categories of registration of insured persons in the Social Insurance Agency thus:

- **Placed on the labour market** – these are registrations as employee, or voluntarily insured person. We assume that if somebody can pay insurance payments, he/she would have the financial resources to do so. There are also mothers and fathers on maternity/paternity leave because they have temporarily interrupted their employment.
- **Partially placed on the labour market** - these are persons employed part-time.

- **Self-employed** persons,
- **Persons who are outside of the labour market due to occurred individual barriers** such as caring for a child, receiving a disability pension, being a personal assistant, etc. These types of registration indicate to us that the jobseeker was forced by a life event to stay out of the labour market and the social aspect is a barrier for his/her entrance into the open labour market.
- **Assessment base/wage** which is the monthly income of the individual in Euros or average income per month of the self-employed.

The table below describes in detail the registrations in the proposed categories of registrations in the Social Insurance Agency which will create the fundamental variables for verification of a jobseeker's employability.

Type of registration	Category of registration
ZEC - employee	placed on the labor market
ZECN - employee with irregular income	placed on the labor market
ZECDN10S - staff who were long-term unemployed	placed on the labor market
ZECD1PR - part-time agreement of service	partially located on the labor market
ZECD1N - part-time irregular income - agreement of service	partially located on the labor market
ZECD2PR - do part-time agreement on work activities	partially located on the labor market
ZECD2N - part-time irregular income - agreement on work activities	partially located on the labor market
ZECD3 - part-time student work	partially located on the labor market
ZECD3N - part-time irregular income - student work	partially located on the labor market
SZC - self-employed	self-employed
DPODP - voluntarily insured person on supplementary insured	placed on the labor market
OVS - person performing SS, NS, ZDS	placed on the labor market
OCS - person performing community service	placed on the labor market
DIE6R - looking after a child under 6 years	due to subjective reasons outside the labor market
DIE7R - looking after a child under 7 years	due to subjective reasons outside the labor market
DIE18R - looking after a child under 18 years	due to subjective reasons outside the labor market
OID - receiving disability pension	due to subjective reasons outside the labor market
DPPS - additional premium payer for supplementary insured (student)	due to subjective reasons outside the labor market
DPPN - additional premium payer for supplementary insured (unemployed)	unemployed
DPPP - additional premium payer for supplementary insured (interruption insurance)	placed on the labor market
PUR - recipient of accident benefit	due to subjective reasons outside the labor market
POP - recipient of care allowance	due to subjective reasons outside the labor market
OSA - Personal Assistant	due to subjective reasons outside the labor market
FOMAT - ZEC, SZC at the time the maternity / parental leave	placed on the labor market

Source: Social Insurance Agency

### 2.1.3 Context data (control variables)

Context data comes from the Slovak Statistical Office of the Slovak Republic and it will be used mostly for description and interpretation of conclusions based on different analysis of the evaluation. There is data about the unemployment rate at the different NUTS.

Other data came from the University of Žilina, in particular a matrix of real distances between Slovak towns and villages in kilometres. The data was used to measure individual distance from the municipality of permanent residence to the regional PES office. That database was fundamental for the creation of one instrumental variable that was used for the estimation model of propensity score matching method.

## 2.2 Selection bias

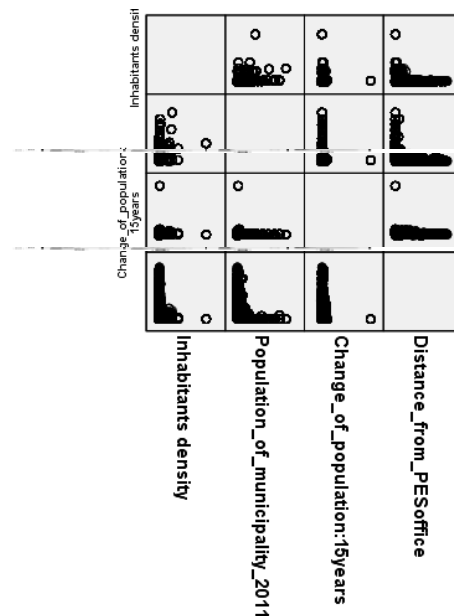
As already mentioned, the intervention promoting traineeship is obligatorily distributed to any eligible jobseekers that apply. This is the reason why the intervention indeed suffers from selection bias, namely that the impact is potentially affected by a self-selection bias effect. Therefore, it is necessary to struggle with unobserved characteristics (variables) which could potentially influence the estimated average treatment effects. One of the most significant unobserved variables could be the motivation to participate in the intervention based on the circumstances of the individuals. We can assume that young jobseekers are primarily motivated to find a job according to the general situation on the labour market in the place where they live. There are also some other important influences on employability such as having good luck, ability to convince people, availability of relevant information, and also random factors, etc. All the named sources are very hard or impossible to quantify and match with individuals in our treated and control samples.

We tried to identify some proxy indicators which would identify the differences between treated and non-treated groups to find the best possible logistic regression model that would help us credibly estimate the individual probability of participants and controls to be covered by the intervention. We focused on the data which could be possibly matched to the individuals according to the available data in the datasets from COLSaF and SIA. We proposed using these four instrumental variables:

- population of the municipality from the last Slovak census in 2011,
- change of the population in the last 15 years in the municipality and
- real distance from permanent residence to the local Public Employment Services office where the individual belong and
- inhabitants' density in the municipality.

The assume that potential instrumental variables describe the local environment of the

individual jobseeker from the potential of the locality to create new jobs, and variety of occupations. In the municipality there could be barriers for individual jobseekers to match with professions which are based on the limited number of employees in the municipality or region. Trends in the population of the municipality over the last 15 years could also provide information about the general motivation of the inhabitants to find a job possibly in another part of the region or Slovakia for many reasons. Some parts of Slovakia have become, in recent years, mainly resources or tourism locations. That indicator should collect information about the socio-economic climate of individuals' municipality. The next important instrumental variable could be the number of kilometres between permanent residence and PES office. Local public employment offices are usually in the cities which are also social, cultural and economic centres in the locality. We assume that





distance could be a problem for some graduates to travel regularly in order to visit the open labour market and to be in touch with it.

In the scatterplot matrix outliers are marked which were identified

- in Bratislava V (part of the capital) where the highest number of permanent inhabitants is situated;
- in Bratislava I (old town) where inhabitant density is extreme and
- Selce (a municipality with a more than 1500 % increase of inhabitants in the last 15 years; it is a municipality near to Banská Bystrica).

These outliers were eliminated and we constructed a new scatterplot matrix which describes the shape of the function of the proposed instrumental variables.

Before calculating the correlation coefficients it is useful to show the relationships between variables graphically. For the input variables can be used scatter plot matrix, which consists of scatterplots for all pairs of given variables.

From the graph, we can check whether the data contains outliers or other kinds of problems that could further distort the results. At the same time, we can create an idea about the relationships between variables.

The correlation matrix contains, for each pair of input variables, Pearson's linear correlation coefficient (Pearson Correlation) values and a significance test of the zero rate (Sig. (2-tailed)).

Correlation coefficients significantly different from zero are indicated with an asterisk in the table (One star corresponds to the non-zero at 95% confidence level, two stars 99% confidence level).

As is obvious in the matrix, all the Pearson coefficients are estimated to be significantly different from zero at 99 % confidence level. Despite the fact that all the correlation coefficients are calculated based on confidence levels of 99%, depending on the

		Correlations			
		Inhabitants density	Population_of_municipality_2011	Change_of_population:15years	Distance_from_PESoffice
Population_of_municipality_2011	Pearson Correlation	,196**			
	Sig. (2-tailed)	0,000			
	N	116292			
Change_of_population:15years	Pearson Correlation	-,130**	-,172**		
	Sig. (2-tailed)	0,000	0,000		
	N	116197	116197		
Distance_from_PESoffice	Pearson Correlation	-,113**	-,338**	,019**	
	Sig. (2-tailed)	0,000	0,000	,000	
	N	116292	116292	116197	
The registered unemployment rate in the district of perm. residence	Pearson Correlation	-,287**	-,228**	,061**	,156**
	Sig. (2-tailed)	0,000	0,000	,000	0,000
	N	116292	116292	116197	116292

\*\* . Correlation is significant at the 0.01 level (2-tailed).

individual instrumental variables they are fading, or very weak. The correlation coefficients were much weaker after elimination of the outliers mentioned in the text above (max. 0.265).

In the next step it should be verified whether there are identified differences between treated and non-treated groups across the designed reference periods for both interventions. Because, if there are significant differences between both groups, there is reason to expect that some of the proposed instrumental variables could be a satisfactory proxy indicator. This indicator quantifies unknown unobservable factors which could determine the participation of the individuals in the intervention.

The table below describes the results of the independent samples from the Kolmogorov-Smirnov tests during the reference periods. At the significance level of 0.05, we can write

the statistical statement that all instrumental variables do not have the same distributions between treated and non-treated groups of jobseekers. In the other words, in the samples of traineeship of participants and their controls, there are significant differences in inhabitants' density, change of the population in the municipality over the last 15 years, population of the municipality or individual real distance to the PES office.

Null Hypothesis	Test	Sig. In ref. Period 1	Sig. In ref. Period 2	Sig. In ref. Period 3	Sig. In ref. Period 4	Decision
The distribution of <i>Inhabitants density</i> is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	0,000	0,000	0,000	Reject the null hypothesis.
The distribution of <i>Population_of_municipality_2011</i> is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	0,000	0,000	0,000	Reject the null hypothesis.
The distribution of <i>Change_of_population:15years</i> is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	0,000	0,000	0,000	Reject the null hypothesis.
The distribution of <i>Distance_from_PESoffice</i> is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	0,000	0,001	0,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

In the other table below are presented the results of the same test which are the same as were in the traineeship. Just one test retains the null hypothesis in the first reference period of the instrumental variable "change of the population in the municipality in the last 15 years." Although the result states that differences between the treated and control groups are not significant, we will use that instrumental variable for the model of logistic regression.

Null Hypothesis	Test	Sig. Ref. Period 1	Decision. Ref. Period 1	Sig.. Ref. Period 2	Decision. Ref. Period 2
The distribution of <i>Inhabitants density</i> is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,008	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of <i>Population_of_municipality_2011</i> is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of <i>Change_of_population:15years</i> is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,277	Retain the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of <i>Distance_from_PESoffice</i> is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,001	Reject the null hypothesis.	0,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Additionally, we decided to eliminate the influence of self-selection bias through a narrower selection of controls for self-employment. We assumed that the motivation to join in the intervention could be partially ensured through the selection of eligible controls which were:

- registered in the register of jobseekers in the reference period<sup>2</sup>,
- not supported through intervention or another intervention and
- self-employed during the reference period plus 2 years, which represents the compulsory sustainable period of self-employment according to the record in the SIA database.

<sup>2</sup> Reference period represents a specific time of intervention homogeneity which was taken into account for evaluation reasons. For example: from 1.1.2007 till 30.4.2008.

The limitation of the presented process of selecting jobseekers into controls is that our counterfactual evaluation should answer just one question: what would have happened if the intervention had not been provided to any jobseeker who intends to become self-employed. The reason is that through this selection we will compare just the controls – jobseekers who really wanted to become self-employed same as the supported jobseekers.

## 3 Methodology

This chapter describes the theoretical approaches which were applied across the provided evaluation of both evaluated active labour market measures (§ 49 and § 51). It is necessary to emphasise that this impact evaluation report should also have a learning purpose which is reflected in the selection methods. Through the use of different types of methods, we would like to use the differences in estimated net-effects. Basically, this report should cover the combination of the counter-factual methods from the less robust to the more robust and technically challenging ones.

### 3.1 Hierarchical cluster analysis

Analysis allows the generation of groups of cases (rows of the data matrix) or variables (columns of the data matrix) such that the elements within the groups were as homogeneous as possible and elements between the groups were as different as possible. Input variables can be numeric, dichotomous or express frequency. Hierarchical clustering is based on the gradual merging of the closest pair of cases or clusters that have formed into one - each step merges one pair and the distance matrix is recalculated for the newly formed group. The algorithm is continued until all of the cases are in clusters.

### 3.2 Parametric and non-parametric tests

In statistics, the Kolmogorov – Smirnov test is a non-parametric test for testing the equality of continuous probability distributions that can be used to compare two samples. The Kolmogorov–Smirnov statistic quantifies a distance between the empirical distribution functions of two samples. The empirical distribution function is a step function, which counts a cumulative share of elements in the sample with ordered values. Two empirical distribution functions of two samples are then compared in each value and the supremum of the differences is compared with a table of critical values of this Kolmogorov – Smirnov test. The null distribution of this statistic is calculated under the null hypothesis that the samples are drawn from the same distribution. This two-sample test is one of the most useful and general non-parametric methods for comparing two samples.

### 3.3 Correlation

Correlation characterizes the relationship of two numeric or ordinal variables. This relationship is expressed by the correlation coefficient.

Pearson's linear correlation coefficient measures the degree of linear dependence of two numeric variables. Before calculation it is necessary to determine whether the data contains outliers that might skew the conclusions reached. This type of rate is not appropriate where, for the variable, there exists another type of addition other than linear.

Pearson's linear correlation coefficient takes values in intervals from -1 up to 1. If the absolute value equals one, the data is exactly on a straight line. A correlation coefficient

equal to one is characterized by a direct proportion (the line is growing); a correlation coefficient equal to minus one corresponds to the inverse (the line is declining). In examining the actual data, however, these cut-off values of the correlation coefficient are almost never encountered (the data does not lie exactly on a straight line), but we are interested in the degree to which a line is closest. The closer one is to the absolute value of the coefficient, the more data the line catches and the stronger the linear relationship between the variables exists. If there is no linear relation between the studied variables, the correlation coefficient is equal to zero.

### **3.4 Post-only non-equivalent comparison design**

The post-only non-equivalent comparison design is a weaker quasi-experimental design than the other one. The method is based on the comparison of post-intervention data. A major problem is that the treatment or intervention group and the controls may not have started at the same place. So, while we know where the two groups ended, we do not know where they began. Differences between the treated and non-treated may reflect differences in where they began rather than the effect of the interventions. To make groups more equivalent, it is necessary to try to match treated and control groups as closely as can be. Still, generally this may be the best design the ex-post situation allows.

#### *Exact matching with the application of post-only non-equivalent comparison design*

This method is very similar to the previous one. However, it is distinguished by the application of exact matching, which is the process of pairing individuals from treated and non-treated samples according to quantified, categorized characteristics which must be the same for both units.

### **3.5 Propensity score matching**

Propensity score matching (PSM) constructs a statistical comparison group that is based on a model of the probability of participating in the treatment, using observed characteristics. Participants are then matched on the basis of their propensity score to non-participants. The average treatment effect of the program is then calculated as the mean difference in outcomes across these two groups.

Different approaches are used to match participants and non-participants on the basis of the propensity score. We used two methods: nearest-neighbour (NN) matching and exact matching based on propensity score.

#### *Propensity score exact matching*

Exact matching based on propensity score was made using a propensity score rounded up to 4 digits. This choice of digits proved to be the most optimal because by its use we have kept the largest number of units, both treated and non-treated. The participants and non-participants with the same propensity score were matched together. Then, the non-participants assume the impact period from matched participants.

### *Propensity score nearest neighbour matching*

Nearest-neighbour matching is one of the most frequently used matching techniques. Here, each treated unit is matched to the comparison with a non-treated unit (or more units) with the closest propensity score. We did matching with the 5 nearest neighbours. Matching can be done with or without replacement; we used matching without replacement. That means that the same non-participant can be used as a match to participants only once.

### **3.6 Cost-benefit analysis**

Cost effectiveness analysis involves comparing the costs of the intervention to its effects that can be achieved from counter impact evaluation approaches. The purpose of cost-benefit analysis is to determine whether the monetised benefits of a programme exceed its net costs.

The other expression of the cost-benefit analysis says that it is a kind of financial statement summing items with a positive and negative influence on public finance.

In the performed cost-benefit analysis, real benefits and costs, as well as costs for lost opportunities and benefits from savings, were taken into account. Cost-benefit analysis work was carried out with the following items:

- 1) **Unemployment allowance** defined by Act No. 461/2003 Coll. on social security is, on the one hand, the cost of the state's passive labour market policy which is paid to the registered jobseeker if the jobseeker is eligible<sup>3</sup>. If the jobseeker is employed and unemployment allowance is not charged, the value of the last paid allowance is a positive effect, because we can generally assume that, due to intervention, the public budget saved the sum of the unpaid unemployment allowance during the period the jobseekers were employed.
- 2) Paid and saved **benefit in material need** is defined by Act No. 599/2003 Coll. on assistance in material need. That item represents the very same philosophy as in the previous unemployment allowance. The positive effect is a saved non-paid benefit while the jobseeker is employed and he/she is not eligible to apply for benefits in material need. Paid benefits of material need are a negative effect on the public budget.
- 3) A **Grant** paid to the treated jobseekers according to the actual rules of the intervention by Act No. 5/2004 Coll. on Employment Services. That item appears in the cost-benefit analysis just as the negative effect on public finance.
- 4) Received and saved payments of **health insurance** according to the updating of Act No. 580/2004 Coll. on health insurance in the two years impact period. In the case a jobseeker is unemployed, health insurance is paid by public finance and it is a cost, i.e. a negative effect. A positive effect is if individuals are employed and pay insurance to the public health service.
- 5) **Social insurance** paid according to actual versions of Act No. 461/2003 Coll.

---

<sup>3</sup> § 104 of Act No. 461/2003 Coll. on social security states: The insured person is entitled to unemployment benefit if, in the four years before registering as an unemployed jobseeker (hereinafter referred to as "registered unemployed") they were covered by unemployment insurance for at least three years.

on social security. That item measures how much money flows into the social service. In the analysis were considered values paid according to average tax assessment based on Social Insurance Agency evidence. We took into account the sum paid by the jobseeker as well as the sum paid by the employer for the employee.

- 6) **Value added taxes** defined by Act No. 222/2004 Coll. on value added tax and amendments and supplements of various acts. We assume that if somebody has a limited and below average income then it is possible that almost all is spent as the consumption of the family. That money comes back to the national budget in the way of paid value added tax. The positive effect is the total of paid value added tax; the negative effect is tax that would be paid if the jobseeker were employed (the difference between average tax assessment base and total of unemployment allowance and benefit in material need).
- 7) Paid/lost **taxes from income** according to Act No. 595/2003 Coll. on income tax. That item describes the amount of money which flows into the public budget if the jobseeker is employed and the negative effect is the lost amount of money which would be paid if the jobseeker were employed.

### 3.7 IT tools applied

All the statistical methods and computation were carried out by:

- IBM SPSS Statistics 22
- IBM SPSS Modeller
- Fusion tables by Google.com
- MS Excel
- MS Access

## 4 Traineeship

Allowance for traineeship is the intervention stated in § 51 of Act No. 5/2004 Coll. This Active labour market policy measure is distributed through regional public employment offices. The intervention was introduced for the first time on 14th April 2004.

### 4.1 Treatment effects of traineeship

The Explanatory Report on Act No. 5/2004 Coll. states that the primary purpose of traineeship is to create the conditions for obtaining the relevant professional skills and practical experience which will be valuable and attractive for an employer or any potential employer on the open labour market to ensure a higher rate of employability for unemployed graduates. The intervention was designed according to the assumption that lower practical experience is a significant barrier for the smooth entrance of graduates to the open labour market.

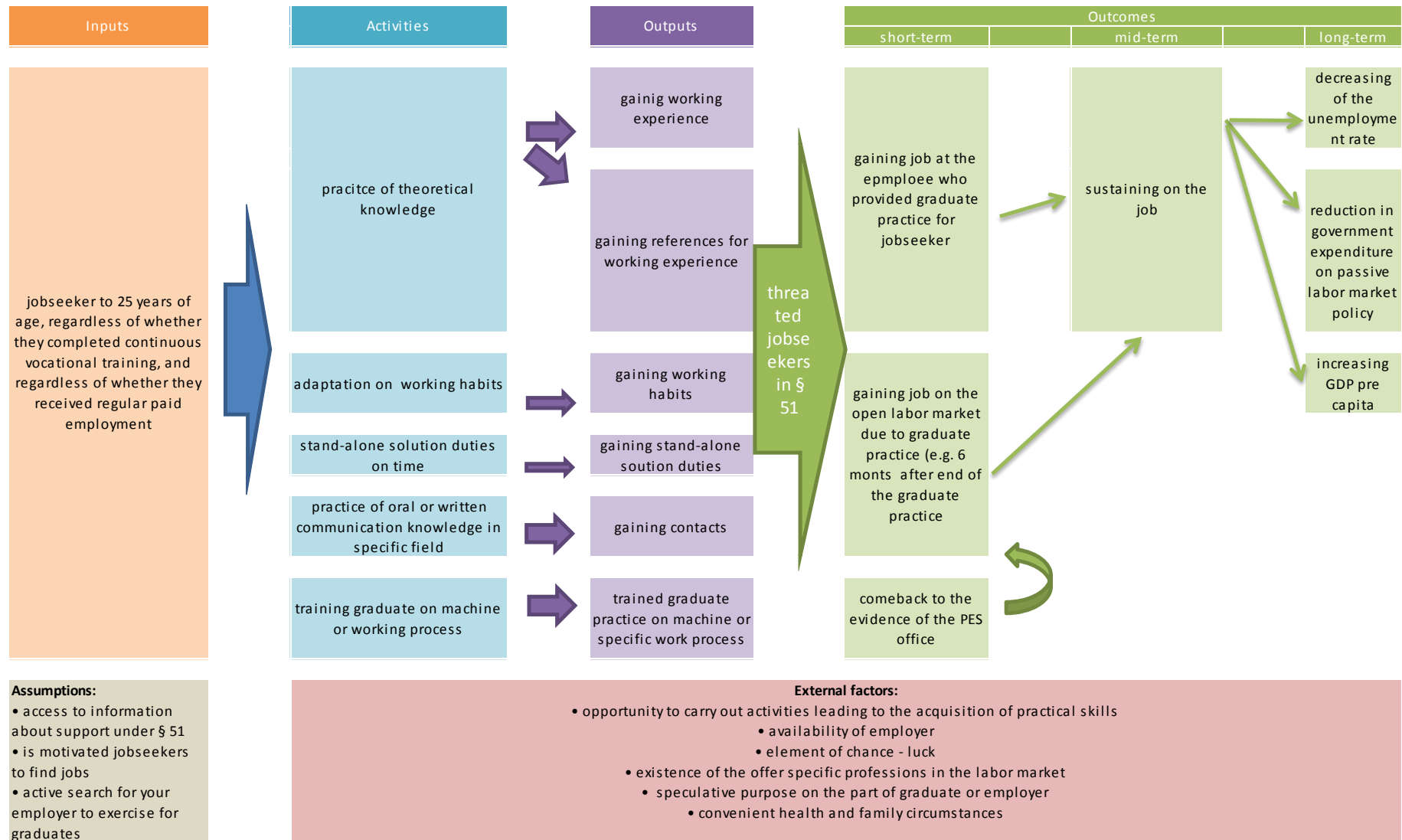
As the scheme shows, the intervention has a number of potential effects. This research will estimate the effects which occurred in the treated target groups due to the intervention in separate reference periods. The report will be focused on their employability and wages earned in the impact period, two years after the end of intervention. Every jobseeker included in the treated or non-treated samples has 24 months of impact period starting from the individual date of the end of intervention<sup>4</sup>. Controls will admit an individual impact period according to treated pairs.

---

<sup>4</sup> This rule is used in CIE methods of exact matching and propensity score matching.



The schemes below the text present the intervention log of traineeship.



Source: authors

## 4.2 Reference periods

As was described in the previous monitoring report, Act No. 5/2004 Coll. on Employment Services and on the amendment and supplement of various acts, traineeship was revised four times between the years 2007 and 2012, which are the evaluated years of the implementation of traineeship. Therefore, our treated and non-treated jobseekers must be divided into a reference period according to changes in intervention conditions, and criteria of eligibility.

Reference period			
1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012
16 months	32 months	6 months	10 months
<b>Criteria for eligibility of jobseekers according Act No. 5/2004 Coll.:</b>			
adequacy of education			
any registered jobseeker until 25 years of age (<=25 years of age)			until 26 years of age (<=26)
jobseeker must fill in the application form			
<b>Terms of the intervention:</b>			
support period up to 6 months		at least 3 months and not more than 6 months	
Eligibility for multiple support: 1 year after the end of previous graduate practice.		no multiple support	
<b>Financial contribution:</b>			
financial support 56,43 Eur for participant per month		living wage	

Source: Act No. 5/2004 Coll., § 51

The allowance for traineeship was distributed in the growth tendency according to the time of increasing unemployment rate in Slovakia. In the first 16 months of the reference period, less than 700 jobseekers per month on average were supported. In the last period, based on the years of 2011 and start of 2012, it was up to 3000 jobseekers per month. This is an increase of more than double in comparison to the first reference period. In total, more than 90 thousand jobseekers from all parts of Slovakia were supported, and more than 1,400 jobseekers per month during the 64 months of the evaluation period of traineeship were treated.

	Reference period				total
	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	
	16 months	32 months	6 months	10 months	
No. of treated jobseekers	10 807	37 954	18 042	24 584	91 387
Average per month	675	1 186	3 007	2 458	1 428

## 4.3 Target groups

Due to changes in the Act on Employment Services – target groups of traineeship were changed over the period. To keep the evaluated intervention homogeneous, it was necessary to identify jobseekers' criteria to be eligible for the intervention. Even when we divided the evaluated period of the traineeship implementation into four periods, it was possible to identify just one significant change of the target group in 2011. That is the reason why we identify two types of target groups which will be of concern in the process of control group design.

- **From 1st January 2007 till 30th June 2011 (54 months)**
  - The Act on Employment Services stated that an eligible person for traineeship was: any registered jobseeker who was 25 years of age and less,

- a jobseeker who had adequate education related to the traineeship he/she wanted to attend and
- **a jobseeker who submits an application for traineeship**
- **From 1st July 2011 till 30th April 2012 (10 months)**
- Eligible for traineeship was every registered jobseeker who fit these conditions:
  - 26 years of age and less,
  - a jobseeker who had adequate education related to the traineeship he/she wanted to attend and
- **a jobseeker who submitted an application for traineeship.**

To summarize these facts, the target group of traineeship consists of every jobseeker that was registered in the database of the Public Employment office, jobseekers to 25/26 years of age, regardless of whether they ended up continuing vocational training, and regardless of whether they received regular paid employment or not.

#### 4.4 Test of representativeness of samples

The samples of the treated and non-treated individuals were created on the basis of the rules of the law and also on the logical time sequence of individual registrations of jobseekers. During the process of creating the samples, some individuals were excluded because they did not have recorded all the values of all relevant variables. We set the rules concerning which variables must be recorded for every individual to be included in the sample. It was necessary to reduce the sample because of missing data records. However, in order to verify that the generated samples sustained were still representative, we compared in detail the distribution of variables for individuals which are included in the final sample with those who were excluded for reason of missing data in some of the variables recorded. For this purpose, a non-parametric alternative to the Chi-squared test was used, which is represented by the Kolmogorov-Smirnov test. With the aforementioned test we compared the probability of distributions distinguishing the two samples. We have preceded this test to compare the distributions probability of several variables in the sample of treated individuals and in the sample of non-treated individuals.

##### 4.4.1 Treated group excluded from the sample

In this part of the evaluation we tested the probability distributions of frequencies for treated individuals included in the sample and excluded from the sample. We verified the equality of frequency distributions in the final sample of treated individuals and the dropped ones. We used the Kolmogorov-Smirnov test as the non-parametric alternative to the Chi-square test.

The results of the testing are in the following table:

Hypothesis Test Summary				
Treated P51				
Variable	Null Hypothesis	Test	Sig.	Decision
Gender	The distribution of values is the same across the categories of selected / non selected	Independent-Samples Kolmogorov-Smirnov Test	0.518	The null hypothesis was confirmed
Marital status			1.000	The null hypothesis was confirmed
Level of education (10 categories)			0.759	The null hypothesis was confirmed
Level of education (5 categories)			0.893	The null hypothesis was confirmed
Disadvantages			0.964	The null hypothesis was confirmed
Evidence before 2007 (in months)			0.699	The null hypothesis was confirmed
Following registration in SIA			0.964	The null hypothesis was confirmed
Driving licence (16 categories)			0.211	The null hypothesis was confirmed
Age			The distribution of Age is the same across categories of selected / non selected	Independent-Samples Mann-Whitney U Test Independent-Samples Kruskal-Wallis Test
Unemployed in months	The distribution of Unemployed in months is the same across categories of selected / non selected	Independent-Samples Mann-Whitney U Test Independent-Samples Kruskal-Wallis Test	0.188	The null hypothesis was confirmed

The null hypothesis is that both groups were sampled from populations with identical distributions. That means, for example, in the case of the variable *Marital status*, that the sample of treated individuals included in the sample came from the same distribution of various levels of *Marital status* as treated individuals excluded from the sample, so that they have the same distribution. The null hypothesis is confirmed in case that the p-value of the test is greater than the significance level used for testing. We used the significance level of 0.05 in all tests. So, for the variables where the p-value of the test is greater than 0.05, we confirmed the null hypothesis.

As we can see in the table above, the distribution of **all variables** listed in the table **is the same for the final sample of selected treated jobseekers and for the sample of dropped ones because of some missing value of some variable**. That means, the sample still remains representative for the whole population of treated jobseekers.

#### 4.4.2 Distributions of frequencies of treated individuals included and excluded from the sample

In the tables below, the distribution of frequencies of the sample of included treated individuals and those excluded is written.

Gender Crosstabulation				
		group		Total
		non selected	select ed	
Gender	men	5385	13566	18951
	women	9963	24463	34426
	unknown	7	0	7
Total		15355	38029	53384

Marital status Crosstabulation				
		group		Total
		non selected	select ed	
Marital status	unknown	0	0	0
	registered partners	3	5	8
	divorced	59	116	175
	single	14132	35242	49374
	widow	0	3	3
	married	1161	2663	3824
Total		15355	38029	53384

Unemployed before 2007 in months Crosstabulation				
		group		Total
		non selected	select ed	
Unemployed before 2007 in	< 1 year	1137	2324	3461
	1 - 3 years	538	1258	1796
	> 3 years	294	360	654
	no evidence	13386	34087	47473
Total		15355	38029	53384

Following registration in SIA Crosstabulation				
		group		Total
		non selected	select ed	
Following registration in SIA	no registration	1139	6458	7597
	following registration	14216	31571	45787
Total		15355	38029	53384

Level of education_10				
		group		Total
		non selected	select ed	
Level of education_10 categories	Not finished education	43	1	44
	Primary education	125	264	389
	Lower secondary professional ed	36	105	141
	Secondary vocational education	1058	3568	4626
	Full secondary vocational educat	6924	22188	29112
	Full secondary comprehensive ed	1451	3763	5214
	Upper vocational education	34	65	99
	Bachelor	815	773	1588
	Master	2733	7280	10013
Doctoral	19	22	41	
Total		13238	38029	51267

Driving licence_16 categories Crosstabulation				
		group		Total
		non selected	select ed	
Driving licence_16 categories	Driving license: group DE	1	8	9
	Driving license: group D	25	75	100
	Driving license: group D1E	1	8	9
	Driving license: group D1	25	75	100
	Driving license: group CE	87	256	343
	Driving license: group C	180	521	701
	Driving license: group C1E	87	256	343
	Driving license: group C1	180	521	701
	Driving license: group BE	87	256	343
	Driving license: group B	4604	10395	14999
	Driving license: group B1	4604	10394	14998
	Driving license: group A	1067	2423	3490
	Driving license: group A2	0	0	0
	Driving license: group A1	1067	2423	3490
	Driving license: group AM	4655	10508	15163
Driving license: group T	234	676	910	
Total		16904	38795	55699

Types of disadvantages Crosstabulation				
		Count		
		non selected	select ed	Total
Types of disadvantages	no disadvantage	10876	26011	36887
	graduate	2543	6626	9169
	long - term unemployed	1906	5350	7256
	low education level	1	0	1
	organizational	3	6	9
	poor working discipline	3	1	4
	care	10	17	27
	disabled	13	18	31
Total		15355	38029	53384

#### 4.4.3 Non-treated group excluded from the sample

In this part of the evaluation, the probability distributions of frequencies for non-treated individuals included and excluded from the sample was tested. As is shown in the table below, the distribution of the frequencies of the variables is the same in the sample of included non-treated individuals and the sample of excluded individuals. That means, through the exclusion of the individuals with some missing value of some variable, the final sample remains representative for the whole population of non-treated jobseekers.

<b>Hypothesis Test Summary</b>				
<b>Non treated P51</b>				
<b>Variable</b>	<b>Null Hypothesis</b>	<b>Test</b>	<b>Sig.</b>	<b>Decision</b>
<b>Gender</b>	The distribution of values is the same across categories of selected / non selected	Independent-Samples Kolmogorov-Smirnov Test	0.996	The null hypothesis was confirmed
<b>Marital status</b>			0.441	The null hypothesis was confirmed
<b>Level of education (10 categories)</b>			0.699	The null hypothesis was confirmed
<b>Level of education (5 categories)</b>			0.441	The null hypothesis was confirmed
<b>Disadvantages</b>			0.699	The null hypothesis was confirmed
<b>Evidence before 2007 (in months)</b>			1.000	The null hypothesis was confirmed
<b>Following registration in SIA</b>			0.964	The null hypothesis was confirmed
<b>Driving licence (16 categories)</b>			0.941	The null hypothesis was confirmed
<b>Age</b>			0.979	The null hypothesis was confirmed
<b>Last occupation</b>			0.269	The null hypothesis was confirmed

#### 4.4.4 Distributions of frequencies of non-treated individuals included and excluded from the sample

In the following tables the frequencies of the variables in the sample on non-treated individuals included in the sample and the excluded ones are written:

Gender Crosstabulation				
		group		Total
		selected	non selected	
Gender	men	35987	50145	86132
	women	26094	35777	61871
	unknown	0	42	42
Total		62081	85964	148045

Marital status Crosstabulation				
		group		Total
		selected	non selected	
Marital status	unknown	0	475	475
	registered partners	9	34	43
	divorced	136	175	311
	single	58576	79172	137748
	widow	7	10	17
	married	3353	6098	9451
Total		62081	85964	148045

Unemployed before 2007 in months Crosstabulation				
		group		Total
		selected	non selected	
Unemployed before 2007 in	< 1 year	1095	830	1925
	1 - 3 years	480	539	1019
	> 3 years	65	136	201
	no evidence	60441	57719	118160
Total		62081	59224	121305

Following registration in SIA Crosstabulation				
		group		Total
		selected	non selected	
Following registration in SIA	no registration	62081	30365	92446
	following registration	0	55599	55599
Total		62081	85964	148045

Level of education_10				
		group		Total
		selected	non selected	
Level of education_10 categories	Not finished education	26	1908	1934
	Primary education	3391	10118	13509
	Lower secondary professional education	420	457	877
	Secondary vocational education	16198	11818	28016
	Full secondary vocational education	36546	26168	62714
	Full secondary comprehensive education	3688	6259	9947
	Upper vocational education	106	98	204
	Bachelor	660	1105	1765
	Master	1046	1479	2525
Doctoral	0	0	0	
Total		62081	59410	121491

Driving licence_16 categories Crosstabulation				
		group		Total
		selected	non selected	
Driving licence_16 categories	Driving license: group DE	21	2	23
	Driving license: group D	134	133	267
	Driving license: group D1E	21	3	24
	Driving license: group D1	134	133	267
	Driving license: group CE	553	451	1004
	Driving license: group C	1137	1008	2145
	Driving license: group C1E	553	451	1004
	Driving license: group C1	1137	1008	2145
	Driving license: group BE	553	451	1004
	Driving license: group B	16050	16581	32631
	Driving license: group B1	16050	16581	32631
	Driving license: group A	3832	4419	8251
	Driving license: group A2	0	1	1
	Driving license: group A1	3832	4419	8251
	Driving license: group AM	16202	16803	33005
Driving license: group T	1424	1296	2720	
Total		61633	63740	125373

Types of disadvantages Crosstabulation				
		group		Total
		selected	non selected	
Types of disadvantages	no disadvantage	56698	68231	124929
	graduate	2682	5370	8052
	long - term unemployed	2640	12168	14808
	low education level	9	25	34
	organizational	9	10	19
	poor working discipline	6	64	70
	care	26	70	96
	age over 50 years	0	7	7
disabled	11	19	30	
Total		62081	85964	148045

## 4.5 Description of samples

This chapter describes some facts about the sample at the time before the creation of the pairs. This is another milestone on the path to gain matched individuals of treated and control groups in four follow-up reference periods, which should ensure homogeneity of intervention and the validity of counter-factual impact evaluation.

The heatmap or intensity map presents the number of individuals that enrol in the program of traineeship. It is obvious that most of the participants in the samples are from those parts of Slovakia which are highly exposed to the unemployment rate, i.e. places where the intervention mostly makes sense and the placement of jobseekers has a much desired effect.

### 4.5.1 Permanent residence

We have covered all districts and regions of Slovakia. Just for interpretation, we will use the regional distribution of individuals. As the table below presents the most treated jobseekers in all four reference periods coming from Prešov region, which is the second most suffering from high and permanent levels of unemployment rate after Banská Bystrica region. Even though Banská Bystrica region has a higher level of unemployment rate, Prešov region is more populous, and that is the reason why, in all reference periods, most jobseekers came from Prešov region. The least treated and non-treated jobseekers are in Bratislava region; the capital region for a long-time has had the lowest level of unemployment rate. In total we have almost 65 thousand treated jobseekers covered by our samples across four reference periods and almost 67 thousand controls.

The table next to the text describes the power of relations among treated, non-treated groups of individuals and the average unemployment rate across the regions of Slovakia and reference periods.

<b>1-st reference period: 1.1.2007 - 30.4.2008</b>		
<i>Correlation</i>	<i>Treated</i>	<i>Non-treated</i>
Average unemployment rate in region (%)	0,808	0,545
<b>2-nd reference period: 1.5.2008 - 31.12.2010</b>		
<i>Correlation</i>	<i>Treated</i>	<i>Non-treated</i>
Average unemployment rate in region (%)	0,849	0,410
<b>3-rd reference period: 1.1.2011 - 30.6.2011</b>		
<i>Correlation</i>	<i>Treated</i>	<i>Non-treated</i>
Average unemployment rate in region (%)	0,874	0,200
<b>4-th reference period: 1.7.2011 - 30.4.2012</b>		
<i>Correlation</i>	<i>Treated</i>	<i>Non-treated</i>
Average unemployment rate in region (%)	0,854	0,308

As is presented in the table, the relation between the unemployment rate in the specific region and number of treated jobseekers is much more related than the number of non-treated jobseekers in the regions.

Maximal differences between treated and non-treated groups in the reference periods are 8.8 %. Through those differences it is obvious that in regions with a higher level of unemployment rate there are higher shares of treated jobseekers than the total treated jobseekers in our samples. The aforementioned indicates a higher probability of being treated in a group of unemployed eligible individuals in regions with a higher level of unemployment rate than in regions with lower unemployment rates in the west of Slovakia.



1-st reference period: 1.1.2007 - 30.4.2008						
Region	Region of permanent residence_treated		Region of permanent residence_non-treated		Difference between groups (%)	Average unemployment rate in region (%)
	Frequency	Percent	Frequency	Percent		
Bratislava region	134	2,1	217	5,3	3,2	2,2
Tnava region	607	9,6	522	12,8	3,2	4,4
Trencin region	613	9,7	495	12,2	2,5	4,7
Nitra region	1091	17,3	613	15,1	-2,2	7,4
Zilina region	703	11,1	548	13,5	2,3	6,7
Banska Bystrica region	935	14,8	508	12,5	-2,3	15,1
Presov region	1145	18,2	639	15,7	-2,4	13,3
Kosice region	1080	17,1	525	12,9	-4,2	12,8
Total	6308	100,0	4067	100,0		9,2

2-nd reference period: 1.5.2008 - 31.12.2010						
Region	Region of permanent residence_treated		Region of permanent residence_non-treated		Difference between groups (%)	Average unemployment rate in residence region (%)
	Frequency	Percent	Frequency	Percent		
Bratislava region	350	2,2	2873	7,9	5,7	3,9
Tnava region	2713	11,2	4450	12,2	1,0	7,2
Trencin region	2275	9,4	4962	13,6	4,2	8,2
Nitra region	3072	12,6	5243	14,3	1,7	10,4
Zilina region	3446	14,2	4974	13,6	-0,6	10,3
Banska Bystrica region	3451	14,2	4004	11,0	-3,2	18,4
Presov region	4756	19,6	5322	14,6	-5,0	17,2
Kosice region	4063	16,7	4737	13,0	-3,8	15,4
Total	24126	100,0	36565	100,0	-	12,3

3-rd reference period: 1.1.2011 - 30.6.2011						
Region	Region of permanent residence_treated		Region of permanent residence_non-treated		Difference between groups (%)	Average unemployment rate in residence region (%)
	Frequency	Percent	Frequency	Percent		
Bratislava region	357	2,2	583	8,1	5,9	5,5
Tnava region	1556	9,6	943	13,1	3,5	8,9
Trencin region	1434	8,8	1013	14,1	5,3	9,6
Nitra region	1887	11,6	1023	14,2	2,6	13,2
Zilina region	2333	14,4	999	13,9	-0,5	12,8
Banska Bystrica region	2240	13,8	744	10,4	-3,4	20,7
Presov region	3501	21,6	1020	14,2	-7,4	19,6
Kosice region	2922	18,0	861	12,0	-6,0	18,2
Total	16230	100,0	7186	100,0	-	14,6

4-th reference period: 1.7.2011 - 30.4.2012						
Region	Region of permanent residence_treated		Region of permanent residence_non-treated		Difference between groups (%)	Average unemployment rate in residence region (%)
	Frequency	Percent	Frequency	Percent		
Bratislava region	415	2,3	1498	7,9	5,6	5,1
Tnava region	1641	9,1	2346	12,3	3,3	8,6
Trencin region	1606	8,9	2573	13,5	4,6	9,5
Nitra region	2282	12,6	2704	14,2	1,6	12,5
Zilina region	2670	14,8	2823	14,8	0,1	12,3
Banska Bystrica region	2449	13,5	1998	10,5	-3,0	20,2
Presov region	4051	22,4	2596	13,6	-8,8	19,1
Kosice region	2978	16,5	2499	13,1	-3,3	17,2
Total	18092	100,0	19037	100,0	-	14,0

## 4.5.2 Gender

These frequency tables show the share of men and women in our samples. As the numbers describe, the average percentage of treated unemployed men is at the level of more than 31%. On the other hand, 2 individuals from the treated unemployed graduates are women in different reference periods, even though the share of women in non-treated groups is almost the reverse. That is why it is possible to deduce that girls have a greater interest to undergo traineeship than women. We verified this difference by a statistical test.

The difference between treated and non-treated groups is at the level of 20 to 25 %.

The differences between the categories of gender were verified using non-parametric tests for testing the equality of the distributions of two samples. The results of the tests are in the following table.

If we compare the p-value of the test with a significance level of 0.05; we could say that the null hypothesis is rejected. The differences between the percentage of men and women between treated individuals are significant. The differences between non-treated individuals in the field of gender are not significant. This result is illustrated in the following table.

This test verified that between treated individuals more women are participating in this program and this difference is statistically significant.

1-st reference period: 1.1.2007 - 30.4.2008					
Gender	Gender_treated		Gender_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
men	1974	31,3	2323	57,1	25,8
women	4334	68,7	1744	42,9	-25,8
Total	6308	100,0	4067	100,0	-

2-nd reference period: 1.5.2008 - 31.12.2010					
Gender	Gender_treated		Gender_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
men	8870	36,5	20808	56,9	20,4
women	15434	63,5	15757	43,1	-20,4
Total	24304	100,0	36565	100,0	-

3-rd reference period: 1.1.2011 - 30.6.2011					
Gender	Gender_treated		Gender_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
men	6027	37,1	4336	60,3	23,2
women	10203	62,9	2850	39,7	-23,2
Total	16230	100,0	7186	100,0	-

4-th reference period: 1.7.2011 - 30.4.2012					
Gender	Gender_treated		Gender_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
men	6588	36,4	11325	59,5	23,1
women	11504	63,6	7712	40,5	-23,1
Total	18092	100,0	19037	100,0	-

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count_gender_non_treated is the same across categories of levels_of_variable_gender_non_treated.	Independent-Samples Mann-Whitney U Test	,343	Retain the null hypothesis.
2	The distribution of count_gender_non_treated is the same across categories of levels_of_variable_gender_non_treated.	Independent-Samples Kolmogorov-Smirnov Test	,211	Retain the null hypothesis.
3	The distribution of count_gender_non_treated is the same across categories of levels_of_variable_gender_non_treated.	Independent-Samples Kruskal-Wallis Test	,248	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count_gender_non_treated is the same across categories of levels_of_variable_gender_non_treated.	Independent-Samples Mann-Whitney U Test	,343	Retain the null hypothesis.
2	The distribution of count_gender_non_treated is the same across categories of levels_of_variable_gender_non_treated.	Independent-Samples Kolmogorov-Smirnov Test	,211	Retain the null hypothesis.
3	The distribution of count_gender_non_treated is the same across categories of levels_of_variable_gender_non_treated.	Independent-Samples Kruskal-Wallis Test	,248	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

### 4.5.3 Marital status

Our database distinguished five types of marital status of the registered jobseekers. The most frequent type of registered jobseekers are single; in different reference periods these make up about 90 % of the eligible sample of treated and non-treated groups. A high share of single individuals in our distribution was expected based on the fact that we evaluate traineeship, i.e. jobseekers a short time after they have graduated from schools. Nevertheless, our samples distributions show 3.8 by more than 11 % of married jobseekers. There is only an insignificant share of widows, divorced individuals or registered partners. Differences between treated and non-treated groups are not greater than 1.8 % of the specific category of marital status across the reference periods. The differences between the percentage of the category *single* and of the category *married* were also verified by non-parametric tests. The results are written below. In both cases, the significance of the differences was not confirmed.

1-st reference period: 1.1.2007 - 30.4.2008					
Type of marital status	Marital status_treated		Marital status_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
registered partners	3	,0	0	,0	,0
divorced	59	0,9	17	0,4	-0,5
single	5546	87,9	3650	89,7	1,8
widow	1	0,0	2	0,0	0,0
married	699	11,1	398	9,8	-1,3
Total	6308	100,0	4067	100,0	

2-nd reference period: 1.5.2008 - 31.12.2010					
Type of marital status	Marital status_treated		Marital status_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
registered partners	3	,0	9	,0	,0
divorced	66	0,3	100	0,3	0,0
single	22527	92,7	34320	93,9	1,2
widow	2	0,0	4	,0	0,0
married	1706	7,0	2132	5,8	-1,2
Total	24304	100,0	36565	100,0	-

3-rd reference period: 1.1.2010 - 30.6.2011					
Type of marital status	Marital status_treated		Marital status_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
registered partners	1	,0	0	,0	,0
divorced	27	0,2	8	0,1	-0,1
single	15243	93,9	6815	94,8	0,9
widow	2	0,0	1	,0	0,0
married	957	5,9	362	5,0	-0,9
Total	24304	100,0	36565	100,0	-

4-th reference period: 1.7.2011 - 30.4.2012					
Type of marital status	Marital status_treated		Marital status_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
registered partners	1	,0	1	,0	,0
divorced	29	0,2	20	0,1	-0,1
single	17082	94,4	18283	96,0	1,6
widow	1	0,0	1	,0	0,0
married	979	5,4	732	3,8	-1,6
Total	18092	100,0	19037	100,0	-

#### Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count_single is the same across categories of treated_single.	Independent-Samples Mann-Whitney U Test	,114 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count_single is the same across categories of treated_single.	Independent-Samples Kolmogorov-Smirnov Test	,699	Retain the null hypothesis.
3	The distribution of count_single is the same across categories of treated_single.	Independent-Samples Kruskal-Wallis Test	,110	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

#### Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count_married is the same across categories of treated_married.	Independent-Samples Mann-Whitney U Test	,343 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count_married is the same across categories of treated_married.	Independent-Samples Kolmogorov-Smirnov Test	,699	Retain the null hypothesis.
3	The distribution of count_married is the same across categories of treated_married.	Independent-Samples Kruskal-Wallis Test	,248	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

#### 4.5.4 Types of disadvantages

We recognize 9 types of disadvantages according to the definitions of Act No. 5/2004 on Employment Services.

This variable shows that in a group of treated jobseekers, only about 20 % are graduates. We verified the eligibility of jobseekers and our samples are composed from eligible individuals currently valid in the reference period.

Distribution in all reference periods indicates the insignificant character of the variable because it only covers on average of less than 20 % of all jobseekers; the rest of the treated and non-treated jobseekers are without any feature of disadvantages.

Never the less, the second biggest category in the presented distributions are graduates and long-term unemployed jobseekers. The other types of disadvantages rarely appeared in our distribution of reference periods.

#### 4.5.5 Age

The average age of treated jobseekers is in the range from 20 to 21 years, while the average age of controls is in the range from 22 to 24 years. The average age of non-treated jobseekers is higher in all reference periods. The youngest eligible treated jobseekers are 16 years old in all reference periods and the youngest non-treated jobseekers are 17 years old.

1-st reference period: 1.1.2007 - 30.4.2008					
Type of disadvantages	Disadvantages_treated		Disadvantages_treated_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
no	5289	83,8	3760	92,5	8,6
graduate	733	11,6	145	3,6	-8,1
long-term unemployed	279	4,4	158	3,9	-0,5
low education level	0	0,0	2	,0	0,0
organizational	2	0,0	1	,0	0,0
unemployed	0	0,0	0	,0	0,0
care	3	0,0	1	,0	0,0
age more than 50	0	0,0	0	,0	0,0
disable	2	0,0	1	,0	0,0
Total	6308	100,0	4067	100,0	-

2-nd reference period: 1.5.2008 - 31.12.2010					
Type of disadvantages	Disadvantages_treated		Disadvantages_treated_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
none	18107	74,5	33685	92,1	17,6
graduate	4390	18,1	1487	4,1	-14,0
unemployed	1785	7,3	1	,0	-7,3
long-term unemployed	0	,0	1349	3,7	3,7
not finished	0	,0	1	,0	0,0
low education	0	,0	6	,0	0,0
organizational	5	0,0	8	,0	0,0
care	12	0,0	19	,1	0,0
problematic situation	0	,0	3	,0	0,0
disabled	5	0,0	6	,0	0,0
Total	24304	100,0	36565	100,0	-

3-rd reference period: 1.1.2011 - 30.6.2011					
Type of disadvantages	Disadvantages_treated		Disadvantages_treated_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
none	10551	65,0	6682	93,0	28,0
graduate	3611	22,2	240	3,3	-18,9
long-term unemployed	2052	12,6	258	3,6	3,6
low education	0	,0	1	,0	0,0
organizational	1	,0	8	,0	0,0
care	6	,0	1	,0	0,0
disabled	9	,1	1	,0	0,0
Total	18092	100,0	0	100,0	-

4-th reference period: 1.7.2011 - 30.4.2012					
Type of disadvantages	Disadvantages_treated		Disadvantages_treated_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
none	9610	53,1	16977	89,2	36,1
graduate	3898	21,5	988	5,2	-16,4
long-term unemployed	4560	25,2	1063	5,6	-19,6
not finished	0	,0	0	,0	0,0
low education	0	,0	1	,0	0,0
organizational	1	0,0	1	,0	0,0
care	5	0,0	6	,0	0,0
disabled	12	0,1	1	,0	-0,1
Total	18092	100,0	19037	100,0	-

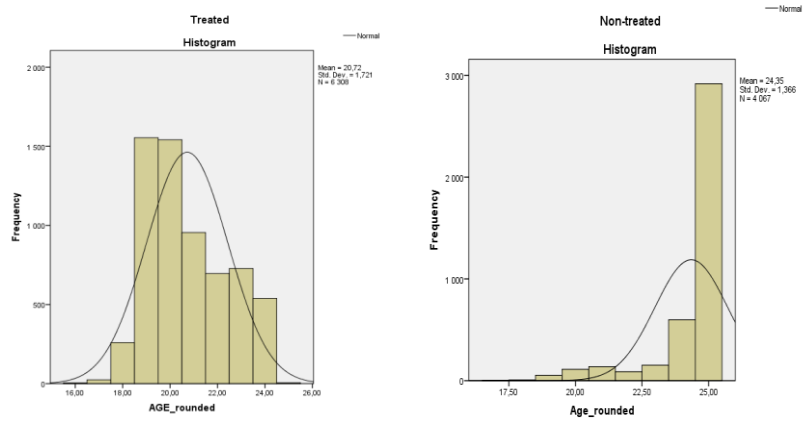
Descriptives_treated: AGE		1-st reference period: 1.1.2007 - 30.4.2008		2-nd reference period: 1.5.2008 - 31.12.2010		3-rd reference period: 1.1.2010 - 30.6.2011		4-th reference period: 1.7.2011 - 30.4.2012	
		Treated	Non-treated	Treated	Non-treated	Treated	Non-treated	Treated	Non-treated
		Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic
Mean		20,7215	24,3479	21,0387	24,0229	21,2969	23,1316	21,2186	22,8624
95% Confidence Interval for	Lower Bound	20,6791	24,3059	21,0145	24,0111	21,2661	23,1043	21,1894	22,8433
	Upper Bound	20,7640	24,3899	21,0629	24,0347	21,3277	23,1590	21,2478	22,8814
5% Trimmed Mean		20,6909	24,5641	21,0015	24,1491	21,2546	23,1880	21,1569	22,8845
Median		20,0000	25,0000	20,0000	24,0000	21,0000	23,0000	21,0000	23,0000
Variance		2,959	1,865	3,699	1,323	4,010	1,400	4,014	1,796
Std. Deviation		1,72031	1,36579	1,92338	1,15015	2,00259	1,18336	2,00340	1,34016
Minimum		16,00	17,00	16,00	17,00	16,00	18,00	16,00	18,00
Maximum		25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00
Range		9,00	8,00	9,00	8,00	9,00	7,00	9,00	7,00
Interquartile Range		3,00	1,00	4,00	2,00	3,00	1,00	3,00	2,00
Skewness		,437	-2,451	,450	-1,434	,324	-,678	,444	,008
Kurtosis		-,808	5,394	-1,093	2,263	-1,248	,662	-1,108	-,665

As is presented in the output table of the normality test below, any distributions of reference periods were not confirmed via a normal distribution of values. Even graphical numbers of distributions do not have symmetric histograms under a normal curve. The shape of distributions reveals that the group of treated jobseekers is created mostly by individuals between 19 and 20 years of age. On the other hand, non-treated groups in the first two reference periods are mostly 25-year old jobseekers and in the last second reference period the majority are 22 and 23-year old jobseekers.

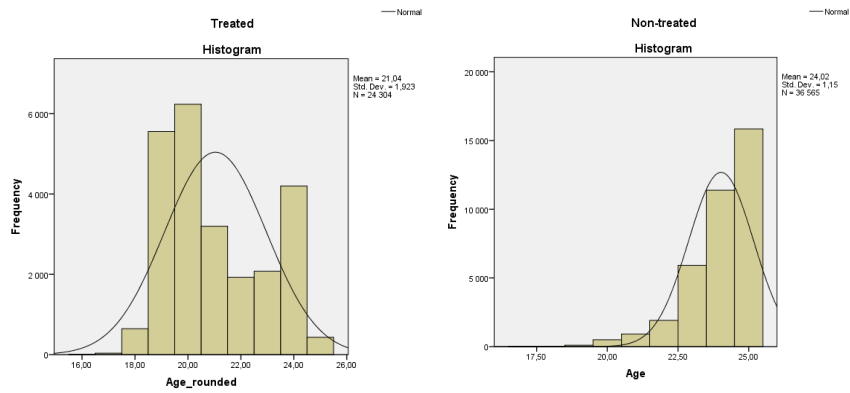
Tests of Normality: Age	Treated			Non-treated		
	Kolmogorov-Smirnov <sup>a</sup>					
	Statistic	df	Sig.	Statistic	df	Sig.
<b>1-st reference period: 1.1.2007 - 30.4.2008</b>	,199	6308	0,000	,401	4067	0,000
<b>2-nd reference period: 1.5.2008 - 31.12.2010</b>	,219	24304	0,000	,237	36565	0,000
<b>3-rd reference period: 1.1.2010 - 30.6.2011</b>	,208	16230	0,000	,216	7186	0,000
<b>4-th reference period: 1.7.2011 - 30.4.2012</b>	,216	18092	0,000	,181	19037	0,000

Boxplots present the number of outliers and extremes in the distributions of controls, which ensures a slight distortion of means to decrease. Extremes and outliers occur just in non-treated samples.

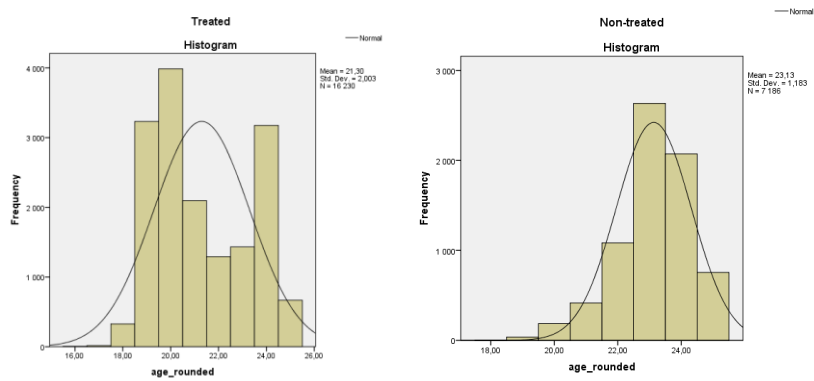
**1-st reference period: 1.1.2007 - 30.4.2008**



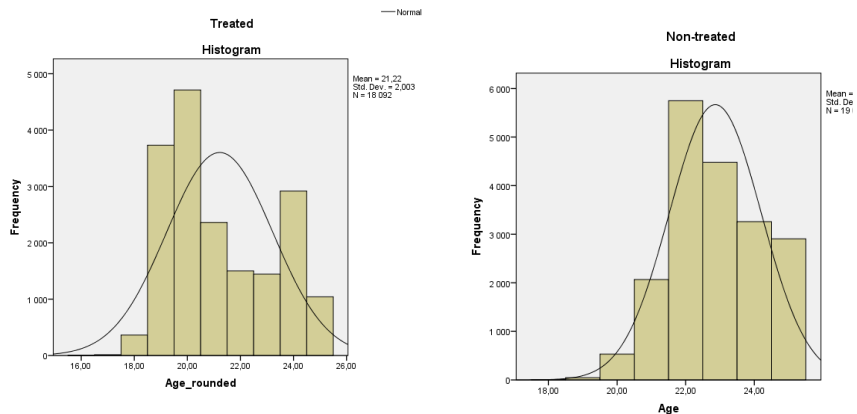
**2-nd reference period: 1.5.2008 - 31.12.2010**



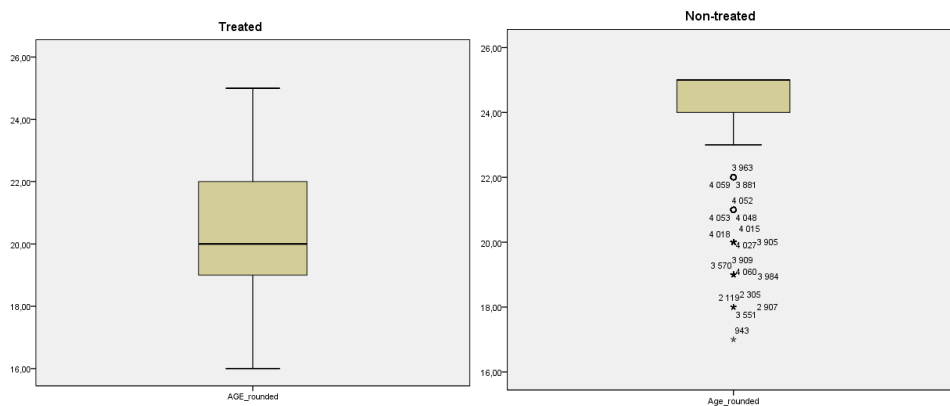
**3-rd reference period: 1.1.2011 - 30.6.2011**



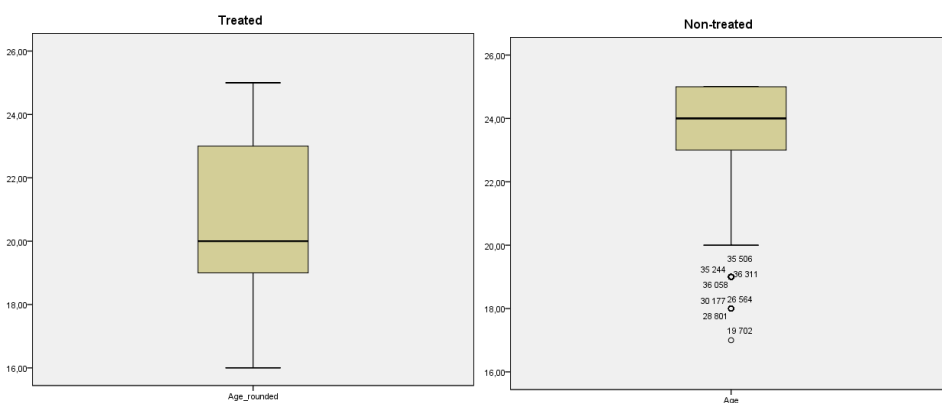
**4-th reference period: 1.7.2011 - 30.4.2012**



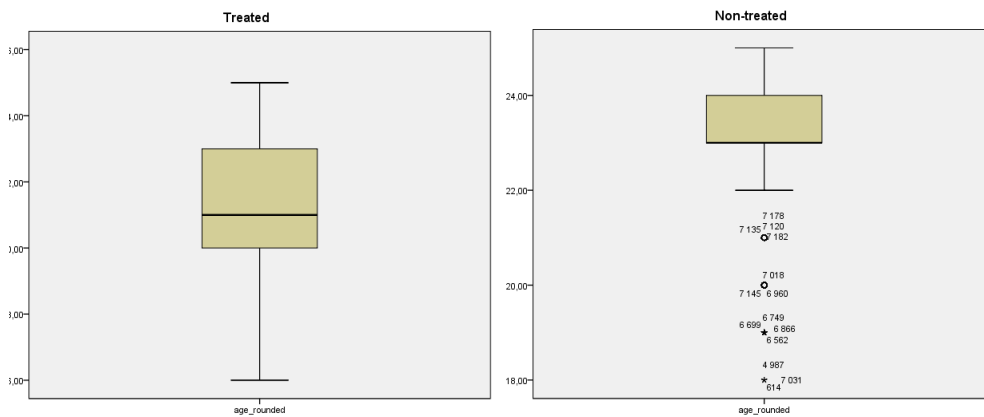
**1-st reference period: 1.1.2007 - 30.4.2008**



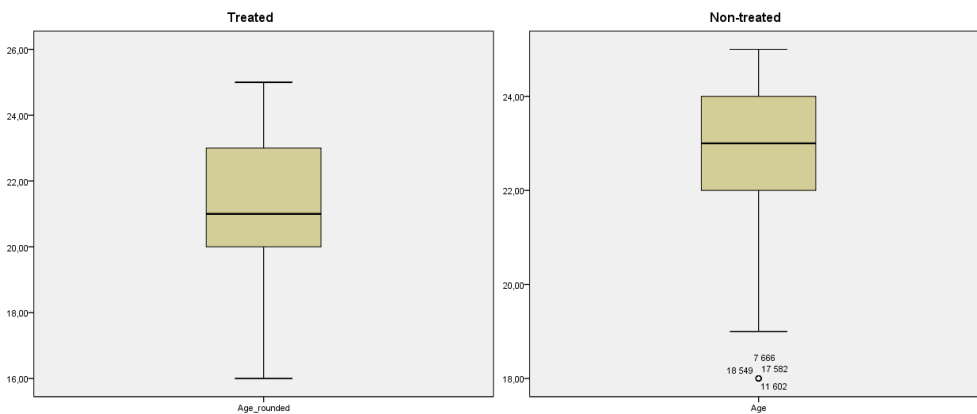
**2-nd reference period: 1.5.2008 - 31.12.2010**



**3-rd reference period: 1.1.2010 - 30.6.2011**



**4-th reference period: 1.7.2011 - 30.4.2012**



#### 4.5.6 Level of education

We distinguished 5 types of highest achieved level of education among treated and non-treated jobseekers. As is presented in the green bar charts in the tables next to the text, the highest frequency of education level is of secondary vocational school graduates, i.e. on average more than half of the sample. The second most frequent levels are college and vocational school graduates who were treated. We can identify an increasing interest from college graduates in traineeship since the second period, i.e. 1.5.2008, when the crisis started in Slovakia and unemployment started to increase.

The biggest differences between treated and controls across the periods are 22 % at the college level of education. We tested the statistical significance of the differences between the sample of treated and non-treated individuals for every level of education. The results of these tests are in the following table:

1-st reference period: 1.1.2007 - 30.4.2008					
Level of education	Treated		Non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
primary school	93	1,5	648	15,9	14,5
secondary vocational school	3475	55,1	1839	45,2	-9,9
vocational school	1269	20,1	1259	31,0	10,8
comprehensive school	519	8,2	189	4,6	-3,6
college	952	15,1	132	3,2	-11,8
Total	6308	100,0	4067	100,0	-

2-nd reference period: 1.5.2008 - 31.12.2010					
Level of education	Treated		Non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
primary school	169	0,7	1821	5,0	4,3
secondary vocational school	13213	54,4	22136	60,5	6,2
vocational school	3789	15,6	9393	25,7	10,1
comprehensive school	2392	9,8	2184	6,0	-3,9
college	4741	19,5	1031	2,8	-16,7
Total	24304	100,0	36565	100,0	-

3-rd reference period: 1.1.2010 - 30.6.2011					
Level of education	Treated		Non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
primary school	109	0,7	408	5,7	3,0
secondary vocational school	8285	51,0	4400	61,2	10,2
vocational school	2265	14,0	1728	24,0	10,1
comprehensive school	1644	10,1	501	7,0	-3,2
college	3927	24,2	149	2,1	-22,1
Total	16230	100,0	7186	100,0	-

4-th reference period: 1.7.2011 - 30.4.2012					
Level of education	Treated		Non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
primary school	110	0,6	835	4,4	3,8
secondary vocational school	9755	53,9	11983	62,9	9,0
vocational school	2555	14,1	4443	23,3	9,2
comprehensive school	2118	11,7	1279	6,7	-5,0
college	3554	19,6	497	2,6	-17,0
Total	18092	100,0	19037	100,0	-



Only in the case of secondary school education was the difference between treated and non-treated individuals insignificant. In the case of other variables, the differences are statistically significant. This we can say by using the p-value of the test, which we compare with the significance level 0.05. In case that the p-value of the test is lower than 0.05, the null hypothesis about the same distribution between treated and non-treated is rejected.

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count_primary_school is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,029 <sup>1</sup>	Reject the null hypothesis.
2	The distribution of count_primary_school is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,037	Reject the null hypothesis.
3	The distribution of count_primary_school is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,020	Reject the null hypothesis.
4	The distribution of count_secondary_school is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,343 <sup>1</sup>	Retain the null hypothesis.
5	The distribution of count_secondary_school is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,211	Retain the null hypothesis.
6	The distribution of count_secondary_school is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,248	Retain the null hypothesis.
7	The distribution of count_vocational_education is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,029 <sup>1</sup>	Reject the null hypothesis.
8	The distribution of count_vocational_education is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,037	Reject the null hypothesis.
9	The distribution of count_vocational_education is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,021	Reject the null hypothesis.
10	The distribution of count_comprehensive is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,029 <sup>1</sup>	Reject the null hypothesis.
11	The distribution of count_comprehensive is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,037	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

#### 4.5.7 Registered before 2007

There are four groups of the variable which inform us about the cumulative period of jobseekers before the first reference period as of 1.1.2007.

It is obvious that most of the treated and non-treated jobseekers were not unemployed before 2007. No more than 20 % (just in the first reference period) are long-term unemployed jobseekers. From the second reference period, long-term unemployed jobseekers are markedly reduced.

The biggest differences between the groups of treated and controls are at the level of 16 %.

1-st reference period: 1.1.2007 - 30.4.2008					
Unemployed before 2007	Treated		Non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
no	3565	56,5	3442	84,6	28,1
< 1 year	1344	21,3	411	10,1	-11,2
> 3 years	354	5,6	192	4,7	-0,9
1 - 3 years	1045	16,6	22	0,5	-16,0
Total	6308	100,0	4067	100,0	-
2-nd reference period: 1.5.2008 - 31.12.2010					
Unemployed before 2007	Treated		Non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
no	22538	92,7	35552	97,2	4,5
< 1 year	1139	4,7	648	1,8	-2,9
> 3 years	153	0,6	323	0,9	0,3
1 - 3 years	474	2,0	42	0,1	-1,8
Total	24304	100,0	36565	100,0	-
3-rd reference period: 1.1.2010 - 30.6.2011					
Unemployed before 2007	Treated		Non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
no	15783	97,2	7076	98,5	1,2
< 1 year	364	2,2	87	1,2	-1,0
> 3 years	7	0,0	21	0,3	0,2
1 - 3 years	76	0,5	2	0,0	-0,4
Total	16230	100,0	7186	100,0	-
4-th reference period: 1.7.2011 - 30.4.2012					
Unemployed before 2007	Treated		Non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
no	17703	97,8	18850	99,0	1,2
< 1 year	315	1,7	161	0,8	-0,9
> 3 years	7	0,0	20	0,1	0,1
1 - 3 years	67	0,4	6	0,0	-0,3
Total	24304	100,0	19037	100,0	-

### 4.5.8 Category of driving licence

Most of the graduates in the samples do not have any driving licence, which could be one of the reasons for their lacking attractiveness to the labour market.

Just about every one in four jobseeker has a driving licence for cars and motorcycles. And, just on average, about 3 % of the treated and non-treated jobseekers are holders of driving licences for buses or lorries, which determines transport working positions.

Between the treated and non-treated groups were identified differences at the maximum level of four present.

We verified the differences between treated and non-treated individuals during these 4 reference periods. The results of the testing are in the following table.

For both levels of variables, the differences between treated and non-treated are not significant. That means we could say that both treated and non-treated individuals come from the same distributions.

Category of driving licence	1-st reference period: 1.1.2007 - 30.4.2008				Difference between groups (%)
	Treated		Non-treated		
	Frequency	Valid percent	Frequency	Valid percent	
Cars and motorcycles	1427	23	920	23	0
Smaller trucks	118	2	94	2	0
Buses	41	1	36	1	0
Trucks	12	0	10	0	0
No driving license	3145	77	27591	75	2
2-nd reference period: 1.5.2008 - 31.12.2010					
Cars and motorcycles	6408	26	8959	25	2
Smaller trucks	415	2	817	2	-1
Buses	158	1	330	1	0
Trucks	44	0	78	0	0
No driving license	5230	73	13138	69	4
3-rd reference period: 1.1.2010 - 30.6.2011					
Cars and motorcycles	4957	31	1952	27	3
Smaller trucks	302	2	149	2	0
Buses	117	1	55	1	0
Trucks	33	0	13	0	0
No driving license	4879	77	17890	74	4
4-th reference period: 1.7.2011 - 30.4.2012					
Cars and motorcycles	5995	33	5890	31	2
Smaller trucks	358	2	498	3	-1
Buses	147	1	192	1	0
Trucks	37	0	46	0	0
No driving license	11270	69	12092	67	3

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count_cars is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,886 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count_cars is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.
3	The distribution of count_cars is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,559	Retain the null hypothesis.
4	The distribution of count_no_driving_licence is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,343 <sup>1</sup>	Retain the null hypothesis.
5	The distribution of count_no_driving_licence is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,699	Retain the null hypothesis.
6	The distribution of count_no_driving_licence is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,306	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

<sup>1</sup>Exact significance is displayed for this test.

### 4.6 Analysis of variance

In the samples of treated and non-treated jobseekers, we verified the equality of means or probability distributions of variables. There we tested the significance of differences between the variable means or between the variable frequency distributions. For the purpose of the testing we used a one-way analysis of variance. Firstly we verified whether the distribution of variable frequencies was normal or not. This was made using the Shapiro-Wilk test of normal distribution. Then, within the second step, we used the independent samples t-test in case of normal

distribution or non-parametric alternative Mann-Whitney U test. We also used the Kruskal-Wallis test and Kolmogorov-Smirnov test as non-parametric alternatives to one-way analysis of variance for two samples.

#### 4.6.1 1<sup>st</sup> reference period

In the following table, the results of verifying normal distribution of variables in the sample of treated and non-treated jobseekers in the first reference period are written. Based on the results from the Shapiro-Wilk test we then used the parametric or non-parametric alternative for testing the equality of means or equality of distributions.

Tests of Normality							
Variable	Treated	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	Df	Sig.
Marital status	non treated	,403	5	,008	,625	5	,001
	treated	,392	5	,012	,638	5	,002
Level of education (10 categories)	non treated	,326	10	,003	,678	10	,000
	treated	,339	10	,002	,583	10	,000
District of School	non treated	,133	78	,002	,886	78	,000
	treated	,179	78	,000	,823	78	,000
Disadvantages	non treated	,476	7	,000	,492	7	,000
	treated	,391	7	,002	,552	7	,000
Last Occasion	non treated	,453	36	,000	,211	36	,000
	treated	,289	36	,000	,619	36	,000
Age	non treated	,388	4067	,000	,550	4067	,000
	treated	,198	6308	,000			
Gender	non treated	,260	2				
	treated	,260	2	-	-	-	-
School (5 categories)	non treated	,204	5	,200 <sup>*</sup>	,910	5	,467
	treated	,298	5	,169	,853	5	,206
Jobseeker before 2007	non treated	,395	4		,719	4	,019
	treated	,317	4		,880	4	,339
Driving licence	non treated	,333	16	,000	,633	16	,000
	treated	,343	16	,000	,618	16	,000

Based on the results of this testing we used the t-test for two variables: *School (5 categories)* and *Jobseeker before 2007*. The results of comparing the means of these two variables between the treated and non-treated individuals are in the following table:

Independent Samples Test									
		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	
School categories) (5	Equal variances assumed	,534	,486	-,667	8	,524	-448,200	672,461	
	Equal variances not assumed			-,667	6,253	,529	-448,200	672,461	
Jobseeker before 2007	Equal variances assumed	,148	,714	-,524	6	,619	-560,250	1068,632	
	Equal variances not assumed			-,524	5,858	,619	-560,250	1068,632	

For both variables, the difference between the means is not significant. So the variable means of these two variables for treated and non-treated jobseekers are statistically equal.

In the following table, the results of non-parametric tests for the other variables are written.

Gender	Marital status																																
<p><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>1,000<sup>1</sup></td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>,964</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>,439</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is ,05.  <sup>1</sup>Exact significance is displayed for this test.</p>	Null Hypothesis	Test	Sig.	Decision	1 The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1,000 <sup>1</sup>	Retain the null hypothesis.	2 The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,964	Retain the null hypothesis.	3 The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,439	Retain the null hypothesis.	<p><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>,690</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>1,000</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>,802</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is ,05.  <sup>1</sup>Exact significance is displayed for this test.</p>	Null Hypothesis	Test	Sig.	Decision	1 The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,690	Retain the null hypothesis.	2 The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	3 The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,802	Retain the null hypothesis.
Null Hypothesis	Test	Sig.	Decision																														
1 The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1,000 <sup>1</sup>	Retain the null hypothesis.																														
2 The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,964	Retain the null hypothesis.																														
3 The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,439	Retain the null hypothesis.																														
Null Hypothesis	Test	Sig.	Decision																														
1 The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,690	Retain the null hypothesis.																														
2 The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.																														
3 The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,802	Retain the null hypothesis.																														
<p><b>Level of education (10 categories)</b></p> <p><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>,796<sup>1</sup></td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>,988</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>,762</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is ,05.  <sup>1</sup>Exact significance is displayed for this test.</p>	Null Hypothesis	Test	Sig.	Decision	1 The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,796 <sup>1</sup>	Retain the null hypothesis.	2 The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,988	Retain the null hypothesis.	3 The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,762	Retain the null hypothesis.	<p><b>Disadvantages</b></p> <p><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>,466<sup>1</sup></td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>,938</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3 The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>,440</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is ,05.  <sup>1</sup>Exact significance is displayed for this test.</p>	Null Hypothesis	Test	Sig.	Decision	1 The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,466 <sup>1</sup>	Retain the null hypothesis.	2 The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,938	Retain the null hypothesis.	3 The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,440	Retain the null hypothesis.
Null Hypothesis	Test	Sig.	Decision																														
1 The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,796 <sup>1</sup>	Retain the null hypothesis.																														
2 The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,988	Retain the null hypothesis.																														
3 The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,762	Retain the null hypothesis.																														
Null Hypothesis	Test	Sig.	Decision																														
1 The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,466 <sup>1</sup>	Retain the null hypothesis.																														
2 The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,938	Retain the null hypothesis.																														
3 The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,440	Retain the null hypothesis.																														

## Last occasion

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,681	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,878	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,681	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

## School district

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,004	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,004	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,004	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

## Age

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
2	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	,000	Reject the null hypothesis.
3	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

## Driving licence

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,590 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,941	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,584	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.  
<sup>1</sup>Exact significance is displayed for this test.

## Summary:

- Variables, for which the means or the probability distributions of their frequencies are statistically the same:
  - School (5 categories)
  - Jobseeker before 2007
  - Gender
  - Marital status
  - Level of education (10 categories)
  - Disadvantages
  - Last occasion
  - Driving licence

- Variables, for which the probability distributions in the sample of treated and non-treated jobseekers are significantly different: School district
- Age.

#### 4.6.2 2<sup>nd</sup> reference period

Similarly with the first reference period, we tested the equality of means or of probability distributions for the variables in a sample of treated and non-treated individuals in the second reference period. As before, the first step was to test whether there is a normal distribution of every variable. The results are in the following table.

Tests of Normality							
Variable	Treated	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	Df	Sig.
Marital status	non treated	.434	5	.002	.594	5	.001
	treated	.425	5	.004	.603	5	.001
Level of education (10 categories)	non treated	.317	5	.113	.786	5	.062
	treated	.310	5	.132	.861	5	.234
District of School	non treated	.126	79	.003	.902	79	.000
	treated	.168	79	.000	.833	79	.000
Disadvantages	non treated	.478	8	.000	.458	8	.000
	treated	.329	8	.011	.579	8	.000
Last Occasion	non treated	.263	39	.000	.661	39	.000
	treated	.289	39	.000	.631	39	.000
Age	non treated	.217	36565	.000			
	treated	.210	24304	.000			
Gender	non treated	.260	2				
	treated	.260	2				
School (5 categories)	non treated	.126	79	.003	.902	79	.000
	treated	.168	79	.000	.833	79	.000
Jobseeker before 2007	non treated	.435	4		.643	4	.002
	treated	.423	4		.664	4	.004
Driving licence	non treated	.338	16	.000	.623	16	.000
	treated	.354	16	.000	.606	16	.000

In this case, only the variable *Level of education (10 categories)* is normally distributed. Based on this result, the second step is to test whether the mean of this variable or the mean and the distribution of the other variables are the same between the samples of treated and non-treated individuals. This is tested by t-test in the case of the variable *Level of education (10 categories)*, which is normally distributed and by a non-parametric alternative for the other variables.

The results are in the following two tables.

Independent Samples Test								
		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
School categories) (5	Equal variances assumed	1.738	.224	.536	8	.607	2452.200	4577.173
	Equal variances not assumed			.536	6.259	.611	2452.200	4577.173

As we can see in the table, based on the significance of the t-test, the means of the variable *Level of education* are the same in the sample of treated and non-treated individuals.

<p>Marital status</p> <p style="text-align: center;"><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th></th> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>.548<sup>1</sup></td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>.819</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>.465</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is .05.  <sup>1</sup>Exact significance is displayed for this test.</p>		Null Hypothesis	Test	Sig.	Decision	1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.548 <sup>1</sup>	Retain the null hypothesis.	2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.819	Retain the null hypothesis.	3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.465	Retain the null hypothesis.	<p>Gender</p> <p style="text-align: center;"><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th></th> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>.333<sup>1</sup></td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>.270</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>.121</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is .05.  <sup>1</sup>Exact significance is displayed for this test.</p>		Null Hypothesis	Test	Sig.	Decision	1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.333 <sup>1</sup>	Retain the null hypothesis.	2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.270	Retain the null hypothesis.	3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.121	Retain the null hypothesis.
	Null Hypothesis	Test	Sig.	Decision																																					
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.548 <sup>1</sup>	Retain the null hypothesis.																																					
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.819	Retain the null hypothesis.																																					
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.465	Retain the null hypothesis.																																					
	Null Hypothesis	Test	Sig.	Decision																																					
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.333 <sup>1</sup>	Retain the null hypothesis.																																					
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.270	Retain the null hypothesis.																																					
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.121	Retain the null hypothesis.																																					
<p>School (5 categories)</p> <p style="text-align: center;"><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th></th> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>1.000<sup>1</sup></td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>.819</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>.917</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is .05.  <sup>1</sup>Exact significance is displayed for this test.</p>		Null Hypothesis	Test	Sig.	Decision	1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1.000 <sup>1</sup>	Retain the null hypothesis.	2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.819	Retain the null hypothesis.	3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.917	Retain the null hypothesis.	<p>Disadvantages</p> <p style="text-align: center;"><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th></th> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>.648<sup>1</sup></td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>.627</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>.599</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is .05.  <sup>1</sup>Exact significance is displayed for this test.</p>		Null Hypothesis	Test	Sig.	Decision	1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.648 <sup>1</sup>	Retain the null hypothesis.	2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.627	Retain the null hypothesis.	3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.599	Retain the null hypothesis.
	Null Hypothesis	Test	Sig.	Decision																																					
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1.000 <sup>1</sup>	Retain the null hypothesis.																																					
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.819	Retain the null hypothesis.																																					
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.917	Retain the null hypothesis.																																					
	Null Hypothesis	Test	Sig.	Decision																																					
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.648 <sup>1</sup>	Retain the null hypothesis.																																					
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.627	Retain the null hypothesis.																																					
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.599	Retain the null hypothesis.																																					
<p>Last occasion</p>	<p>Jobseeker before 2007</p>																																								



Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,005	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,050	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,005	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1,000 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,773	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.  
<sup>1</sup>Exact significance is displayed for this test.

Age

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
2	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	,000	Reject the null hypothesis.
3	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Driving licence

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,341 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,415	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,336	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.  
<sup>1</sup>Exact significance is displayed for this test.

School district

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,004	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,021	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,004	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

- Summary:**
- Variables, for which the means or the probability distributions of their frequencies are statistically the same:
    - School (5 categories)
    - Jobseeker before 2007
    - Gender
    - Marital status
    - Level of education (10 categories)
    - Disadvantages
    - Driving licence
  - Variables, for which the probability distributions in the sample of treated and non-treated jobseekers are significantly different:
    - Last occasion School district
  - Age

### 4.6.3 3<sup>rd</sup> reference period

The results of testing normality using the Shapiro-Wilk test are in the following table:

Tests of Normality							
Variable	Treated	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	Df	Sig.
		Marital status	non treated	.440	5	.002	.588
treated	.433		5	.003	.594	5	.001
Level of education (10 categories)	non treated	.377	10	.000	.605	10	.000
	treated	.303	10	.010	.651	10	.000
District of School	non treated	.121	79	.006	.896	79	.000
	treated	.170	79	.000	.774	79	.000
Disadvantages	non treated	.483	8	.000	.453	8	.000
	treated	.332	8	.010	.650	8	.001
Last Occasion	non treated	.402	41	.000	.226	41	.000
	treated	.321	37	.000	.628	37	.000
Age	non treated	.181	7186	.000			
	treated	.196	16230	.000			
Gender	non treated	.260	2	.000			
	treated	.260	2	.000			
School (5 categories)	non treated	.302	5	.153	.793	5	.072
	treated	.223	5	.200 <sup>*</sup>	.913	5	.487
Jobseeker before 2007	non treated	.436	4		.639	4	.002
	treated	.432	4		.647	4	.002
Driving licence	non treated	.347	16	.000	.615	16	.000
	treated	.358	16	.000	.607	16	.000

The only variable that has a normal distribution is *School (5 categories)*. For this variable we used the t-test to compare means for treated and non-treated individuals. The results are in the next table.

Independent Samples Test									
		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	
School categories)	(5	Equal variances assumed	1.142	.316	-1.125	8	.293	-1808.800	1607.525
		Equal variances not assumed			-1.125	6.308	.301	-1808.800	1607.525

According to the significance of the test we can say that the means of this variable between treated and non-treated individuals are not different.

In the next table are the results of non-parametric tests of equality of the variables

distributions between treated and non-treated individuals.

**Gender**

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,333 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,270	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,121	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

**Marital status**

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,548 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,530	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

**Level of education (10 categories)**

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,631 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,988	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,623	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

**District of school**

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,000	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

**Disadvantages**

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,798 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,964	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,751	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

**Last occasion**

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,195	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,276	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,195	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

**Driving licence**

**Summary:** Variables, for which the means or the probability distributions of their frequencies are statistically the same:

- School (5 categories)
- Jobseeker before 2007
- Gender
- Marital status
- Level of education (10 categories)

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.341 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.415	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.336	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

<sup>1</sup>Exact significance is displayed for this test.

- Disadvantages
- Driving licence
- Last occasion
- Age
- Variables, for which the probability distributions in the sample of treated and non-treated jobseekers are significantly different:
  - District of school

#### 4.6.4 4<sup>th</sup> reference period

In the last reference period the results of testing the normality of variables frequencies in the sample of treated and non-treated jobseekers are the following:

Tests of Normality							
Variable	Treated	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	Df	Sig.
		Marital status	non treated	.448	5	.001	.579
treated	.437		5	.002	.591	5	.000
Level of education (10 categories)	non treated	.374	10	.000	.578	10	.000
	treated	.293	10	.015	.619	10	.000
School district	non treated	.141	79	.001	.899	79	.000
	treated	.153	79	.000	.827	79	.000
Disadvantages	non treated	.461	7	.000	.511	7	.000
	treated	.328	7	.022	.769	7	.020
Last Occasion	non treated	.274	38	.000	.615	38	.000
	treated	.332	38	.000	.607	38	.000
Age	non treated	.125	19037	.000			
	treated	.205	18092	.000			
Gender	non treated	.260	2	.000			
	treated	.260	2	.000			
School (5 categories)	non treated	.300	5	.162	.774	5	.049
	treated	.307	5	.139	.856	5	.213
Jobseeker before 2007	non treated	.438	4	0	.636	4	.002
	treated	.434	4	0	.643	4	.002
Driving licence	non treated	.340	16	.000	.615	16	.000
	treated	.359	16	.000	.609	16	.000

Similarly with the third reference period, only one variable has a normal distribution, *School (5 categories)*. For this variable, we used the t-test to compare the means for treated and non-treated individuals. The results are in the next table.

Independent Samples Test									
		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	
School categories)	(5	Equal variances assumed	.420	.535	.070	8	.946	189.000	2709.096
		Equal variances not assumed			.070	7.445	.946	189.000	2709.096

Based on the significance of the test we can say that the means of this variable is not different between the samples of treated and non-treated individuals. The results of the other variables testing are in the following table:

<p><b>Gender</b></p> <p style="text-align: center;"><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th></th> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>1,000<sup>1</sup></td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>,964</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>1,000</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is ,05.  <sup>1</sup>Exact significance is displayed for this test.</p>		Null Hypothesis	Test	Sig.	Decision	1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1,000 <sup>1</sup>	Retain the null hypothesis.	2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,964	Retain the null hypothesis.	3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	1,000	Retain the null hypothesis.	<p><b>Marital status</b></p> <p style="text-align: center;"><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th></th> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>1,000<sup>1</sup></td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>1,000</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>,914</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is ,05.  <sup>1</sup>Exact significance is displayed for this test.</p>		Null Hypothesis	Test	Sig.	Decision	1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1,000 <sup>1</sup>	Retain the null hypothesis.	2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,914	Retain the null hypothesis.
	Null Hypothesis	Test	Sig.	Decision																																					
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1,000 <sup>1</sup>	Retain the null hypothesis.																																					
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,964	Retain the null hypothesis.																																					
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	1,000	Retain the null hypothesis.																																					
	Null Hypothesis	Test	Sig.	Decision																																					
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1,000 <sup>1</sup>	Retain the null hypothesis.																																					
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.																																					
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,914	Retain the null hypothesis.																																					
<p><b>Level of education (10 categories)</b></p> <p style="text-align: center;"><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th></th> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>,853<sup>1</sup></td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>,988</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>,850</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is ,05.  <sup>1</sup>Exact significance is displayed for this test.</p>		Null Hypothesis	Test	Sig.	Decision	1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,853 <sup>1</sup>	Retain the null hypothesis.	2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,988	Retain the null hypothesis.	3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,850	Retain the null hypothesis.	<p><b>District of school</b></p> <p style="text-align: center;"><b>Hypothesis Test Summary</b></p> <table border="1"> <thead> <tr> <th></th> <th>Null Hypothesis</th> <th>Test</th> <th>Sig.</th> <th>Decision</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Mann-Whitney U Test</td> <td>,543</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>2</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kolmogorov-Smirnov Test</td> <td>,551</td> <td>Retain the null hypothesis.</td> </tr> <tr> <td>3</td> <td>The distribution of count is the same across categories of treated.</td> <td>Independent-Samples Kruskal-Wallis Test</td> <td>,543</td> <td>Retain the null hypothesis.</td> </tr> </tbody> </table> <p>Asymptotic significances are displayed. The significance level is ,05.</p>		Null Hypothesis	Test	Sig.	Decision	1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,543	Retain the null hypothesis.	2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,551	Retain the null hypothesis.	3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,543	Retain the null hypothesis.
	Null Hypothesis	Test	Sig.	Decision																																					
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,853 <sup>1</sup>	Retain the null hypothesis.																																					
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,988	Retain the null hypothesis.																																					
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,850	Retain the null hypothesis.																																					
	Null Hypothesis	Test	Sig.	Decision																																					
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,543	Retain the null hypothesis.																																					
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,551	Retain the null hypothesis.																																					
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,543	Retain the null hypothesis.																																					

## Disadvantages

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1,000 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,938	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,949	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

## Last occasion

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,003	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,023	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,003	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

## Jobseeker before 2007

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	1,000 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,773	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

## Driving licence

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,897 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,941	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,895	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

## Age

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
2	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	,000	Reject the null hypothesis.
3	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

## Summary:

- Variables, for which the means or the probability distributions of their frequencies are statistically the same:
  - School (5 categories)
  - Jobseeker before 2007
  - Gender
  - Marital status
  - Level of education (10 categories)
  - Disadvantages
  - Driving licence
  - District of school
- Variables, for which the probability distributions in the sample of treated and non-treated jobseekers are significantly different:
  - Last occasion
- Age

There is just one problematic variable which is not possible to eliminate, and that is age; significant differences between treated and non-treated samples have been identified. As

was obvious already in the histograms in the sub-chapter which described age, there was a left-side distribution of treated ages and right-side distribution of non-treated ages in three reference periods. That fact will influence the shrinkage of the matched samples, because the intersection between treated and non-treated distributions is too low.

## 4.7 Qualitative survey of graduate work experience

This qualitative part was incorporated in the evaluation because the evaluators wanted to outline even partial motivations, aspirations, real outputs and the results of the treated individuals. The main reason for this part of the research was verifying a theory about the change of traineeship. Qualitative research was carried out through interviews by phone. COLSaF provided a database of 48 contacts for treated individuals who were asked for an interview. The database contained individuals from every region of the SR (i.e. 8 regions) and three individuals for men and women, in total 48 contacts.

Finally we carried out 41 interviews composed of 23 women and 18 men from every Slovak region.

In the scheme below is described the expected theory of the change of the intervention and the prepared topics for interviews which came from three basic parts:

### A. Activities of the intervention

In the first section of the questions which were posed to our respondents, we wanted to uncover the motivation to take part in the intervention and identify activities which could lead to immediate service for the jobseeker and to increase his/her employability on the open labour market.

During the interview we asked questions such as:

- 1) Where did you learn about the intervention?
- 2) Did you find an employer for traineeship on your own, or did PES assist you?
- 3) How did you find the employer?
- 4) Why did you decide to participate in the traineeship?
- 5) Have you matched your graduated profession with the profession of the place where you performed traineeship?
- 6) Did you do what was agreed upon with the employer in the agreement?
- 7) Have you attended any training during traineeship?

### B. Immediate outputs of the intervention

Through those sorts of questions we wanted to identify the provided services products that jobseekers carried out during their traineeship. We wanted to lead a dialogue with the jobseeker about their emotions coming out from completing the intervention.

- Have you met with your initial aspiration of traineeship?
- Which skills and knowledge have you gained during traineeship?
- Have you gained any contacts for other employers or references for any job?

### C. Outcomes

This last group of questions should identify the perception of short-term and mid-term effects of traineeship.

- Do you think your traineeship was successful? Why? Do you think that traineeship helped you to get a job?
- Which knowledge and skills have you used for your work?

What would you change?



#### 4.7.1 Conclusions from the interviews

One third of asked respondents had information about the intervention before this was offered by the Public Employment Services office. In other words, one third of those treated applied for intervention without any impulse, they knew that they were eligible and they wanted to attend traineeship.

The rest of the eligible jobseekers were informed about the intervention by the PES office and, afterwards, the interventions were offered as well.

Most of the respondents answered that they chose their placing for traineeship from the list provided by the PES office and only about 2 jobseekers from 5 had selected a place for intervention before they applied at the PES office for intervention. These are the same jobseekers which stated that they knew about the intervention before the PES office informed them.

Just one third of treated jobseekers answered that they would like to find a job through intervention, the rest of the respondents had the aspiration just to have some practical experience or deepen existing skills.

All respondents admit that they really carried out work that was agreed upon in the agreement before they started traineeship; all participants denied any abuse. But only in a few cases did jobseekers work in the field from which they graduated. Just about 10 % of respondents admitted that they worked in a business matching the type and specialization of the education they had completed.

Overall, most of the respondents were satisfied with the provided intervention. They are sure that they have met with the expectations of traineeship, even though these expectations were minimal and, in most cases, they did not aspire to find a job and stay employed in the field in which they carried out traineeship.

For the question focused on gained skills, most of the respondents answered directly that they learned to communicate with people, they gained some interpersonal skills in the working environment because that was their first experience in almost a real job without support of school, or schoolmates, and that is why they feel this intermediate step was important. A few of the asked respondents mentioned that although the area in which they worked provided some sort of course, training was provided to the participants only rarely. In those few cases, the trainings were focused on MS Excel, or Access, internal supply system, or work with a cash machine. In most of the cases, graduates were distributed into public institutions; only a few participants stated that they attended traineeship in the private sector or third sector (about 10 %).

Only about 1 treated individual from 10 stayed at the public institution where they worked during the traineeship, this was the Public Employment Services office – registration of jobseekers, or Social Insurance Agency – as an administrator. Finally, those treated found other jobs and used those public institutions as waystations which helped them to gain references for other employers. Another identified benefit from traineeship was contacts and friendship created during traineeship, but none of the treated identified that through those contacts they would find a job.

About 20 % of asked jobseekers wanted to go to traineeship just because they were waiting for another year to enrol in another school.

#### **What treated jobseekers would like to change?**

The most frequent proposals of interviewed jobseekers concern the motivation allowance based on living wage that is not sufficient motivation for traineeship. This is one identified barrier for wider use of this active labour market policy measure.

Another frequent proposal is based on the treatment period. About 3 asked graduates from 10 states that 6 months is not long enough a period to show what they know, to present their real potential even though they have just 4 hours per day and they depend on the decisions of their tutor. The tutor mostly has in the mornings some urgent work and, only after he has finished what he must, then he can care about graduates. The respondents are sure that **longer working days and increased allowance** during traineeship would also increase their chances of employment.

Treated jobseekers identify the need to gain something tangible through traineeship, something like a **recommendation, or certificate**, which could enforce the positions of seeking jobseekers in a job interview and would upgrade the intervention to a more serious level. Treated jobseekers would like to seriously make an effort to gain the chance of a job through preparing as much as possible.

Traineeship should be **better fitted to the type of education or working positions** which are attractive for graduates or where they see themselves. It is possible to expect that a higher involvement of subjects from the private sector would also increase the efficiency of the intervention. There should be prepared a motivation tool for firms and organizations to offer traineeship. For instance, traineeship could be extensive in case of financial contribution to the allowance for graduates. Intervention could be varied into some types of *modus operandi* with some motivation of companies to offer graduates jobs, sustained for some protected period.

The last type of comment from respondents was about administration whilst applying for traineeship. Some data should be electronically exchanged between PES offices if traineeship is to be carried out in another city than where the jobseeker is registered.

## 4.8 Net effects of traineeship

### 4.8.1 Analysis of influences on employability

In the table next to the text are correlation coefficients and their significance on the dependent variable Placed on LM and Assessment Base and other independent variables that are the characteristics of treated and non-treated units and their living environment.

For the dependent variable **Placed on the labour market** we can see that:

- men are placed on the labour market longer than women, the correlation is significant but weak;
- older individuals are placed longer, but the correlation coefficient is not very high, although it is significant, in other words its power or influence is not very strong;
- a longer period of registration and total period of all registrations cause a shorter placement on LM, which is confirmation of the expected assumption, and the correlation is moderate;
- the classification of an individual into Reference period 2 has a negative impact on placement on LM; other reference periods have a positive impact on placement;
- single jobseekers are unemployed for a shorter time than other marital statuses, the correlation is weak;

Variable	Placed_on_LM_shares		Average assessment base	
	Pearson Correlation	Sig. (2-tailed)	Pearson Correlation	Sig. (2-tailed)
Placed_on_LM_shares	1,000		,615**	0,000
Average assessment base	,615**	0,000	1,000	
Gender	-,019**	0,000	-,085**	0,000
Age	,113**	0,000	,249**	0,000
Unemployed in months	-,306**	0,000	-,335**	0,000
Total period of all registrations in months (colsaf)	-,412**	0,000	-,387**	0,000
The average gross wage in the region of perm. residence	,089**	0,000	,140**	0,000
The proportion of women in the district of perm. residence	,042**	0,000	,058**	0,000
Surface of district of permanent residence	-,048**	0,000	-,092**	0,000
The density of population in the district of perm. residence	,042**	0,000	,081**	0,000
The number of municipalities in the district of perm. residence	-,069**	0,000	-,123**	0,000
The number of cities in the district of perm. residence	-0,005	0,088	-,032**	0,000
The registered unemployment rate in the district of perm. residence	-,135**	0,000	-,184**	0,000
Inhabitants density	,033**	0,000	,100**	0,000
Population_of_municipality_2011	,022**	0,000	,031**	0,000
Change_of_population:15years	-0,004	0,129	-,012**	0,000
Distance_from_PESoffice	-,027**	0,000	-,040**	0,000
period=1.0	,033**	0,000	-,047**	0,000
period=2.0	-,061**	0,000	,064**	0,000
period=3.0	,029**	0,000	-,025**	0,000
period=4.0	,023**	0,000	-,022**	0,000
marital_status=registered partners	-,007**	0,012	-0,002	0,558
marital_status=divorced	-,012**	0,000	-,008**	0,004
marital_status=single	,087**	0,000	,008**	0,004
marital_status=widow	-0,002	0,418	0,001	0,755
marital_status=married	-,086**	0,000	-,006*	0,022
education_STUPEN=Not finished education	-,009**	0,002	-0,003	0,225
education_STUPEN=Primary education	-,115**	0,000	-,085**	0,000
education_STUPEN=Lower secondary professional education	-,017**	0,000	-,017**	0,000
education_STUPEN=Secondary vocational education	-,039**	0,000	-,038**	0,000
education_STUPEN=Full secondary vocational education	,058**	0,000	,025**	0,000
education_STUPEN=Full secondary comprehensive education	-,105**	0,000	-,092**	0,000
education_STUPEN=Upper vocational education	-0,004	0,110	-0,005	0,077
education_STUPEN=Bachelor	-,048**	0,000	-,047**	0,000
education_STUPEN=Master	,144**	0,000	,167**	0,000
education_STUPEN=Doctoral	-0,001	0,607	-0,005	0,093
school=primary school	-,115**	0,000	-,085**	0,000
school=secondary vocational school	,025**	0,000	-0,002	0,571
school=vocational school	-0,002	0,544	-0,008**	0,005
school=comprehensive school	-,103**	0,000	-,091**	0,000
school=colledge	,116**	0,000	,137**	0,000
disadvantage=no disadvantage	,154**	0,000	,123**	0,000
disadvantage=graduate	-,064**	0,000	-,042**	0,000
disadvantage=long - term unemployed	-,146**	0,000	-,127**	0,000
disadvantage=low education level	-0,002	0,411	0,000	0,897
disadvantage=organizational	,006*	0,032	0,005	0,070
disadvantage=poor working discipline	-0,005	0,097	-,007**	0,015
disadvantage=care	-,009**	0,001	-0,004	0,178
disadvantage=disabled	-,012**	0,000	-,014**	0,000
unemployed_before_2007=< 1 year	,023**	0,000	-0,004	0,181
unemployed_before_2007=1 - 3 years	0,003	0,252	-,027**	0,000
unemployed_before_2007=> 3 years	-,017**	0,000	-,026**	0,000
unemployed_before_2007=no evidence	-,015**	0,000	,026**	0,000
region=Bratislavský region	,054**	0,000	,122**	0,000
region=Tmavský region	,069**	0,000	,084**	0,000
region=Trenčiansky region	,056**	0,000	,057**	0,000
region=Nitriansky region	,015**	0,000	-,008**	0,005
region=Žilinský region	,015**	0,000	,022**	0,000
region=Banskobystrický region	-,050**	0,000	-,057**	0,000
region=Prešovský region	-,088**	0,000	-,104**	0,000
region=Košícký region	-,033**	0,000	-,050**	0,000
Treated/non-treated	-,052**	0,000	-,197**	0,000

- from significant education levels, Master's education (positive impact) and primary education (negative impact) have the greatest impact
- from school types, college has a positive impact and primary school has a negative impact;
- almost all significant disadvantages types have a negative impact; if an individual has no disadvantage, he is placed on LM for longer;
- all regions have significant correlations but the correlation is weak.

For the **Assessment base** we can state that:

- older individuals have a higher assessment base, the correlation is significant and moderate;
- men have a higher assessment base than women, the correlation is significant and moderate;
- the period of registration and total period of all registrations have a negative impact on the assessment base, which means that if an individual is unemployed for a longer time, then he has a smaller assessment base, the correlation is significant and is moderate;
  - in Period 2 the assessment base is higher, but the correlations are weak;
- marital status has very weak correlations with the assessment base;
- Master's education level has the biggest correlation from all education levels and college is similar,
- with the disadvantaged long term unemployed the assessment base is lower, with no disadvantage the assessment base is higher;
- if an individual has no registration before 2007, the assessment base is higher, all other unemployment durations have a negative impact, but the correlation is weak;
- Bratislava region has the biggest positive correlation from all regions;

In the following table the coefficient of linear regression for 5 types of dependent variables are listed. If there is no coefficient, that variable wasn't significant in a linear regression model.

Dependent variable	Placed_on_LM	Self employed	Full time employed	Part time job	Individual barrier	Assement base
(Constant)	,925	,146	,932	-,057	,101	109,825
Treated/non-treated	,156	,004	,152	,014	-,010	-14,905
Gender	-,028	-,006	-,022	,004	,054	
Age	,005	,000	,004	,002		4,810
Unemployed in months	-,004		-,004	,001	,001	1,454
Total period of all registrations in months (colsaf)	-,008	,000	-,008	,001	-,001	-3,313
The proportion of women in the district of perm. residence		-,003				
The average gross wage in the region of perm. residence	-,001		-,001		,000	,359
Surface of district of permanent residence						,018
The density of population in the district of perm. residence						-,008
The number of cities in the district of perm. residence	,005		,005	-,001		
The number of municipalities in the district of perm. residence						-,299
The registered unemployment rate in the district of perm. residence	-,001		-,001	,000	,001	-,702
Inhabitants density						,009
Change_of_population:15years				,002		
District of permanent residence						,852
period=1	,042		,041	-,004	-,007	-31,633
period=3	,104	-,003	,108	,018	,007	
period=4	,104	-,003	,109	,045	,004	
marital_status=registered partners					,180	
marital_status=divorced	-,068		-,066	-,012	,143	-33,128
marital_status=widow					,176	
marital_status=married	-,109		-,109	-,006	,168	-27,368
education_level=Primary education	-,108		-,107			-40,144
education_level=Lower secondary professional education	-,029		-,030		,019	-30,964
education_level=Secondary vocational education	-,035		-,036		,020	-28,446
education_level=Full secondary comprehensive education	-,099		-,100	,014		
education_level=Upper vocational education	-,054		-,052			
education_level=Bachelor	-,132	-,024	-,132	,019		
education_level=Master		-,024				120,306
education_level=Doctoral		-,030				
school=primary shool				-,006	,074	
school=comprehensive school				-,008		
school=vocational school		,001				-4,823
school=colledge	,093	,024	,094	-,020	-,023	
disadvantage=graduate	-,038	-,001	-,037	-,008	,004	
disadvantage=long - term unemployed	-,095	-,002	-,093	,008	-,005	-21,192
disadvantage=organizational	,196		,200			
disadvantage=care	-,102		-,097	,025	,087	
disadvantage=disabled	-,120		-,116			
unemployed_before_2007=< 1 year	,063	,002	,061	-,014	,010	14,982
unemployed_before_2007=1 - 3 years	,141		,141	-,021	,012	11,881
unemployed_before_2007=> 3 years	,252		,251	-,049	,015	39,531
region=Bratislavský region	,389	-,004	,402	-,005	,056	-48,032
region=Trnavský region	,142	-,004	,148	-,003	,009	48,953
region=Trenčiansky region	,102	-,003	,106	-,002	,007	30,706
region=Nitriansky region	,066	-,002	,070			20,864
region=Žilinský region	,095		,095	-,003	,022	
region=Banskobystrický region	,065		,068		,005	7,387
region=Košícký region	,126	-,003	,132	-,004	,021	-49,870

## 4.8.2 “Post-only non-equivalent comparison design” Method

There are several methodologies on how to estimate the net effect of the interventions, one of the most simplistic methodologies is the difference of average treatment effects between a treated group and control without the matching of individuals' characteristics. That is why this method is not sufficiently robust.

We decided to apply this methodology to the data because we intend to provide different results of net estimated treatment effects on the promotion of self-employment.

As can be seen on the bottom line of the table, the non-treated group is composed of more than 66 thousand individuals and the treated group is composed of almost 65 jobseekers which were exposed to the intervention. In total, there are 131 thousand individuals, which is already a serious number of items of jobseekers; in reality, that sample is almost 5 % of the working-age Slovak population. For this method there was used the biggest possible number of jobseekers that had recorded correct and exhaustive data. This fact ranks among the advantages of the Post-only non-equivalent design.

Ref. Period			non-treated	treated
1	N	Valid	4067	6308
		Missing	0	0
2	N	Valid	36565	24304
		Missing	0	0
3	N	Valid	7186	16230
		Missing	0	0
4	N	Valid	19037	18092
		Missing	0	0
In total			66855	64934

### *Measuring of employability*

The frequency table below the text represents at a glance the average probability of treated and control groups across the set reference periods sustained in the impact period on the labour market in the first column. To recap, being placed on the open market means, for the purposes of this evaluation report, to be placed in a full-time job, or to be self-employed. The numbers there are ranked from 0 to 1. 0 means that jobseekers were not placed on the labour market. In the last part of the table (i.e. the last three columns) are presented the total average percentage of probabilities to be placed on the labour market for different parts of the impact period.

While on average for all reference periods 34 % of the treated did not find any job during the two years long impact period, just more than 15 % of non-treated jobseekers did not find a job. There was 10 % probability that one treated jobseeker was sustained on the LM for the entire impact period of the first reference sample, while one control sustained on the labour market in the same period had less than 1 % probability.

The yellow bar chart integrated into the table represents the tendency of the jobseekers in the different samples to be employed and sustained on the labour market in a full-time job or to be self-employed.

To state simply, the more successful are those cumulative percent columns that have more yellow area. In the first reference periods the treated have more individuals that were sustained on the labour market mainly longer than the controls. For instance, in the first reference period it was indicated that almost 29 % were employed for 70 % of the impact period for the treated, while it was just 25 % for the non-treated.

Group of sample	Share of impact period sustained on LM	Reference period 1			Reference period 2			Reference period 3			Reference period 4			Ref. Periods in total		
		Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent
non treated	0	633	15,6	15,6	6472	17,7	17,7	819	11,4	11,4	2513	13,2	13,2	10437	15,6	15,6
	0,1	600	14,8	84,4	4580	12,5	82,3	537	7,5	88,6	1742	9,2	86,8	7459	11,2	84,4
	0,2	485	11,9	72,5	3546	9,7	72,6	622	8,7	79,9	1327	7,0	79,8	5980	8,9	75,4
	0,3	428	10,5	62,0	3614	9,9	62,7	467	6,5	73,4	1187	6,2	73,6	5696	8,5	66,9
	0,4	326	8,0	54,0	4463	12,2	50,5	460	6,4	67,0	1714	9,0	64,6	6963	10,4	56,5
	0,5	313	7,7	46,3	4948	13,5	37,0	467	6,5	60,5	1143	6,0	58,8	6871	10,3	46,2
	0,6	425	10,4	35,8	4037	11,0	25,9	484	6,7	53,8	1649	8,7	49,9	6595	9,9	36,4
	0,7	494	12,1	23,7	2406	6,6	19,4	538	7,5	46,3	1872	9,8	40,1	5310	7,9	28,4
	0,8	335	8,2	15,4	1385	3,8	15,6	891	12,4	33,9	2682	14,1	26,0	5293	7,9	20,5
	0,9	23	0,6	14,9	1087	3,0	12,6	1412	19,6	14,3	2914	15,3	10,7	5436	8,1	12,4
	1	5	0,1	14,8	27	0,1	12,5	489	6,8	7,5	294	1,5	9,2	815	1,2	11,2
	Total	4067			36565			7186			19037			66855	100	
treated	0	1407	22,3	22,3	7384	30,4	30,4	5801	35,7	35,7	7542	41,7	41,7	22134	34,1	34,1
	0,1	458	7,3	77,7	1665	6,9	69,6	1076	6,6	64,3	1123	6,2	58,3	4322	6,7	65,9
	0,2	351	5,6	72,1	1391	5,7	63,9	878	5,4	58,8	901	5,0	53,3	3521	5,4	60,5
	0,3	489	7,8	64,4	1913	7,9	56,0	1334	8,2	50,6	1365	7,5	45,8	5101	7,9	52,6
	0,4	303	4,8	59,6	1152	4,7	51,3	743	4,6	46,1	829	4,6	41,2	3027	4,7	48,0
	0,5	501	7,9	51,6	1757	7,2	44,1	1212	7,5	38,6	1369	7,6	33,6	4839	7,5	40,5
	0,6	330	5,2	46,4	1242	5,1	38,9	818	5,0	33,5	937	5,2	28,5	3327	5,1	35,4
	0,7	351	5,6	40,8	1200	4,9	34,0	752	4,6	28,9	778	4,3	24,2	3081	4,7	30,7
	0,8	604	9,6	31,3	2069	8,5	25,5	1206	7,4	21,5	1174	6,5	17,7	5053	7,8	22,9
	0,9	509	8,1	23,2	1644	6,8	18,7	849	5,2	16,2	819	4,5	13,1	3821	5,9	17,0
	1	1005	15,9	7,3	2887	11,9	6,9	1561	9,6	6,6	1255	6,9	6,2	6708	10,3	6,7
	Total	6308			24304			16230			18092			64934	100	

In the next table is presented the net effects of the intervention where there are subtracted frequency tables of the treated and controls for the reference periods.

It appears that across all reference periods there was a higher probability of the treated in finding a job in comparison to the controls - about 6 to almost 29% probability. On the other side, in the table it is visible that if somebody found a job from the treated it was for a longer time on average, because there is, in the bottom part of the impact period sustained on the labour market, on the bottom of the table positive differences. In the first period, for instance, one treated jobseeker had almost a 6 % higher probability to be employed for the entire two years after finishing traineeship.

Share of impact period sustained on LM	Reference period 1	Reference period 2	Reference period 3	Reference period 4	Ref. Periods in total
	Percent	Percent	Percent	Percent	Percent
0	6,7	12,7	24,3	28,5	18,5
0,1	-7,5	-5,7	-0,8	-2,9	-4,5
0,2	-6,4	-4,0	-3,2	-2,0	-3,5
0,3	-2,8	-2,0	1,7	1,3	-0,7
0,4	-3,2	-7,5	-1,8	-4,4	-5,8
0,5	0,2	-6,3	1,0	1,6	-2,8
0,6	-5,2	-5,9	-1,7	-3,5	-4,7
0,7	-6,6	-1,6	-2,9	-5,5	-3,2
0,8	1,3	4,7	-5,0	-7,6	-0,1
0,9	7,5	3,8	-14,4	-10,8	-2,2
1	15,8	11,8	2,8	5,4	9,1

On the table below the text are presented the estimated average performances of the traineeship by PES offices. There are six different dependent variables which should refer to the effects of the intervention. The first dependent variable which was measured is average wage translated from the average assessment base in Euros based on the records of the Social Insurance Agency. The other effects are devoted to the placement of the jobseekers on the labour market in the form of part-time, full-time job, or self-employed. With that kind of registration we can consider that particular jobseeker to be a success because it is out of the registration of the jobseekers and has a financial resource, even if in the registration "part-time job" it isn't comprehensive success of employability, but the

jobseeker keeps in touch with the labour market. The other registration refers to individual barriers for entrance to the labour market due to the need to do personal assistance for family relatives or care for a child. The last dependent variable describes total average months registration with the Social Insurance Agency (SIA) i.e. out of the jobseeker database of the PES office.

Effect of the intervention	Group of sample	Ref. Period 1 Mean	Ref. Period 2 Mean	Ref. Period 3 Mean	Ref. Period 4 Mean	In total Mean
Average assessment base	non treated	429	494	501	526	421
Average assessment base	treated	432	483	499	490	324
Placed on labour market	non treated	0,39	0,38	0,57	0,53	0,46
Placed on labour market	treated	0,56	0,50	0,45	0,41	0,43
Self-employed	non treated	0,00	0,01	0,00	0,00	0,00
Self-employed	treated	0,01	0,01	0,00	0,00	0,00
Part-time job	non treated	0,00	0,00	0,01	0,04	0,01
Part-time job	treated	0,00	0,02	0,04	0,08	0,03
Full-time job	non treated	0,39	0,38	0,56	0,52	0,41
Full-time job	treated	0,55	0,49	0,45	0,40	0,38
Individual barrier for entrance to LM	non treated	0,05	0,04	0,05	0,03	0,03
Individual barrier for entrance to LM	treated	0,04	0,03	0,03	0,03	0,03
Average assessment base	net effect	3	-11	-3	-36	-97
Placed on labour market		0,17	0,12	-0,11	-0,12	-0,03
Self-employed		0,01	0,00	0,00	0,00	0,00
Part-time job		0,00	0,02	0,03	0,04	0,02
Full-time job		0,16	0,12	-0,12	-0,12	-0,03
Individual barrier for entrance to LM		-0,01	0,00	-0,02	-0,01	-0,01

The averages of wages and months of different types of registrations are presented in the table and there is also a difference between the treated and controls of jobseekers which are all target groups of the evaluated intervention. That statistical statement is confirmed through the results from the last table. There are significant differences between treated and controls in the achieved average assessment base, during the impact period of 24 months after the intervention finished. Only in the first reference period was the assessment base of the placed participants of the intervention higher (almost 3 Euros per month) on average within 24 months after intervention in comparison with the controls, but in total average across all reference periods the treated had on average almost 100 Euros lower assessment base.

As is obvious from the table above, graduates prefer to find a job on the labour market and it was not frequent to establish their business in self-employment immediately after leaving school. There is just a very small but significant difference between the treated and non-treated in part-time jobs where the group of treated is a little more successful. The biggest difference is obvious in the ability of jobseekers placed in a full-time job on the open labour market. In the first two reference periods there are positive net effects which brought an increased average probability of the treated to be employed on the labour market longer than between 12 to 17 % of the impact period. But in the last two reference periods the net effects were estimated as being negative. In this respect it is necessary to mention that in those reference periods there started a total influence of the highest unemployment rate on the labour market and the influence was maximum.

In total we can say that across the reference periods, treated jobseekers were placed on the labour market for 46 % of the impact period of 24 months and controls were sustained on the labour market on average for 43 % of the reference period.



In the table below the text are presented the results of the statistical tests of dependence variables which should reject or retain the null hypothesis: whether the distribution of the particular dependent variable which demonstrates the effect in the impact period of 24 months base is the same across all categories of treated/non-treated jobseekers. The statistical tests were carried out at a 95 % confidence level. Quite simply, yellow cells represent the statement that differences between treated and non-treated samples are significant for the particular dependent variable and reference period.

There are obvious significant differences between the treated and non-treated across all the reference periods in placement on the open labour market, namely in full-time jobs and in the achieved average assessment base in the database of SIA.

In the three final reference periods there were identified significant differences between treated and controls in placement in part-time jobs: individual jobseekers that completed traineeship were working in part-time jobs at a higher frequency than their controls, even placement in part-time jobs was quite rare.

Summary test hypothesis		Ref. Period 1		Ref. Period 2		Ref. Period 3		Ref. Period 4	
Null Hypothesis	Test	Sig.	Decision	Sig.	Decision	Sig.	Decision	Sig.	Decision
The distribution of self-employment is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,53	Retain the null hypothesis.	0	Reject the null hypothesis.	0,97	Retain the null hypothesis.	1	Retain the null hypothesis.
The distribution of part-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1	Retain the null hypothesis.	0	Reject the null hypothesis.	0	Reject the null hypothesis.	0	Reject the null hypothesis.
The distribution of full-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0	Reject the null hypothesis.	0	Reject the null hypothesis.	0	Reject the null hypothesis.	0	Reject the null hypothesis.
The distribution of barrier for entrance to LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,18	Retain the null hypothesis.	0	Reject the null hypothesis.	0	Reject the null hypothesis.	0,07	Retain the null hypothesis.
The distribution of placed on the LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0	Reject the null hypothesis.	0	Reject the null hypothesis.	0	Reject the null hypothesis.	0	Reject the null hypothesis.
The distribution of Average assessment base is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,02	Reject the null hypothesis.	0	Reject the null hypothesis.	0	Reject the null hypothesis.	0	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

### Cost-benefit analysis

In the next table there are presented the financial effects on the national budget, which should be representative according to the results learnt from the Post-only non-equivalent comparison design. All the numbers in the table are counted per jobseeker which was treated and non-treated for the particular reference period, which was set for traineeship. There are also presented the net effects of the intervention against the reference periods. The whole cost benefit analysis proceeded according to the methodology introduced in the previous chapter.

The next table contains items which are fundamental at the moment of possibly measured influences or flows on the national budget. Every item is divided into the situation when the treated or non-treated jobseeker is employed. Only the items “grant” and Social Insurance do not distinguish between employed and non-employed statuses because the grant was paid only to the treated individuals when they were unemployed. Social

insurance did not pay when a jobseeker was unemployed according to the evidence from the PES office.

Net effect	Reference period	Treated				Non-treated				Difference between treated and non-treated			
		1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012
Average time share on open market/labour market		56%	50%	45%	41%	39%	38%	57%	53%	17%	12%	-11%	-12%
Average time share out of open market/labour market		44%	50%	55%	59%	61%	62%	43%	47%	-17%	-12%	11%	12%
Unemployment allowance	employed	2 269 €	2 962 €	3 043 €	2 451 €	1 058 €	2 344 €	3 270 €	3 055 €	1 211 €	618 €	-228 €	-604 €
	unemployed	-1 814 €	-2 978 €	-3 704 €	-3 577 €	-1 656 €	-3 771 €	-2 510 €	-2 763 €	-147 €	793 €	-1 194 €	-834 €
Benefit in material need	employed	1 606 €	1 416 €	1 281 €	1 154 €	1 032 €	1 088 €	856 €	1 491 €	574 €	328 €	424 €	-336 €
	unemployed	-1 284 €	-1 423 €	-1 559 €	-1 685 €	-1 615 €	-1 751 €	-657 €	-1 348 €	331 €	328 €	-901 €	-336 €
Grant		-348 €	-1 082 €	-1 112 €	-1 121 €	0 €	0 €	0 €	0 €	-348 €	-1 082 €	-1 112 €	-1 121 €
Health insurance	employed	802 €	828 €	760 €	719 €	566 €	622 €	948 €	865 €	235 €	205 €	-188 €	-146 €
	unemployed	-297 €	-335 €	-367 €	-396 €	-408 €	-412 €	-290 €	-317 €	111 €	77 €	-77 €	-79 €
Social insurance		1 940 €	2 003 €	1 839 €	1 739 €	1 370 €	1 506 €	2 294 €	2 093 €	570 €	497 €	-455 €	-354 €
Taxes from consumption	employed	816 €	842 €	773 €	732 €	576 €	633 €	965 €	880 €	240 €	209 €	-191 €	-149 €
	unemployed	-394 €	-423 €	-414 €	-558 €	-666 €	-482 €	-383 €	-402 €	272 €	59 €	-31 €	-115 €
Income tax	employed	298 €	307 €	282 €	267 €	210 €	231 €	352 €	321 €	87 €	76 €	-70 €	-54 €
	unemployed	-238 €	-309 €	-344 €	-390 €	-329 €	-372 €	-270 €	-291 €	91 €	63 €	-73 €	-99 €
Total / Difference		3 357 €	1 808 €	479 €	-665 €	140 €	-363 €	4 576 €	3 583 €	3 217 €	2 171 €	-4 096 €	-4 248 €

From the results presented above, a positive impact on the state budget was estimated in the first three reference periods. For instance, in the first reference period the estimated effectiveness ratio shows that one invested Euro to the jobseeker returned 10 Euros over the 2 year long impact period. That extremely positive effectiveness is based on the very limited grants which were provided to the treated jobseekers (on average about 60 Euros per month). In the next update the value of the grant increased based on the living wage stated for that particular year. That is why the grant increased on average 3 times. The final row describes total flows which were on average produced by one treated or non-treated jobseeker and the differences between these groups. In the first three reference periods the treated should have created, on average, positive flows into the state budget. They returned to the state the grant that was invested to them and also they produced on average some extra money over the grant. However, in the last period the treated jobseekers were not able to repay the grant and part of the unemployment allowance. That is the reason why the last reference period is in the red.

Non-treated jobseekers were mostly successful in the last two reference periods, where they produced for the state budget up to 4600 Euros. As can be seen in the last columns in the first two months there are positive net effects of the intervention, i.e. the treated produced more money for the state budget than the non-treated - up to 3200 Euros per one treated. In the last reference periods the situation changed and the non-treated were less successful in placement on the open labour market, as well as the grant was increased and that is the reason why the treated jobseekers were much more difficult to get into the green numbers.

#### 4.8.3 Exact matching with the application of Post-only non-equivalent comparison design

This method is based on the creation of pairs of treated and non-treated jobseekers which have the same characteristics of independent variables:

- reference period,
- gender,
- age,

- marital status,
- permanent residence,
- level of education (5 categories),
- unemployed before 2007 (4 categories) ,
- driving licence: cars and motorcycles,
- driving licence: vans and trucks,
- driving licence: bus,
- driving licence: trucks.

After matching the individuals from both samples, the impact of the intervention was estimated through subtraction of the individual dependent variables of the treated and non-treated. We measured 6 types of dependent variables which should estimate the financial status of the individual and employability in the impact period of 24 months:

- 1) placed on the labour market, which is a total of registrations for full-time work and self-employment
- 2) individual barrier for entrance to LM,
- 3) part-time job,
- 4) full-time job,
- 5) self-employed,
- 6) average assessment base in Euros.

The first five variables are measured as shares of the particular type of registration in SIA during 24 months of the impact period. The coefficient was designed because it will be needed to provide a comparison of results estimated based on the different types of carried out methods.

In total, more than 100 thousand individuals for exact matching in all set reference periods were used. Every treated jobseeker was matched to individuals from controls, which should help in estimating the net effect of traineeship in different reference periods. For instance, in the first reference period there were established 359 pairs, where 702 treated individuals were matched to 2439 non-treated individuals.

Case Processing Summary			
Reference period		Cases	
		Valid	
		N	Percent
1	non treated	702	100,0%
	treated	2439	100,0%
2	non treated	31800	100,0%
	treated	20038	100,0%
3	non treated	6258	100,0%
	treated	11266	100,0%
4	non treated	17220	100,0%
	treated	14110	100,0%
In total	non treated	55980	100,0%
	treated	47853	100,0%

## Measuring of employability

Group of sample	Share of sustained time on LM on the time of impact period	Reference period 1			Reference period 2			Reference period 3			Reference period 4			Total estimated net effect		
		Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent
non treated	0	396	56,4	56,4	15925	50,1	50,1	2458	39,3	39,3	7353	42,7	42,7	26132	46,7	46,7
	0,1	43	6,1	43,6	1745	5,5	49,9	378	6,0	60,7	729	4,2	57,3	2895	5,2	53,3
	0,2	25	3,6	40,0	1529	4,8	45,1	359	5,7	55,0	682	4,0	53,3	2595	4,6	48,7
	0,3	27	3,8	36,2	2012	6,3	38,8	588	9,4	45,6	1069	6,2	47,1	3696	6,6	42,1
	0,4	20	2,8	33,3	1343	4,2	34,6	368	5,9	39,7	706	4,1	43,0	2437	4,4	37,7
	0,5	41	5,8	27,5	1769	5,6	29,0	490	7,8	31,9	1205	7,0	36,0	3505	6,3	31,5
	0,6	23	3,3	24,2	1174	3,7	25,3	267	4,3	27,6	829	4,8	31,2	2293	4,1	27,4
	0,7	15	2,1	22,1	1192	3,7	21,6	263	4,2	23,4	770	4,5	26,7	2240	4,0	23,4
	0,8	40	5,7	16,4	1864	5,9	15,7	400	6,4	17,0	1083	6,3	20,5	3387	6,1	17,3
	0,9	29	4,1	12,3	1128	3,5	12,2	240	3,8	13,2	808	4,7	15,8	2205	3,9	13,4
	1	43	6,1	6,1	2119	6,7	5,5	447	7,1	6,0	1986	11,5	4,2	4595	8,2	5,2
Total	702	100		31800	100		6258	100		17220	100		55980	100		
treated	0	325	13,3	13,3	4031	20,1	20,1	2423	21,5	21,5	3399	24,1	24,1	10178	21,3	21,3
	0,1	182	7,5	86,7	1540	7,7	79,9	856	7,6	78,5	1010	7,2	75,9	3588	7,5	78,7
	0,2	125	5,1	81,5	1296	6,5	73,4	753	6,7	71,8	853	6,0	69,9	3027	6,3	72,4
	0,3	188	7,7	73,8	1934	9,7	63,8	1180	10,5	61,3	1372	9,7	60,1	4674	9,8	62,6
	0,4	128	5,2	68,6	1078	5,4	58,4	665	5,9	55,4	846	6,0	54,1	2717	5,7	57,0
	0,5	221	9,1	59,5	1653	8,2	50,1	1029	9,1	46,3	1344	9,5	44,6	4247	8,9	48,1
	0,6	138	5,7	53,9	1174	5,9	44,3	741	6,6	39,7	995	7,1	37,6	3048	6,4	41,7
	0,7	161	6,6	47,3	1129	5,6	38,6	630	5,6	34,1	931	6,6	31,0	2851	6,0	35,8
	0,8	265	10,9	36,4	1951	9,7	28,9	1000	8,9	25,3	1376	9,8	21,2	4592	9,6	26,2
	0,9	224	9,2	27,2	1552	7,7	21,2	708	6,3	19,0	796	5,6	15,6	3280	6,9	19,3
	1	482	19,8	7,5	2700	13,5	7,7	1281	11,4	7,6	1188	8,4	7,2	5651	11,8	7,5
Total	2439	100		20038	100		11266	100		14110	100		47853	100		

In total we estimated the net impact with a sample of almost 56 thousand non-treated individuals and almost 48 thousand treated jobseekers that were matched according to the same values of the independent variables.

The table above presents the distribution of the samples treated and controls in different reference periods across a share of sustainability on the labour market within the impact period of 24 months. From up to down there are presented non-treated groups, treated groups and differences across set reference periods. The highest intensity in the group of controls was identified as being the jobseekers that did not find any job during the whole impact period. The frequency of non-employed jobseekers depended on the reference period. The highest numbers are in the first two periods - more than 50 % of samples. The other shares of sustaining time on the total impact period (i.e. 24 months) are equally distributed among the shares.

The sample of treated jobseekers shows that the highest frequencies are in the interval with zero share of time spent on the LM or all 2 years employed on the labour market. On average, there is almost 36 % probability that the treated jobseeker sustained employment more than 17 months after he finished intervention, while in the group of controls it is just less than 23 % probability that the same jobseeker remained 17 months placed on the

Share of sustained time on LM on the time of impact period	Net effect				Total estimated net effect
	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	
0	-43,1	-30,0	-17,8	-18,6	-25,4
0,1	1,3	2,2	1,6	2,9	2,3
0,2	1,6	1,7	0,9	2,1	1,7
0,3	3,9	3,3	1,1	3,5	3,2
0,4	2,4	1,2	0,0	1,9	1,3
0,5	3,2	2,7	1,3	2,5	2,6
0,6	2,4	2,2	2,3	2,2	2,3
0,7	4,5	1,9	1,4	2,1	2,0
0,8	5,2	3,9	2,5	3,5	3,5
0,9	5,1	4,2	2,4	0,9	2,9
1	13,6	6,8	4,2	-3,1	3,6

labour market.

The yellow histogram in the columns “cumulative percent” should help to picture the speed of decreasing probability of individuals staying on the labour market. While on average almost 80% of the treated found a job for just 2 and half months of the impact period, more than 50 % of the controls had the same performance in the same impact period. To compare numbers among set reference periods of the treated shows the tendency of worsening of employability and vice versa - the non-treated had the reverse trend of employability improving.

In the next table is presented the net effect on employability of treatments by the traineeship. The numbers are the results of the subtraction of treated and non-treated. That is why the first red bar chart shows the decreasing probability of the treated that they will not find a job during the whole impact period. The result shows that, in total, the average in all reference periods is about 25% higher probability that a jobseeker that is treated will not get a place on the open labour market at all within two years after the intervention finished.

### Types of registrations in SIA

In the table below it is possible to see the types of registration in the SIA. There are outlined 5 basic types of registration. The treated and non-treated that were placed on the labour market were placed in full-time jobs. In the table it is obvious that, generally, the unemployed graduates didn't have any interest in establishing a business and becoming self-employed. Also, there were found minimum individuals registered in part-time jobs in the database of SIA, and on average for about 2% of the time of the impact period, the treated and non-treated found individual barriers for entrance to LM. These individuals received accident benefits, care allowance or were temporary personal assistants.

The assessment base reveals treated and non-treated individuals that were even once per period placed on the open labour market and who achieved on average a monthly assessment base higher than the stated minimal wage, which is on average about 100 Euros.

Effect of the intervention	Group of sample	Ref. Period 1	Ref. Period 2	Ref. Period 3	Ref. Period 4	total
		Mean	Mean	Mean	Mean	Mean
Self-employment	non treated	0,4%	0,8%	0,3%	0,2%	0,6%
	treated	0,7%	0,4%	0,3%	0,2%	0,3%
Full-time job	non treated	20%	24%	27%	29%	26%
	treated	52%	42%	38%	34%	39%
Individual barrier for entrance to LM	non treated	4%	2%	1%	2%	2%
	treated	2%	3%	2%	2%	2%
Part-time job	non treated	0,0%	1%	3%	4%	2%
	treated	0,0%	1%	3%	6%	3%
Placed on LM	non treated	21%	25%	27%	29%	26%
	treated	53%	43%	38%	34%	40%
Average assessment base	non treated	433	504	510	522	510
	treated	432	488	501	496	490
Self-employment	Difference/ estimated net effect	0%	0%	0%	0%	0%
Full-time job		32%	18%	11%	5%	13%
Individual barrier for entrance to LM		-1%	1%	1%	0%	1%
Part-time job		0%	1%	0%	2%	1%
Placed on LM		32%	18%	11%	5%	13%
Average assement base		-1	-16	-9	-26	-20

To analyse the net effects of the intervention through the methodology it is obvious from the bottom part of the table that treated graduates seem to be, with a higher probability,

placed on the labour market longer by about 13% in total effect across the reference periods. The treated sustained employment for about 32% of the impact period, longer in comparison with the controls in the first reference period. In the table there is also visible a trend of decreasing efficiency over time.

It was estimated a negative effect on the assessment base in the impact period, because according to the values it is reasonable to assume that if the intervention had not been granted the graduates would have achieved a higher assessment base, on average about 20 Euros per month.

Null Hypothesis	Test	Reference period 1		Reference period 2		Reference period 3		Reference period 4	
		Sig.	Decision	Sig.	Decision	Sig.	Decision	Sig.	Decision
The distribution of self-employment is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	0,744	Retain the null hypothesis.	1,000	Retain the null hypothesis.	1,000	Retain the null hypothesis.
The distribution of full-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of barrier for on the LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,994	Retain the null hypothesis.	0,001	Reject the null hypothesis.	0,744	Retain the null hypothesis.	0,543	Retain the null hypothesis.
The distribution of part-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	0,000	Reject the null hypothesis.	0,935	Retain the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of placed on LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Average assement base is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,419	Retain the null hypothesis.	0,000	Reject the null hypothesis.	0,004	Reject the null hypothesis.	0,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

In the table above are presented the results of the carried out Kolmogorov-Smirnov tests of variables, which should reject or retain a null hypothesis: whether it is the distribution of the particular dependent variable which demonstrates the effect in the impact period of 24 months base, the same across categories of treated/non-treated jobseekers. The tests were carried out at 95% confidence level. The yellow cells represent variables in particular reference periods where the treated and non-treated differ from each other significantly.

Significant differences were estimated namely in the full-time job registrations, which is the reason why also in the dependent variable “placement on LM” significant differences were counted between treated and controls in the assessment base they achieved in the last three reference periods .

### Cost-benefit analysis

The table below presents the estimated values of the financial effects of the intervention outcomes, which have been applied to the results on the basis of the exact matching method. The cost-benefit analysis is a kind of financial statement summing up items with a positive and a negative influence on public finance. The table is divided into reference periods of treated and non-treated with a final counting of the difference between these two groups. In the vertical distribution of the table, presented in the first two lines, is the average effect on employability and in the next lines are presented the different types of the items which affect the public budget. All the values are counted for the impact period of 2 years after realisation of the intervention.

In the cost-benefit analysis there were taken into account real benefits and costs as well as costs for lost opportunities and benefits from savings. The cost-benefit analysis was carried out with items outlined in the first column of the table below.

The insured person is entitled to unemployment benefit if, in the four years before registering as unemployed jobseekers (hereinafter referred to as "registered

unemployed"), he/she was covered by unemployment insurance for at least three years<sup>5</sup>. All the values presented in the table below are estimated on one treated individual

Net effect	Reference period	Treated				Non-treated				Difference between treated and non-treated			
		1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012
Average time share on open market/labour market		53%	43%	38%	34%	21%	25%	27%	29%	32%	18%	11%	5%
Average time share out of open market/labour market		47%	57%	62%	66%	79%	75%	73%	71%	-32%	-18%	-11%	-5%
Unemployment allowance	employed	2 155 €	2 534 €	2 589 €	2 041 €	562 €	1 509 €	1 578 €	1 697 €	1 593 €	1 025 €	1 011 €	344 €
	unemployed	-1 928 €	-3 407 €	-4 157 €	-3 987 €	-2 152 €	-4 606 €	-4 202 €	-4 121 €	224 €	1 200 €	45 €	134 €
Benefit in material need	employed	1 525 €	1 211 €	1 090 €	961 €	548 €	701 €	413 €	828 €	977 €	510 €	676 €	133 €
	unemployed	-1 364 €	-1 628 €	-1 749 €	-1 878 €	-2 098 €	-2 139 €	-1 100 €	-2 011 €	734 €	510 €	-649 €	133 €
Grant		-348 €	-1 082 €	-1 112 €	-1 121 €	0 €	0 €	0 €	0 €	-348 €	-1 082 €	-1 112 €	-1 121 €
Health insurance	employed	819 €	754 €	685 €	575 €	310 €	457 €	492 €	542 €	510 €	297 €	193 €	33 €
	unemployed	-315 €	-383 €	-412 €	-442 €	-530 €	-503 €	-486 €	-473 €	214 €	120 €	74 €	31 €
Social insurance		1 983 €	1 825 €	1 657 €	1 392 €	749 €	1 106 €	1 190 €	1 313 €	1 234 €	719 €	467 €	79 €
Taxes from consumption	employed	834 €	768 €	697 €	585 €	315 €	465 €	501 €	552 €	519 €	303 €	196 €	33 €
	unemployed	-471 €	-547 €	-526 €	-575 €	-899 €	-763 €	-734 €	-754 €	428 €	217 €	208 €	178 €
Income tax	employed	304 €	280 €	254 €	214 €	115 €	170 €	183 €	201 €	189 €	110 €	72 €	12 €
	unemployed	-272 €	-377 €	-408 €	-417 €	-440 €	-518 €	-486 €	-489 €	168 €	111 €	78 €	72 €
Total / Difference		2 923 €	-50 €	-1 393 €	-2 653 €	3 519 €	-4 122 €	-2 652 €	-2 715 €	6 442 €	4 072 €	1 258 €	63 €

Nevertheless, while the most important results are presented in the last green line in the table above, it is important to take a look at the values in the last four columns of the table. Naturally, the most negative influenced item of cost-benefit analysis which decreases the net financial effect of the intervention is the grant, namely in the last three reference periods.

The first part of the columns takes into account only the treated samples in the reference periods. From the result in the last row, it is estimated that one treated brought in, in just the first reference period, positive cash flows of almost three thousand Euros. This is just one positive result which influences the state budget due mainly to the lowest average amount of grant – financial allowance provided during traineeship for a maximum of 6 months. The other reason why this is just one positive result, is the high average level of the time placed on the labour market. In the whole line of reference periods, the difference between treated and non-treated was estimated in the range of 63 Euros to 6,442 Euros per jobseeker. This means one treated can generate from 63 to 6,500 Euros more cash for the state budget in a 2-year long impact period after completion of the traineeship than the same non-treated jobseeker.

#### 4.8.4 Propensity score exact matching

This method of approach is composed of:

- estimation of a logistics model with its application on individuals on the samples of treated and control individuals,
- matching only those individuals which have the same value of propensity score,
- individual non-treated adopted individual impact periods of the treated individual who was matched with the non-treated,
- enforcement of post-only comparison design,

Case Processing Summary					
Ref. period	Type of sample	Cases			
		Valid		Missing	
		N	Percent	N	Percent
1	non treated	249	100,0%	0	0,0%
	treated	762	100,0%	0	0,0%
2	non treated	12303	100,0%	0	0,0%
	treated	6335	100,0%	0	0,0%
3	non treated	823	100,0%	0	0,0%
	treated	3724	100,0%	0	0,0%
4	non treated	4486	100,0%	0	0,0%
	treated	3599	100,0%	0	0,0%
In total		32281			

- statistical tests between treated and non-treated results of dependent variables.

The logistic model was estimated using all independent variables that were measured for the participants and non-participants. We used the following independent variables:

- 1) Gender
- 2) Age
- 3) Marital status – used as a categorical variable
- 4) Level of education\_10 categories
- 5) Level of education\_5 categories – used as a categorical variable
- 6) Types of disadvantages
- 7) Unemployed in months
- 8) Total period of all registrations in months (COLSaF)
- 9) Unemployed before 2007 in months – used as a categorical variable
- 10) The average gross wage in the region of permanent residence
- 11) The proportion of women in the district of permanent residence
- 12) Surface area of district of permanent residence
- 13) The density of population in the district of permanent residence
- 14) The number of municipalities in the district of permanent residence
- 15) The number of cities in the district of permanent residence
- 16) The registered unemployment rate in the district of permanent residence
- 17) Inhabitants density
- 18) Population of municipality in 2011
- 19) Change of population: 15 years
- 20) Distance from PES office
- 21) District of permanent residence
- 22) Region of permanent residence
- 23) Driving licence: cars and motorcycles, buses, trucks, vans

Categorical Variables Codings	
Level of education_5 categories	primary school
	secondary vocational school
	vocational school
	comprehensive school
	colledge
Marital status	registered partners
	divorced
	single
	widow
Unemployed before 2007 in months	married
	< 1 year
	1 - 3 years
	> 3 years
	no evidence

The dependent variable in logistic regression was the variable *Treated / non-treated*, with values 1 for participants and 0 for non-participants.

In the logistic regression procedure we used the Backward conditional stepwise method, with the condition of entry probability 0.01 and removal probability 0.05. Three variables were used as categorical variables with categories defined in the table above.

Using the Backward conditional method we get the final best logistic regression for modelling the probability (or odds) of participating in the programme with the given independent variables. This model was created separately for every reference period. In the following tables there are the results of the final logistic models.



In the first reference period the independent variables *Age*, *Education level*, *Unemployed in month*, *Total period of all registrations*, *Unemployed before 2007* (categorical), *Marital status* (categorical) and *School* (categorical), are significant with a significance level of

Variables in the Equation, period 1						
	B	S.E.	Wald	df	Sig.	Exp(B)
Age (rounded)	-1,012	,027	1385,019	1	,000	,363
Education level	1,313	,113	134,792	1	,000	3,716
Unemployed in months	,158	,009	287,423	1	,000	1,171
Total period of all registrations in months (colsaf)	-,032	,003	121,789	1	,000	,969
Unemployed before 2007 in months_category_1	,599	,114	27,830	1	,000	1,821
Unemployed before 2007 in months_category_2	1,231	,142	75,033	1	,000	3,424
Unemployed before 2007 in months_category_3	3,166	,348	82,829	1	,000	23,719
Marital_status_category_1	1,774	,480	13,628	1	,000	5,892
Marital_status_category_2	-,385	,144	7,162	1	,007	,681
School_category_1	3,184	,846	14,163	1	,000	24,139
School_category_2	2,981	,485	37,832	1	,000	19,702
School_category_3	3,086	,517	35,593	1	,000	21,896
School_category_4	1,782	,390	20,874	1	,000	5,940
Constant	13,361	1,235	116,991	1	,000	634458,661

0.05. The values of coefficient B and Exp(B) we can interpret in the following way: if the value of *Age* increases by 1 year, that means the participant or non-participant will be 1 year older, and all other variables stay the same, and the unit will be 0.363 times more likely to belong to the treated units. The odds of *Age* are smaller than 1, so increasing *Age* changes the probability that the unit is treated 0.363 times (e.g. decreases the probability). For education level the influence is stronger. With a higher education level of 1 degree the probability of being treated will increase 3,716 times, if the values of all other variables stay the same. Similarly for *Unemployed in months* and *Total period of all registrations*. If the value of odds Exp(B) is greater than 1, increasing that variable by 1 will increase the probability of being treated. If the value of Exp(B) is smaller than 1 (as for *Total period of all registrations*), increasing this period by 1 month will change the probability of being treated by 0.969 times (decrease it).

Variables in the Equation, period 2						
	B	S.E.	Wald	df	Sig.	Exp(B)
Gender	,319	,039	66,881	1	,000	1,375
Age (rounded)	-1,168	,012	8910,415	1	0,000	,311
Education level	1,481	,048	946,876	1	,000	4,398
School_category_1	2,588	,369	49,203	1	,000	13,307
School_category_2	2,137	,195	119,501	1	,000	8,473
School_category_3	2,114	,206	104,883	1	,000	8,281
School_category_4	1,104	,158	48,795	1	,000	3,016
Disadvantages	,265	,037	51,625	1	,000	1,304
Unemployed in months	,156	,003	2682,115	1	0,000	1,168
Total period of all registrations in months (colsaf)	-,023	,002	180,901	1	,000	,977
Unemployed before 2007 in months_category_1	1,668	,081	420,584	1	,000	5,301
Unemployed before 2007 in months_category_2	1,467	,123	141,891	1	,000	4,335
Unemployed before 2007 in months_category_3	1,184	,393	9,092	1	,003	3,267
The average gross wage in the region of perm. residence	-,001	,000	9,690	1	,002	,999
Surface of district of permanent residence	,000	,000	15,195	1	,000	1,000
The density of population in the district of perm. residence	,000	,000	17,450	1	,000	1,000
The number of cities in the district of perm. residence	-,078	,024	10,650	1	,001	,925
The registered unemployment rate in the district of perm. residence	,016	,004	19,700	1	,000	1,016
Inhabitants density	,000	,000	6,037	1	,014	1,000
Distance_from_PESoffice	-,012	,002	31,615	1	,000	,988
Constant	15,041	,538	781,794	1	,000	3405655,736

Variables in the Equation, period 3						
	B	S.E.	Wald	df	Sig.	Exp(B)
Gender	,726	,063	130,887	1	,000	2,066
Age (rounded)	-,956	,022	1950,409	1	0,000	,384
Education level	1,132	,028	1631,454	1	0,000	3,100
School	1,042	,134	60,458	1	,000	2,835
Disadvantages	,193	,060	10,456	1	,001	1,213
Unemployed in months	,289	,007	1632,648	1	0,000	1,335
Total period of all registrations in months (colsaf)	-,010	,003	11,861	1	,001	,990
Unemployed before 2007 in months_category EVIDENCIE PRED 2007			88,940	2	,000	
Unemployed before 2007 in months_category1	1,987	,226	77,597	1	,000	7,291
Unemployed before 2007 in months_category_2	1,450	,395	13,465	1	,000	4,262
The density of population in the district of perm. residence	,000	,000	6,974	1	,008	1,000
The registered unemployment rate in the district of perm. residence	,027	,005	27,331	1	,000	1,027
Inhabitantsdensity	,000	,000	13,190	1	,000	1,000
Population_of_municipality_2011	,000	,000	11,378	1	,001	1,000
Distance_from_PESoffice cars	-,011	,004	7,950	1	,005	,990
	,168	,073	5,282	1	,022	1,183
Constant	8,990	,593	229,611	1	,000	8020,558

For categorical variables such as *School*, the value of odds Exp(B) can be interpreted for every category compared to the reference category as follows: for example, for *School* category 1 (primary school) the value 24,139 means that the units with primary school have 24,139 times higher probability of belonging to the group of

treated individuals than the reference category (the last/highest category of school, which is College).

Another categorical variable, *Unemployed before 2007*, for example category 3 (registered more than 3 years), has 23,719 times higher probability of belonging to the group of treated units than the reference category (that is *No evidence before 2007*).

In reference period 2, the following variables are significant in logistic regression with a significance level of 0.05. The variables *Age*, *Total period of all registrations*, *Average gross wage in the region*, *Number of cities in the region* and *Distance from PES office* decrease the probability of being treated, because their odds  $\text{Exp}(B)$  are smaller than 1; although only *Age* has odds significantly smaller than 1, the others are very close to 1 (therefore do not change the probability of being treated very much). Other variables have the value of odds  $\text{Exp}(B)$  greater than 1, so in case they change by 1 and the other variables stay the same, the probability of being the unit in the treatment group will increase  $\text{Exp}(B)$ -times. For categorical variables *School* and *Unemployed before 2007* the odds  $\text{Exp}(B)$  means a change in multiples of the probability of every category compared to the reference category. All categories of these variables increase the probability of classification of the unit into the group of treated individuals.

In the third reference period the situation is similar. Only *Age* has the odds  $\text{Exp}(B)$  significantly smaller than 1, so a change in *Age* of 1 decreases the probability of being in the treatment group. Some variables have odds very close to 1, so we can say that a change in these does not change the probability of being in the treatment group. All other variables, categorical too, have odds more than 1, so with a change of 1 the probability of being treated increases in multiples of  $\text{Exp}(B)$ .

For the last reference period *Age*, *Marital status* category 2 (divorced), *Driving licence* category Vans and category Motorcycles have odds less than 1, so a change of 1 of these variables will cause a decrease in the probability that the unit will be in the treatment group multiple by the  $\text{Exp}(B)$ . In the case of *Marital status* it means the ratio of probability of being treated compared to the reference category. All other variables have odds more than 1, so with a change of them of 1, the probability that the unit belongs to the treatment group increases by multiples of  $\text{Exp}(B)$ .

The coefficient of all variables is significant with a significance level of 0.05 in the Wald test for logistic regression coefficients. All created

models have very good classification ability. The percentage correctly classified is always more than 90%; the Nagelkerke R-square (alternative for R-square in linear regression) is, in all 4 models, more than 0.80.

Variables in the Equation, period 4						
	B	S.E.	Wald	df	Sig.	Exp(B)
Gender	,561	,046	150,048	1	,000	1,752
Age (rounded)	-,868	,016	3085,285	1	0,000	,420
Marital status_category			12,433	2	,002	
Marital status_category(1)	1,297	,568	5,221	1	,022	3,659
Marital status_category(2)	-,233	,102	5,271	1	,022	,792
Education level	1,472	,059	612,059	1	,000	4,357
School			114,603	4	,000	
School_category_1	1,568	,449	12,218	1	,000	4,797
School_category_2	1,653	,244	45,914	1	,000	5,225
School_category_3	1,665	,256	42,181	1	,000	5,287
School_category_4	,733	,198	13,752	1	,000	2,082
Disadvantages	,450	,031	212,018	1	,000	1,568
Unemployed in months	,197	,004	2266,850	1	0,000	1,218
Total period of all registrations in months (colsaf)	,010	,002	18,166	1	,000	1,010
Unemployed before 2007 in months			135,291	2	,000	
Unemployed before 2007 in months_category_(1)	2,038	,192	112,106	1	,000	7,673
Unemployed before 2007 in months_category_(2)	2,140	,419	26,075	1	,000	8,497
The average gross wage in the region of perm. residence	-,001	,000	19,413	1	,000	,999
The proportion of women in the district of perm. residence	-,074	,035	4,375	1	,036	,929
Surface of district of permanent residence	,000	,000	12,940	1	,000	1,000
The number of cities in the district of perm. residence	-,076	,028	7,563	1	,006	,927
Population_of_municipality_2011	,000	,000	15,783	1	,000	1,000
Distance_from_PESoffice	-,010	,003	12,950	1	,000	,990
motorcycles	-,103	,050	4,202	1	,040	,902
small_trucks	-,443	,163	7,385	1	,007	,642
Constant	11,405	1,810	39,723	1	,000	89745,687

## *Measuring employability*

As was used in the previous method, we divided samples of the treated and non-treated into four reference periods. The smallest samples were matched in the first reference period. And, on the other side, the biggest samples of treated and non-treated were matched in the second reference periods.

In total, the results of the method were estimated according to more than 32 thousand individuals from the treated and non-treated samples, which is about one third of the available total sample. Exact matching is a significant rule which directly influences the shrinking of available samples before matching.

As already described twice before, the next table provides a view of the frequencies table of the most important dependent variable which has a role in bringing the view on employability after the intervention finishes after the two years impact period of treated and control groups.

In the lines there are presented the shares of the time of the impact period when individuals were sustained on the labour market from 0 to 100 % in a coefficient from 0 up to 1. The table is again divided into five parts; the first four describe employability in the particular impact periods of set reference periods and the fifth part informs about the average effect without taking into consideration the homogeneity of the intervention.

As presented in the table, in all four reference periods there are high frequencies of non-treated jobseekers that were not employed during the entire impact period. This fact will increase the net-effect of the intervention. On average, more than 60 % of the non-treated remained unemployed throughout the entire two years of the impact period. From the shape of the yellow bar chart, particularly from the sharpness and bluntness of the shape which is created from the bar chart, it is possible to deduce that if a non-treated placed was placed once on the labour market, they remained there for a longer time with higher probability. For instance, in the second reference period, about 40 % of controls found a job at least for 10 % of the impact period, but from those individuals there remained about every second one employed for the entire impact period. While almost 84 % of treated identical graduates placed on the labour market for at least for 10 % of the whole impact period (2 months), from those remained employed only every 7th treated individual.

Group of sample	Share of impact period sustained on LM	Reference period 1			Reference period 2			Reference period 3			Reference period 4			Ref. Periods in total		
		Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent
non treated	0	199	79,9	79,9	7404	60,2	60,2	475	57,7	57,7	2893	64,5	64,5	10971	61,4	61,4
	0,1	10	4,0	20,1	549	4,5	39,8	48	5,8	42,3	222	4,9	35,5	829	4,6	38,6
	0,2	8	3,2	16,1	443	3,6	35,4	44	5,3	36,5	160	3,6	30,6	655	3,7	33,9
	0,3	5	2,0	12,9	553	4,5	31,8	52	6,3	31,1	192	4,3	27,0	802	4,5	30,3
	0,4	2	0,8	10,8	267	2,2	27,3	15	1,8	24,8	84	1,9	22,7	368	2,1	25,8
	0,5	4	1,6	10,0	317	2,6	25,1	29	3,5	23,0	112	2,5	20,8	462	2,6	23,7
	0,6	2	0,8	8,4	194	1,6	22,5	17	2,1	19,4	49	1,1	18,3	262	1,5	21,1
	0,7	3	1,2	7,6	177	1,4	20,9	9	1,1	17,4	40	0,9	17,3	229	1,3	19,7
	0,8	5	2,0	6,4	198	1,6	19,5	15	1,8	16,3	66	1,5	16,4	284	1,6	18,4
	0,9	0	0,0	4,4	112	0,9	17,9	3	0,4	14,5	87	1,9	14,9	202	1,1	16,8
	1	11	4,4	4,4	2089	17,0	17,0	116	14,1	14,1	581	13,0	13,0	2797	15,7	15,7
Total	249	100		12303	100		823	100		4486	100		17861	100		
treated	0	167	21,9	21,9	1022	16,1	16,1	1018	27,3	27,3	992	27,6	27,6	3199	22,2	22,2
	0,1	54	7,1	78,1	556	8,8	83,9	399	10,7	72,7	276	7,7	72,4	1285	8,9	77,8
	0,2	39	5,1	71,0	498	7,9	75,1	313	8,4	61,9	250	6,9	64,8	1100	7,6	68,9
	0,3	49	6,4	65,9	656	10,4	67,2	465	12,5	53,5	370	10,3	57,8	1540	10,7	61,3
	0,4	41	5,4	59,4	367	5,8	56,9	242	6,5	41,1	225	6,3	47,5	875	6,1	50,6
	0,5	70	9,2	54,1	582	9,2	51,1	339	9,1	34,6	317	8,8	41,3	1308	9,1	44,5
	0,6	43	5,6	44,9	391	6,2	41,9	221	5,9	25,5	226	6,3	32,5	881	6,1	35,5
	0,7	45	5,9	39,2	365	5,8	35,7	175	4,7	19,5	176	4,9	26,2	761	5,3	29,3
	0,8	73	9,6	33,3	608	9,6	30,0	216	5,8	14,8	281	7,8	21,3	1178	8,2	24,1
	0,9	72	9,4	23,8	482	7,6	20,4	132	3,5	9,0	191	5,3	13,5	877	6,1	15,9
	1	109	14,3	14,3	808	12,8	12,8	204	5,5	5,5	295	8,2	8,2	1416	9,8	9,8
Total	762	100		6335	100		3724	100		3599	100		14420	100		

The next table informs us about the types of registration of treated and non-treated on average during the impact periods. There are again five types of registration which have been identified for graduates, infrequent self-employment, full-time job, part-time job, individual barrier for entrance to LM and placement on LM which is a full-time job, or self-employment. A part-time job is not considered as real placement on LM. Another presented independent variable is the average of the assessment base.

In the context of the test results between the results of independent variables of treated and non-treated groups, it is possible to note that treated and control individuals earned less money across all reference periods by about 10 to 20 %.

From the presented results it is again obvious that graduates do not have any significant interest in self-employment.

There are significant differences across all reference periods between the distribution of treated and controls for the independent variables full-time job and placement on the labour market. This means that the presented net effect is confirmed. In the first reference period the treated remained on average on LM for about 37% of the entire impact period longer than the controls. In the second reference period it was about 21% and in the others it was 9% and 17%.

Graduates achieved that effect mostly due to placement in full-time jobs; only in few exceptions did graduates establish self-employment. From the group of non-treated none

Effect of the intervention	Group of sample	Ref. Period 1 Mean	Ref. Period 2 Mean	Ref. Period 3 Mean	Ref. Period 4 Mean
Average assessment base	non treated	461	539	510	564
	treated	422	476	480	482
Self-employment	treated	0,00	0,00	0,00	0,00
	treated	0,01	0,01	0,00	0,00
Full-time job	non treated	0,10	0,25	0,24	0,21
	treated	0,47	0,47	0,33	0,38
Individual barrier for entrance to LM	non treated	0,01	0,03	0,03	0,02
	treated	0,04	0,04	0,03	0,03
Part-time job	non treated	0,00	0,00	0,00	0,00
	treated	0,00	0,02	0,06	0,08
Placed on LM	non treated	0,10	0,26	0,24	0,22
	treated	0,48	0,47	0,34	0,38
Average assessment	net effect	-39	-62	-30	-82
Self-employment		0,01	0,00	0,00	0,00
Full-time job		0,37	0,21	0,09	0,17
Individual barrier for entrance to LM		0,02	0,00	-0,01	0,01
Part-time job		0,00	0,02	0,06	0,08
Placed on LM		0,38	0,21	0,10	0,17

had any interest in self-employment.

Summary test hypothesis		Ref. Period 1		Ref. Period 2		Ref. Period 3		Ref. Period 4	
Null Hypothesis	Test	Sig.	Decision	Sig.	Decision	Sig.	Decision	Sig.	Decision
The distribution of Average assessment base is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,038	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,001	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Self-employment is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,999	Retain the null hypothesis.	0,211	Retain the null hypothesis.	1,000	Retain the null hypothesis.	1,000	Retain the null hypothesis.
The distribution of Full-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Individual barrier for entrance to LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,564	Retain the null hypothesis.	0,000	Reject the null hypothesis.	0,727	Retain the null hypothesis.	0,005	Reject the null hypothesis.
The distribution of Part-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Placed on LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

### Cost-benefit analysis

The next table presents the net effect of the traineeship on the state budget calculated per jobseeker. The net effects in the last four columns are calculated by subtraction of the treated and their controls financial balance based on the average measured success in placement on the open labour market. From the table can be seen that the treated returned on average the grants that were distributed through intervention and they also brought a “net profit” from 1,200 up to 1,800 Euros per 24 months long impact period. That situation was estimated from just the first two reference periods, the other had a negative estimated financial influence on the national budget. All controls have an estimated negative financial influence on the national budget due to their low level of employability. On average, the net effects are very positive because, across all reference periods, the treated generated from 700 Euros up to almost 7,500 Euros per impact period more finance due to employability and paid taxes and saved allowance and benefits.

Net effect	Reference period	Treated				Non-treated				Difference between treated and non-treated			
		1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012
Average time share on open market/labour market		48%	47%	34%	38%	10%	26%	24%	22%	38%	21%	10%	17%
Average time share out of open market/labour market		52%	53%	66%	62%	90%	74%	76%	78%	-38%	-21%	-10%	-17%
Unemployment allowance	employed	1 963 €	2 799 €	2 262 €	2 312 €	274 €	1 570 €	1 383 €	1 254 €	1 689 €	1 129 €	879 €	1 058 €
	unemployed	-2 120 €	-3 142 €	-4 484 €	-3 716 €	-2 441 €	-4 546 €	-4 397 €	-4 563 €	320 €	1 404 €	-87 €	848 €
Benefit in material need	employed	1 389 €	1 338 €	952 €	1 089 €	267 €	729 €	362 €	612 €	1 122 €	609 €	590 €	477 €
	unemployed	-1 501 €	-1 501 €	-1 887 €	-1 750 €	-2 380 €	-2 110 €	-1 151 €	-2 227 €	879 €	609 €	-736 €	477 €
Grant		-348 €	-1 082 €	-1 112 €	-1 121 €	0 €	0 €	0 €	0 €	-348 €	-1 082 €	-1 112 €	-1 121 €
Health insurance	employed	746 €	833 €	598 €	651 €	151 €	475 €	431 €	401 €	595 €	358 €	157 €	250 €
	unemployed	-347 €	-353 €	-444 €	-412 €	-601 €	-497 €	-508 €	-524 €	254 €	113 €	64 €	112 €
Social insurance		1 806 €	2 016 €	1 447 €	1 576 €	365 €	1 150 €	1 043 €	970 €	1 441 €	866 €	404 €	606 €
Taxes from consumption	employed	759 €	848 €	609 €	663 €	153 €	484 €	439 €	408 €	606 €	354 €	170 €	255 €
	unemployed	-518 €	-504 €	-568 €	-536 €	-1 019 €	-753 €	-768 €	-835 €	501 €	249 €	200 €	298 €
Income tax	employed	277 €	309 €	222 €	242 €	56 €	177 €	160 €	149 €	221 €	133 €	62 €	93 €
	unemployed	-299 €	-347 €	-440 €	-389 €	-499 €	-511 €	-509 €	-542 €	200 €	154 €	69 €	153 €
Total / Difference		1 807 €	1 214 €	-2 046 €	-1 390 €	5 673 €	4 833 €	-3 515 €	4 895 €	7 480 €	5 047 €	669 €	3 505 €

### 4.8.5 Propensity score nearest neighbour matching

This method is very similar to the previous one. The difference is based on the rule of

pairing treated and non-treated individuals, where each treated unit is matched to the control unit with the closest propensity score. The method was applied without replacement, i.e. one participant or non-participant can be used for matching just once.

As can be seen in the reported table next to the text, due to the carried out method of matching, the samples' size in the reference periods significantly shrank. For example, in the first reference period the size of the sample is lower than the ideal estimated size of a representative sample. Every treated and non-treated were used just once and one individual was matched with his/her 5 nearest neighbours.

Reference period	Non-treated	Treated
	Valid	Valid
	N	N
1	49	72
2	1015	854
3	290	424
4	780	660
In total	2134	2010

### Measuring employability

The table below informs about frequencies of shares of time spent on the open labour market in the impact period of 24 months, as already explained. The results show that, in all reference periods, more treated individuals remained on the labour market for the entire time period. From the control group at least 72% of jobseekers were not placed on the labour market in the whole impact period. Again, there is the obvious tendency that one non-treated who finds a job will have a greater probability of remaining on the labour market in comparison with participants of traineeship. On the other hand, in the group of controls they are not frequently placed on the labour market according to the carried out methodological approach.

Group of sample	Share of impact period sustained on LM	Reference period 1			Reference period 2			Reference period 3			Reference period 4			Ref. Periods in total		
		Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent
Non-treated	0	32	65,3	65,3	678	66,8	66,8	158	54,5	54,5	511	65,5	65,5	1379	64,6	64,6
	0,1	3	6,1	34,7	39	3,8	33,2	23	7,9	45,5	32	4,1	34,5	97	4,5	35,4
	0,2	3	6,1	28,6	27	2,7	29,4	13	4,5	37,6	35	4,5	30,4	78	3,7	30,8
	0,3	2	4,1	22,4	37	3,6	26,7	17	5,9	33,1	36	4,6	25,9	92	4,3	27,2
	0,4	1	2,0	18,4	15	1,5	23,1	10	3,4	27,2	17	2,2	21,3	43	2,0	22,9
	0,5	3	6,1	16,3	19	1,9	21,6	9	3,1	23,8	18	2,3	19,1	49	2,3	20,9
	0,6	0	0,0	10,2	8	0,8	19,7	2	0,7	20,7	3	0,4	16,8	13	0,6	18,6
	0,7	0	0,0	10,2	15	1,5	18,9	3	1,0	20,0	5	0,6	16,4	23	1,1	17,9
	0,8	0	0,0	10,2	14	1,4	17,4	5	1,7	19,0	11	1,4	15,8	30	1,4	16,9
	0,9	1	2,0	10,2	8	0,8	16,1	1	0,3	17,2	11	1,4	14,4	21	1,0	15,5
	1	4	8,2	8,2	155	15,3	15,3	49	16,9	16,9	101	12,9	12,9	309	14,5	14,5
Total		49	100		1015	100		290	100		780	100		2134	100	
Treated	0	20	27,8	27,8	110	12,9	12,9	61	14,4	14,4	136	20,6	20,6	327	16,3	16,3
	0,1	2	2,8	72,2	70	8,2	87,1	29	6,8	85,6	46	7,0	79,4	147	7,3	83,7
	0,2	5	6,9	69,4	58	6,8	78,9	36	8,5	78,8	38	5,8	72,4	137	6,8	76,4
	0,3	7	9,7	62,5	79	9,3	72,1	61	14,4	70,3	67	10,2	66,7	214	10,6	69,6
	0,4	1	1,4	52,8	45	5,3	62,9	22	5,2	55,9	36	5,5	56,5	104	5,2	59,0
	0,5	7	9,7	51,4	85	10,0	57,6	34	8,0	50,7	58	8,8	51,1	184	9,2	53,8
	0,6	5	6,9	41,7	47	5,5	47,7	28	6,6	42,7	47	7,1	42,3	127	6,3	44,6
	0,7	2	2,8	34,7	55	6,4	42,2	18	4,2	36,1	35	5,3	35,2	110	5,5	38,3
	0,8	8	11,1	31,9	87	10,2	35,7	36	8,5	31,8	54	8,2	29,8	185	9,2	32,8
	0,9	4	5,6	20,8	83	9,7	25,5	27	6,4	23,3	53	8,0	21,7	167	8,3	23,6
	1	11	15,3	15,3	135	15,8	15,8	72	17,0	17,0	90	13,6	13,6	308	15,3	15,3
Total		72	100		854	100		424	100		660	100		2010	100	

The next table summarizes the types of registration of the participants and controls for different reference periods. This information should explain where our units were placed

and how they were successful financially.

In the first lines it is possible to see that participants of traineeship earned a lower assessment base per month compared to non-treated groups, except for the first reference period where the treated achieved a predominantly higher assessment base than the controls by about 22 Euros per month. Even that difference was tested with the results: non-significant differences are between the treated and non-treated in the first reference

Effect of the intervention	Group of sample	Ref. Period 1	Ref. Period 2	Ref. Period 3	Ref. Period 4
		Mean	Mean	Mean	Mean
Assessment base	non treated	440	559	543	575
	treated	461	488	496	511
Self-employment	non treated	0,00	0,00	0,00	0,00
	treated	0,01	0,01	0,01	0,00
Full-time job	non treated	0,17	0,22	0,26	0,21
	treated	0,45	0,51	0,48	0,46
Part-time job	treated	0,00	0,02	0,03	0,07
Individual barriers for entrance to LM	non treated	0,05	0,04	0,05	0,02
	treated	0,04	0,04	0,03	0,03
Placed on LM	non treated	0,17	0,22	0,26	0,21
	treated	0,45	0,52	0,49	0,46
Assessment base	net effect/ difference	22	-71	-47	-64
Self-employment		0,01	0,00	0,01	0,00
Full-time job		0,28	0,30	0,22	0,26
Individual barriers for entrance to LM		-0,02	0,00	-0,02	0,01
Placed on LM		0,28	0,30	0,23	0,26

period. Significant differences in the assessment base/wage of treated and non-treated were identified only in the second and fourth reference periods.

According to the following results, both groups of eligible jobseekers did not have any significant interest in starting with self-employment. That information has been confirmed by the statistical test presented in the table below. Also, individual barriers for entrance to the open LM did not occur very frequently according to the result in both treated and non-treated groups of samples.

One of the most important pieces of information was the dependent variable “placement on the labour market” which collates full-time job and self-employment. In accordance with the values presented in the table, it was estimated that the treated remained placed on the open LM for about 6 months longer than the controls in the period of 2 years immediately after traineeship finished. This result was similar for all the set reference periods.

Summary test hypothesis		Ref. Period 1		Ref. Period 2		Ref. Period 3		Ref. Period 4	
Null Hypothesis	Test	Sig.	Decision	Sig.	Decision	Sig.	Decision	Sig.	Decision
The distribution of Assessment base/wage is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,781	Retain the null hypothesis.	0,000	Reject the null hypothesis.	0,283	Retain the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Self-employment is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	0,995	Retain the null hypothesis.	0,997	Retain the null hypothesis.	1,000	Retain the null hypothesis.
The distribution of Full-time jobs is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Part-time jobs is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Individual barrier for entrance to LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	0,602	Retain the null hypothesis.	0,930	Retain the null hypothesis.	0,788	Retain the null hypothesis.
The distribution of Placed on LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

### Cost-benefit analysis

The estimated numbers of the analysis ensure the information about the financial influence on treated and non-treated jobseekers on the state budget in the 2 years long impact period. The methodology was applied to the numbers of the dependent variable “placement on the LM”.

On average, the treated were able to repay into the state budget in the first reference period more than 3 times more money than the grant that the state had invested in the traineeship. This trend decreases and, in the last fourth reference period, the treated returned, due to paid taxes and saved allowance, about 80 Euro-cents for each invested Euro into the traineeship.

While the non-treated are in red numbers due to their weak ability to find a job, the participants of the program in every reference period returned to the public budget grant and produced some extra money on average. The table in the last green line presents the financial positive net effects of traineeship in all reference periods.

Net effect	Reference period	Treated				Non-treated				Difference between treated and non-treated			
		1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012
Average time share on open market/labour market		45%	52%	49%	46%	17%	22%	26%	21%	28%	30%	23%	26%
Average time share out of open market/labour market		55%	48%	51%	54%	83%	78%	74%	79%	-28%	-30%	-23%	-26%
Unemployment allowance	employed	1 848 €	3 088 €	3 288 €	2 798 €	459 €	1 348 €	1 509 €	1 203 €	1 388 €	1 740 €	1 779 €	1 595 €
	unemployed	- 2 235 €	- 2 852 €	- 3 459 €	- 3 229 €	- 2 255 €	- 4 767 €	- 4 271 €	- 4 615 €	20 €	1 915 €	813 €	1 385 €
Benefit in material need	employed	1 308 €	1 476 €	1 384 €	1 318 €	448 €	626 €	395 €	587 €	860 €	850 €	988 €	731 €
	unemployed	- 1 582 €	- 1 363 €	- 1 456 €	- 1 521 €	- 2 199 €	- 2 213 €	- 1 119 €	- 2 252 €	617 €	850 €	- 337 €	731 €
Grant		- 348 €	- 1 082 €	- 1 112 €	- 1 121 €	- €	- €	- €	- €	- 348 €	- 1 082 €	- 1 112 €	- 1 121 €
Health insurance	employed	702 €	919 €	869 €	789 €	253 €	408 €	470 €	385 €	450 €	511 €	399 €	404 €
	unemployed	- 366 €	- 321 €	- 342 €	- 358 €	- 555 €	- 521 €	- 494 €	- 530 €	189 €	200 €	151 €	172 €
Social insurance		1 700 €	2 224 €	2 104 €	1 908 €	612 €	988 €	1 138 €	931 €	1 088 €	1 236 €	966 €	977 €
Taxes from consumption	employed	715 €	936 €	885 €	803 €	257 €	416 €	479 €	391 €	458 €	520 €	406 €	411 €
	unemployed	- 546 €	- 458 €	- 438 €	- 466 €	- 942 €	- 790 €	- 746 €	- 844 €	396 €	332 €	308 €	378 €
Income tax	employed	261 €	341 €	323 €	293 €	94 €	152 €	175 €	143 €	167 €	190 €	148 €	150 €
	unemployed	- 316 €	- 315 €	- 340 €	- 338 €	- 461 €	- 536 €	- 494 €	- 548 €	145 €	221 €	155 €	210 €
Total / Difference		1 141 €	2 594 €	1 705 €	875 €	4 289 €	4 889 €	1 958 €	5 149 €	5 430 €	7 483 €	4 664 €	6 024 €

### 4.8.6 Comparisons of the method results

To put all of what we have learnt from the values from the provided previous four impact evaluation designs, we prepared an overview of the outcomes. The following tables provide a comparison of the established net effects between the treated and non-treated groups of samples.

In the 5 and half years which were evaluated, more than 91 thousand eligible jobseekers were treated by traineeship. The available databases allowed us to use a limited number of correct records, which were in total 72% of all the treated in the different reference periods. The table next to the text presents a comparison of the methods from the representativeness of the samples

		Reference period				total
		1.1.2007 - 30.4.2008 16 months	1.5.2008 - 31.12.2010 32 months	1.1.2011 - 30.6.2011 6 months	1.7.2011 - 30.4.2012 10 months	
Frequencies	Minimal estimated size of samples (confidence level 95 %)	371	380	376	379	383
	No. of treated jobseekers	10 807	37 954	18 042	24 584	91 387
	post-only non-equivalent comparison	6 308	24 304	16 230	19 037	65 879
	exact matching	2 439	20 038	11 226	14 110	47 813
	propensity exact score matching	762	6 335	3 724	3 599	14 420
	propensity score nearest neighbour matching	72	854	424	660	2 010
Share on treated jobseekers	post-only non-equivalent comparison	58%	64%	90%	77%	72%
	exact matching	23%	53%	62%	57%	52%
	propensity exact score matching	7%	17%	21%	15%	16%
	propensity score nearest neighbour matching	1%	2%	2%	3%	2%



of treated jobseekers used for the estimation of net effects. Post-only non-equivalent comparison design uses every available record without elimination due to matching. The method is not very accurate but is simple to use. It is important to emphasise that the samples for this method were tested and the samples do not differ from the basic set of data that has been obtained from COLSaF. Due to the performed exact matching, samples were eliminated on average from about one third of the records that were not matched. One of the most rigorously provided methods is propensity exact score matching, which on average covers 16% of the total treated eligible graduates. The fewest samples of treated units were used for propensity score nearest neighbour matching; it is possible to say that these were used for just 2% of all treated jobseekers.

The values in the table are sorted by the used CIE design and type of the analysed dependent variables. In the final columns are summarized the significances which represent the results of the statistical tests between treated and non-treated units in the samples. Yellow cells inform us that a null hypothesis has been rejected. A null hypothesis assumes that the distribution of the values for a particular independence variable is the same across the categories of participants and non-participants.

For average wage or assessment base during the impact period were estimated mostly significant differences. Jobseekers that underwent the traineeship earned monthly less money than non-treated jobseekers by 1 to 82 Euros per month on average.

Coefficients inform us about the share of sustainability on the labour market during the whole impact period of 24 months after the intervention. The values are differences between treated and non-treated jobseekers. The blue bar charts in the cells of the table represent positive effects.

One of the most important constructed indicators that estimate employability as a net effect is represented by the dependent variable "Placed on LM." The values show that the estimated net effect is positive. In the other words, this means that the **traineeship is meaningful for employability and sustainability of the graduates as the target group**. Differences between the treated and non-treated in placement on LM were confirmed by the statistical tests. The results show that in the data are significant differences across the methods and reference periods. Only the results from the method of exact matching estimated for the last two reference periods indicate a negative impact that could be due to the used method of the net-effect estimation without using any of the matching methods of the same treated and control individuals. The Post-only non-equivalent comparison design method is not a very accurate method which uses large samples of the data which just simply compare but, on the other hand, uses also individuals who essentially differ from one another. There were also problems setting the exact impact period of the non-participants, which takes over the impact period from the matched exact or neighbour twin. That is the reason the impact period was set as the initial date of the reference period plus 2 years of impact period after the last date of the reference period. For instance, in the first reference period it was 40 months (16 months of reference period and 24 months for the impact period). That could also be a very important aspect which affects the result, because while the treated do not look for a job during the traineeship period for a maximum of 6 months, controls could be very active in seeking placement on the labour market.

Placement on LM means that the jobseeker was registered as an employee in a full-time job or was self-employed. The values show that mostly graduates were placed in full-time jobs and the numbers indicate the weak efforts of jobseekers up to 26 years of age to establish self-employment. This knowledge is contrary to the analysis of samples of the

treated by grant for self-employment in the next chapter. There we learn that young jobseekers are more willing to start self-employment if they receive a grant for it. It is possible to state that older graduates were more successful in the sustaining of self-employment; age is not a significant independent variable which influences the number of months sustaining self-employment or being placed on LM in the 2 years after the intervention had finished according to the results of the estimated linear regression model. Age was identified as a significant independent variable which had an influence on sustainability and employability although the coefficient is a very low number, which means a one year increase initiated only a minimal change in placement on the labour market.

The results of the statistical tests also present significant differences in placement in part-time jobs of the treated mainly in the last three reference periods. In the first reference period, it looks like that both the treated and non-treated were not willing to work in part-time jobs. In the first reference period, no rule in the Act on employment services, according to § 6, Art. 2 which states: The jobseeker can engage in gainful employment on the basis of employment or legal relationship under a special regulation, if the wage or remuneration for carrying out these activities shall not exceed 75% of the amount of subsistence minimum for one adult person, was identified. In the last three periods, the treated were more able to find a part-time job than their controls, even if it was just for a short time of the impact period on average.

Only rarely did the treated or non-treated have individual barriers for entrance onto the labour market because they were, for example, recipients of accident benefit, care allowance, or they were personal assistants for relatives.

Net effect	CIE design	Differences in reference periods				Sig. in reference periods			
		1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012
Average assessment base	Post-only non-comparison design	3€	-11€	-3€	-36€	0,017	0,000	0,000	0,000
	Exact matching	-1€	-16€	-9€	-26€	0,419	0,000	0,004	0,000
	Propensity score nearest neighbour matching	22€	-71€	-47€	-64€	0,781	0,000	0,283	0,000
	Propensityscore exact matching	-39€	-62€	-30€	-82€	0,038	0,000	0,001	0,000
Full-time job	Post-only non-comparison design	0,16	0,12	-0,12	-0,12	0,000	0,000	0,000	0,000
	Exact matching	0,32	0,18	0,11	0,05	0,000	0,000	0,000	0,000
	Propensity score nearest neighbour matching	0,28	0,30	0,22	0,26	0,000	0,000	0,000	0,000
	Propensityscore exact matching	0,37	0,21	0,09	0,17	0,000	0,000	0,000	0,000
Individual barriers for entrance to LM	Post-only non-comparison design	-0,01	0,00	-0,02	-0,01	0,181	0,000	0,000	0,069
	Exact matching	-0,01	0,01	0,01	0,00	0,994	0,001	0,744	0,543
	Propensity score nearest neighbour matching	-0,02	0,00	-0,02	0,01	1,000	0,602	0,930	0,788
	Propensityscore exact matching	0,02	0,00	-0,01	0,01	0,564	0,000	0,727	0,005
Part-time job	Post-only non-comparison design	0,00	0,02	0,03	0,04	1,000	0,000	0,000	0,000
	Exact matching	0,00	0,01	0,00	0,02	1,000	0,000	0,935	0,000
	Propensity score nearest neighbour matching	0,00	0,02	0,06	0,08	1,000	0,000	0,000	0,000
	Propensityscore exact matching	0,00	0,00	0,00	0,00	1,000	0,000	0,000	0,000
Placed on LM	Post-only non-comparison design	0,17	0,12	-0,11	-0,12	0,000	0,000	0,000	0,000
	Exact matching	0,32	0,18	0,11	0,05	0,000	0,000	0,000	0,000
	Propensity score nearest neighbour matching	0,28	0,30	0,23	0,26	0,000	0,000	0,000	0,000
	Propensityscore exact matching	0,38	0,21	0,10	0,17	0,000	0,000	0,000	0,000
Self-employment	Post-only non-comparison design	0,01	0,00	0,00	0,00	0,525	0,010	0,973	0,999
	Exact matching	0,00	0,00	0,00	0,00	1,000	0,744	1,000	1,000
	Propensity score nearest neighbour matching	0,01	0,00	0,01	0,00	1,000	0,995	0,997	1,000
	Propensityscore exact matching	0,01	0,00	0,00	0,00	0,999	0,211	1,000	1,000

The outcomes of the performed cost benefit analysis indicate the financial influence of the treated and non-treated eligible graduates on the national budget, as well as the net-effect of the traineeship. In the first part of the table are presented the number of treated

jobseekers, then non-treated and finally the net financial average effects in the impact periods. Values in the table are calculated per jobseeker, per impact period.

Treated jobseekers in the first reference period achieved, on average, a positive influence on the financial budget, they were able to repay the grant back to the national budget multiply, and in the second reference period as well. In the last two reference periods the situation changed and the treated had an estimated negative average effect on public financial sources mainly due to the achieved employability.

Observations on the outcomes of the financial influences of controls on the national budget indicate a generally negative effect. Only the first method of Post-only non-comparison design increased the estimated employability of the non-treated and through that influence were estimated some positive trends. But, as already mentioned, the impact periods were set for the whole time of the reference period and 2 years after, i.e. a total of 40 months for the first reference period and individuals are not matched. That approach of measurement of the registrations in the impact period may overestimate the controls.

The last part of the table presents net effects. It is obvious that in most of the cells appear yellow bar charts which inform us about the positive net effect of the treated when compared to the financial average outcomes of the non-treated. The financial outcomes are positive almost for every reference period across the carried out methods.

CIA design	Treated				Non-treated				Difference between treated and non-treated			
	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012
Post-only non-comparison design	3 357,00 €	1 807,88 €	479,19 €	- 664,84 €	139,67 €	- 363,09 €	4 575,55 €	3 583,01 €	3 217,33 €	2 170,98 €	-4 096,36 €	-4 247,85 €
Exact matching	2 922,89 €	- 50,40 €	-1 393,50 €	-2 652,52 €	-3 519,20 €	-4 121,98 €	-2 652,00 €	-2 715,09 €	6 442,09 €	4 071,58 €	1 258,50 €	62,57 €
Propensityscore exact matching	1 806,95 €	1 213,75 €	-2 846,14 €	-1 389,91 €	-5 673,35 €	-3 832,92 €	-3 514,86 €	-4 895,36 €	7 480,30 €	5 046,67 €	688,73 €	3 505,45 €
Propensity score nearest neighbour matching	1 141,13 €	2 593,53 €	1 705,33 €	875,37 €	-4 289,00 €	-4 889,10 €	-2 958,18 €	-5 148,91 €	5 480,13 €	7 482,63 €	4 663,51 €	6 024,28 €

#### 4.8.7 Identification of the successful target group for traineeship

In this sub-chapter are presented the outcomes which describe the average net effects of the key and available independent variables and their categories. Averages are divided into four reference periods of the intervention and there were also carried out statistical tests between the treated and non-treated groups of samples. The null hypothesis of the test was that distribution of the variable across the categories would be the same.

The objective of this sub-chapter is to identify the target group which was the most successful in the fundamental net-effect: placement on the open labour market and its sustainability during the impact periods across the reference periods.

**The men and women** who participated in the traineeship achieved on average a positive net effect which is significantly better than the control one. But, still on average, women had a higher share of placement on the labour market during the impact periods after the intervention finished.

Gender	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4
	Difference of means of Placement on LM between treated and non-treated individuals				Test of the difference across categories of Treated and Non-treated			
men	0,35	0,22	0,08	0,15	reject	reject	reject	reject
women	0,39	0,22	0,11	0,18	reject	reject	reject	reject

From the **marital status** characteristic, it appears that divorced individuals are more

effective than single or married participants, but their difference between treated and non-treated is insignificant and their representation in every group of samples and reference periods was less than 0.9% from all samples. Single and married graduates achieved positive placement on LM when compared to their controls. In the first two reference periods, single and married individuals achieved on average the same performance in the placement on LM, but in the two last reference periods single units were slightly more successful in remaining on the labour market.

Marital status	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4
	Difference of means of Placement on LM between treated and non-treated individuals				Test of the difference across categories of Treated and Non-treated			
divorced	0,48	0,31	0,16	0,18	retain	retain	retain	retain
single	0,38	0,22	0,09	0,17	reject	reject	reject	reject
married	0,38	0,23	0,06	0,16	reject	reject	reject	reject
registered partners	-	0,25	-	-	-	retain	-	-

College graduates had the highest potential to be employed and remain on the open labour market during the impact period, i.e. the highest grade of **education**. It is noteworthy that the overall average share of placement in the impact period of jobseekers that achieved a primary school education is in the second and third reference period as the second highest. The success rate of secondary educated jobseekers is characterized by irregular values.

Level of education	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4
	Difference of means of Placement on LM between treated and non-treated individuals				Test of the difference across categories of Treated and Non-treated			
primary school	0,13	0,22	0,19	0,10	retain	reject	reject	reject
secondary vocational school	0,40	0,18	0,04	0,15	reject	reject	reject	reject
vocational school	0,34	0,18	0,05	0,15	reject	reject	reject	reject
comprehensive school	0,28	0,22	0,09	0,11	reject	reject	reject	reject
colledge	0,39	0,37	0,21	0,26	reject	reject	reject	reject

**Jobseekers** that were **unemployed** for less than three years **before 2007** were more successful in sustaining a placement on the labour market than jobseekers that were not registered and are new in the evidence of jobseekers. This statement was possible to make thanks to the values which we learn from the outcomes of three from four reference periods. These statements indicate that the placement and sustainability on LM of some unemployed graduates is caused by the time factor. After some months in the database of jobseekers, graduates finally find a job.

Unemployed before 2007	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4
	Difference of means of Placement on LM between treated and non-treated individuals				Test of the difference across categories of Treated and Non-treated			
no evidence	0,37	0,21	0,09	0,16	reject	reject	reject	reject
< 1 year	0,39	0,27	0,19	0,21	reject	reject	reject	reject
1 - 3 years	0,42	0,18	0,13	0,28	reject	reject	retain	reject
> 3 years	0,36	0,20	0,08	0,25	reject	reject	retain	retain

Jobseekers in the western **regions** were generally, across the reference periods, more successful; in the prepared maps the detailed average differences are much more visible.

Region of permanent residence	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4
	Difference of means of Placement on LM between treated and non-treated individuals				Test of the difference across categories of Treated and Non-treated			
Bratislava region	0,48	0,26	0,17	0,15	reject	reject	reject	reject
Trnava region	0,24	0,22	0,06	0,23	reject	reject	reject	reject
Trenčín region	0,49	0,20	0,08	0,18	reject	reject	reject	reject
Nitra region	0,40	0,24	0,09	0,15	reject	reject	reject	reject
Žilina region	0,38	0,23	0,13	0,17	reject	reject	reject	reject
Banská Bystrica region	0,34	0,19	0,08	0,18	reject	reject	reject	reject
Prešov region	0,40	0,23	0,12	0,16	reject	reject	reject	reject
Košice region	0,40	0,24	0,12	0,20	reject	reject	reject	reject

Mostly higher years of **age** (particularly 23 – 24 years) are characterized across most of the reference periods as the categories that determined the success of sustaining jobseekers in employment in the impact periods. It is interesting that jobseekers about the age of 19 years had higher success in placement on the open labour market.

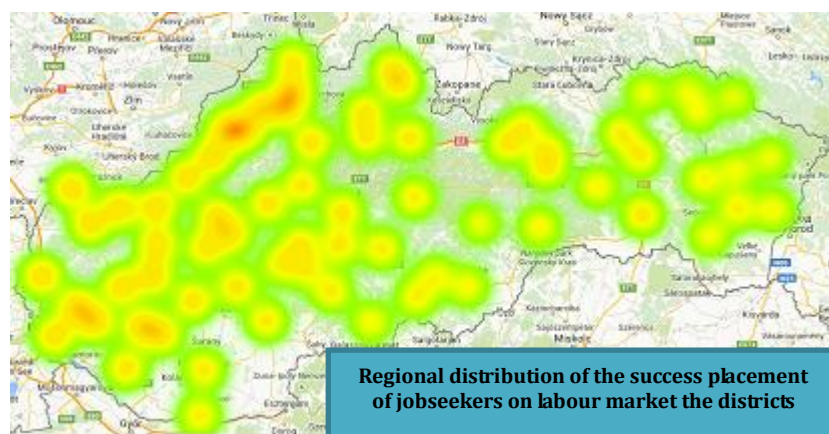
Age	Difference of means of Placement on LM between treated and non-treated individuals				Test of the difference across categories of Treated and Non-treated			
	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4	Ref. period 1	Ref. period 2	Ref. period 3	Ref. period 4
18	0,42	0,17	-0,14	0,17	reject	retain	retain	retain
19	0,39	0,30	0,14	0,26	reject	reject	reject	reject
20	0,40	0,19	0,10	0,19	reject	reject	reject	reject
21	0,38	0,25	0,03	0,13	reject	reject	reject	reject
22	0,35	0,28	0,11	0,13	reject	reject	reject	reject
23	0,41	0,28	0,09	0,16	reject	reject	reject	reject
24	0,47	0,31	0,20	0,29	reject	reject	reject	reject
25	0,00	0,13	0,29	0,27	retain	reject	reject	reject

Almost the half of the jobseekers attended the traineeship in the organisation of the public sector, such as in offices, in public administration, in health care and social centres or in educational institutions. The most frequent category in **the economic classification** of private organisations where the traineeship was carried out was wholesalers and retail traders, manufacturers, employers offering accommodation and food services, real estate traders, etc. From the values presented in the table below it is not clear that a particular economic activity generally determined jobseekers to be employed for a longer time in the impact period. But it is possible to state that individuals that carried out traineeship in public institutions achieved just a share of the sustainability on LM very close to the average of the overall sample across the reference periods, while jobseekers that carried out traineeship in the organisation of the private sector achieved slightly higher performances in placement and sustainability on LM.

Averages of "Placement on LM"	Ref. period 1			Ref. period 2			Ref. period 3			Ref. period 4		
Category of SK NACE	Mean	N	Percent	Mean	N	Percent	Mean	N	Percent	Mean	N	Percent
Public administration and defense; compulsory social security	.46	296	39%	.48	2408	38%	.34	1295	35%	.39	1257	35%
Wholesale and retail trade; repair of motor vehicles and motorcycles	.50	137	18%	.44	1060	17%	.30	667	18%	.37	611	17%
Education	.44	52	7%	.45	449	7%	.35	315	8%	.39	305	8%
Manufacturing	.52	51	7%	.52	525	8%	.38	293	8%	.39	289	8%
Other activities	.49	46	6%	.41	349	6%	.32	186	5%	.32	186	5%
Accommodation and food services	.52	45	6%	.45	358	6%	.28	208	6%	.35	183	5%
Real estate activities	.47	42	6%	.53	326	5%	.37	230	6%	.41	268	7%
Construction	.40	26	3%	.47	204	3%	.32	118	3%	.35	108	3%
Health care and social assistance	.48	21	3%	.49	197	3%	.35	128	3%	.46	120	3%
Administrative and support services	.41	10	1%	.50	130	2%	.35	96	3%	.36	88	2%
Arts, entertainment and recreation	.65	10	1%	.46	111	2%	.26	47	1%	.41	61	2%
Transport and Storage	.59	8	1%	.54	79	1%	.40	43	1%	.47	37	1%
Information and communication	.53	8	1%	.49	74	1%	.40	54	1%	.43	39	1%
Financial and insurance activities	.50	6	1%	.35	38	1%	.27	24	1%	.27	23	1%
Water supply, cleaning and waste-water treatment, waste management and remediation activities	.76	3	0%	.62	20	0%	.30	11	0%	.43	14	0%
Electricity, gas, steam and air conditioning supply	.08	1	0%	.38	7	0%	.42	8	0%	.26	9	0%
Activities of extraterritorial organizations and bodies	.00	0	0%	.00	0	0%	1.00	1	0%	1.00	1	0%
Total	.48	762	100%	.47	6335	100%	.34	3724	100%	.38	3599	100%

The table below is a correlation matrix which describes the relationship between characteristics (independent variables) of participants and a dependent variable, treatment effect – placement on LM. Blue coloured cells represent the tested significant relations.

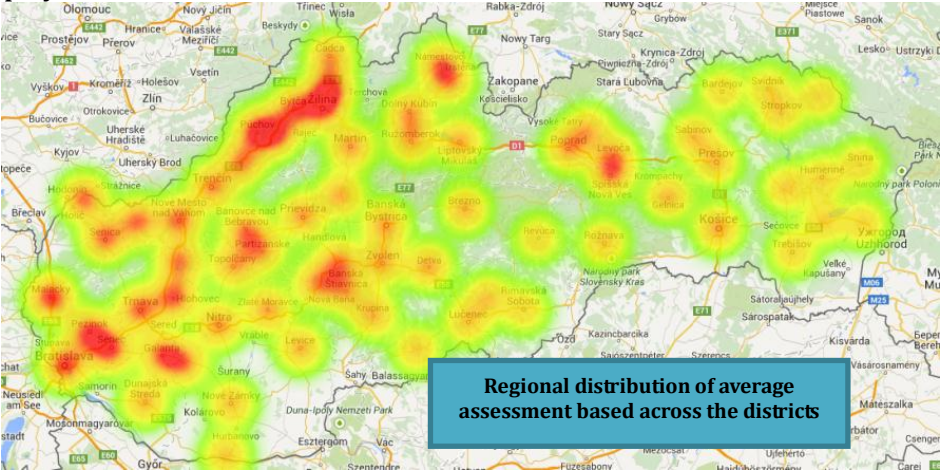
Ref. Period	Group of samples	Statistics	Gender	Age	Marital status	Level of education	Unemployed before 2007 in months	Unemployed in months	Region of permanent residence
1	non treated	Pearson Correlation	-.078	.015	-.096	-.033	.016	-.169**	-.203**
	treated	Pearson Correlation	.001	-.027	-.064	-.065	.000	-.318**	-.106**
2	non treated	Pearson Correlation	-.024**	.085**	-.075**	-.007	.009	-.153**	-.087**
	treated	Pearson Correlation	-.012	.120**	-.094**	.115**	-.011	-.392**	-.078**
3	non treated	Pearson Correlation	-.034	.061	-.058	.002	.012	-.215**	-.085*
	treated	Pearson Correlation	.008	.202**	-.099**	.178**	-.015	-.509**	-.055**
4	non treated	Pearson Correlation	-.036*	.023	-.052**	.011	-.005	-.224**	-.094**
	treated	Pearson Correlation	-.002	.195**	-.081**	.120**	-.032	-.440**	-.094**



The heat map in the figure above describes the regional differences of the achieved average differences of placement on the labour market of treated and controls. Red and orange colours represent districts with the highest net sustainability share of placement on the labour market of eligible jobseekers. From the map it is obvious that in the west of

Slovakia are districts which are more successful in placement, probably mainly due to a lower unemployment rate and the industrial concentration of the employers in the districts. These districts which are in the blue ellipse are characterized as the catchment areas of the automotive and electrical industries. The highest average net sustainability rates of the traineeship are in the Bytča and Považská Bystrica districts.

In the figure below is a map which describes the average assessment base or wage which was achieved by graduates during the impact periods. Again, the highest wage was achieved in the districts which are coloured by red spots on the map. As can be seen from the map, there are differences between the west and east of the country. While the west of the country earns a higher wage, in the east where there is a higher unemployment rate the condition on the labour market drops. Only in the districts of Poprad, Spišská Nová Ves and Levoča does the average wage achieved by jobseekers appear to be independent from the unemployment rate.



**4.8.8 Impact of the traineeship**

This part of the evaluation report describes the estimated average influence of the intervention on decreasing the unemployment rate, or number of registered jobseekers. Impact is calculated year by year according to average estimated placement on LM as the one of outcome variables. In particular, the shares of placement on LM of Propensity exact matching period method that were applied on the number of treated jobseekers in the years were used for estimation of the impact .

We measured 2 years of impact, which is the reason why the impact is also cumulative and estimated just for the number of treated jobseekers in the years from 2007 to half of 2012. In other words, it means that we calculate with the same jobseekers in two consecutive years.

To emphasise the distortion which occurs without using the counterfactual impact evaluation approach, we decided to calculate impact as the gross effect and net effect. Net effect or impact informs about the real estimated percentage of influence due to the traineeship, i.e. with subtraction of the effect which would occur if the intervention did not exist.

At least the provided traineeship decreased the number of the target group of young jobseekers by about 4% and to a maximum of about 19%. In some years were treated on average one jobseeker from 5 registered jobseekers in the 15-24 years of age range. The annual net impacts on decreasing the registered number of jobseekers of 15-24 years of age were estimated on the level from 4 to 12% (i.e. a decrease in jobseekers).

The difference between gross and net effect in this case is about 1/3rd. That is reason we can assume that, without the counter-factual impact evaluation method, impacts would be overestimated by about 33%.

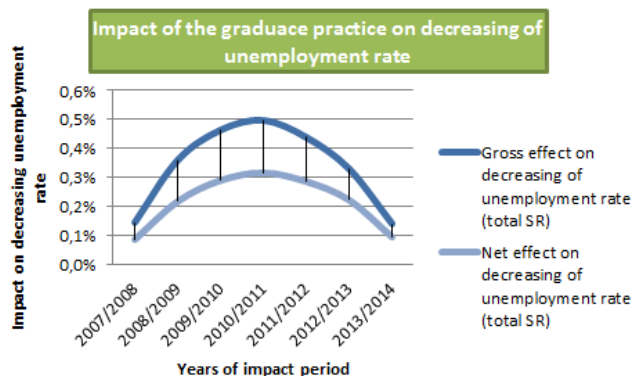
Additionally, we estimated the annual impact on decreasing the number of all registered jobseekers. The gross effect of the traineeship is from 1 to 3.4%, depending on the commutation of the previously treated jobseekers. The net impact on the number of registered jobseekers is lower and achieved values from 0.7 to 2.1%.

	2007/2008	2008/2009	2009/2010	2010/2011	2011/2012	2012/2013	2013/2014
No. of jobseekers in Slovakia (15-24 years of age)	41 873	65 989	67 462	68 782	84 372	72 629	65 469
No. of treated jobseekers	8 000	12 000	14 000	14 000	33 000	10 000	-
Estimated number of jobseekers placed on LM: gross effect	3 896	9 698	12 508	13 412	11 822	8 887	3 772
Estimated number of jobseekers placed on LM: net effect	2 296	5 848	7 813	8 522	7 703	5 974	2 532
Gross effect on decreasing of no. of registered jobseekers (15 - 24 years of age)	9%	15%	19%	19%	14%	12%	6%
Net effect on decreasing no. of registered jobseekers (15 - 24 years of age)	5%	9%	12%	12%	9%	8%	4%
No. of registered jobseekers (total SR)	248 556	379 553	381 209	399 800	425 858	398 876	373 754
Gross effect on decreasing no. of registered jobseekers (total SR)	1,6%	2,6%	3,3%	3,4%	2,8%	2,2%	1,0%
Net effect on decreasing no. of registered jobseekers (total SR)	0,9%	1,5%	2,0%	2,1%	1,8%	1,5%	0,7%
Gross effect on decreasing of unemployment rate (total SR)	0,1%	0,4%	0,5%	0,5%	0,4%	0,3%	0,1%
Net effect on decreasing of unemployment rate (total SR)	0,1%	0,2%	0,3%	0,3%	0,3%	0,2%	0,1%

Source: Statistics office of Slovak Republic, authors

Overall, the traineeship as an intervention had the power to decrease annually the unemployment rate by about 0.1 to 0.3 % in the period 2007 – 2014. This is one of the most important pieces of information from the carried out impact evaluation. That is why it is possible to assume that intervention makes sense for decreasing the unemployment rate and it is a part of the diversity of ALMP which is offered to a specific group of unemployed jobseekers.

### The financial impact of traineeship



The intention of this part of the evaluation is to estimate the overall financial impact of the ALMP measure, taking into account all the participated individuals. We count with the numbers from the performed cost-benefit analysis.

The table below composed from the two parts first shows the financial effect of the intervention according to the gross effects and the second part refers to the financial impact with consideration of the net effects. We estimated that treated individuals were able to bring to the national budget about 150 mil. Euros across the reference periods.

If we consider the estimated net effect of the intervention, the participants of the graduate practice generated for the national budget about 3.6 times more money than the same eligible jobseekers. This means the treated jobseekers brought to the national budget about 540 mil. Euros more than non-treated jobseekers in total for all reference



periods.

	Treated_gross effect in two years of impact period				
	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	In total
Propensity score nearest neighbour matching	1 141 EUR	2 594 EUR	1 705 EUR	875 EUR	1 579 EUR
No. of treated jobseekers in ref. period	10 807	37 954	18 042	24 584	91 387
Total effect on national budget	12 000 000 EUR	98 000 000 EUR	31 000 000 EUR	22 000 000 EUR	144 000 000 EUR
Estimated annual financial effect	6 000 000 EUR	49 000 000 EUR	15 500 000 EUR	11 000 000 EUR	-
	Net-effect in two years of impact period				
	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2011 - 30.6.2011	1.7.2011 - 30.4.2012	In total
Propensity score nearest neighbour matching	5 430 EUR	7 483 EUR	4 664 EUR	6 024 EUR	5 900 EUR
No. of treated jobseekers in ref. period	10 807	37 954	18 042	24 584	91 387
Total effect on national budget	59 000 000 EUR	284 000 000 EUR	84 000 000 EUR	148 000 000 EUR	539 000 000 EUR
Estimated annual financial effect	29 500 000 EUR	142 000 000 EUR	42 000 000 EUR	74 000 000 EUR	-

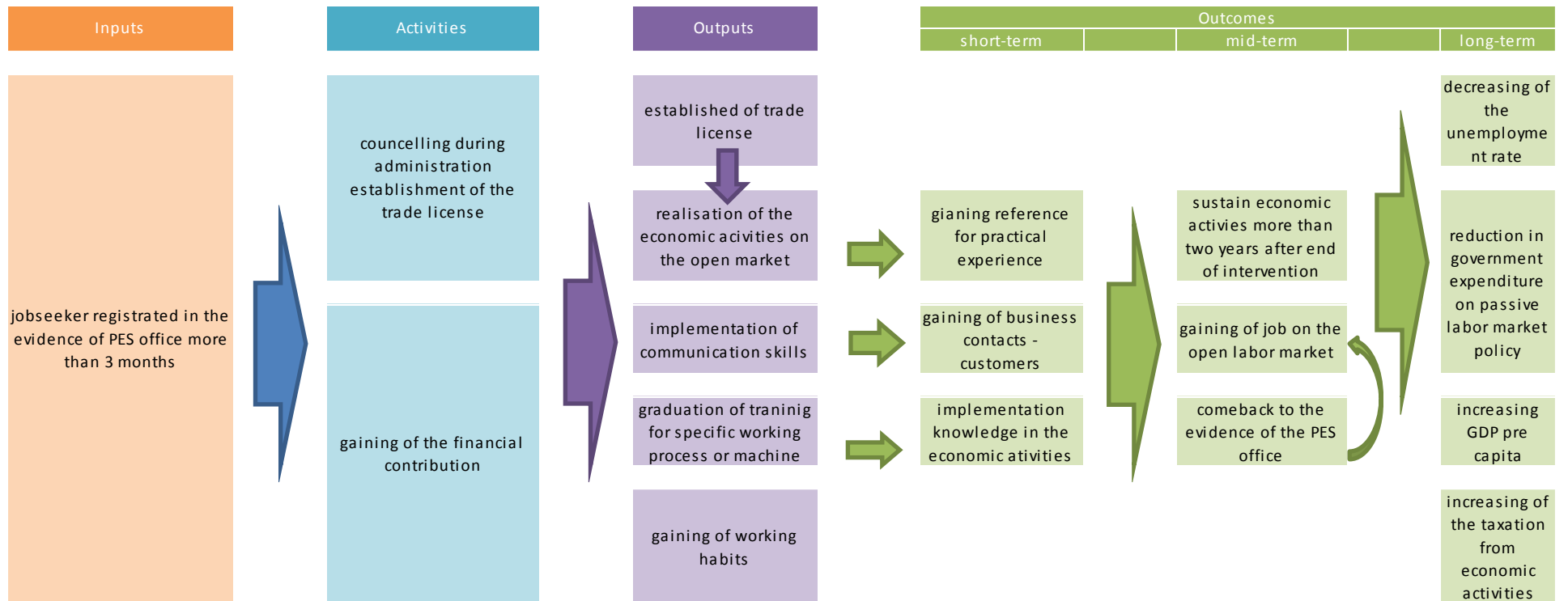
## **5 Self-employment**

The allowance for self-employment is the intervention stated in §51 of Act No. 5/2004 Coll. This ALMP measure is distributed through regional public employment offices. The intervention was introduced for the first time on 14-th April 2004.

### **5.1 Treatment effects of self-employment**

As the Explanatory Report on the Act on employment services states, the aim of the intervention is to motivate a jobseeker to launch the operation or implementation of self-employment with the possibility of using a financial grant. The grant was distributed to jobseekers that were registered for more than 3 months and submitted an application in written form with an attached business plan and budget. The intervention is accompanied by training which is focused on the practical information of establishment of self-employment (intervention by § 46 of Act No. 5/2004 Coll).

The schemes below the text present the general intervention log of self-employment intervention.



**Assumptions:**

- access to information about support under § 49
- AFE is motivated conduct business
- Preparing a business plan with the calculation of the cost of inputs

Source: authors

**External factors:**

- demand for the products or services of economic activity sole trader
  - element of chance - luck
- offer employers within a particular profession in the labor market (job)
  - speculative purpose on the part of jobseekers
  - convenient health and family circumstances

The intervention has several potential measured treatment effects on the target groups:

- sustainability of self-employment/employability on the open labour market,
- profit generated during the impact period which should indicate level of success in the business or wages earned by the employed,
- secondary effects on increasing the employment rate of the self-employed through additional job creation for employers. Unfortunately, this data was not provided to us, which is the reason the evaluation will concern just the measurable primary effects of the intervention.

We will focus on those which can be possibly measured according to data from SIA based solely on employability and amount of money which was made during the impact period. Every jobseeker included in the treated or non-treated sample had 24 months of impact period starting from the individual date of the end of intervention. Controls will admit the individual impact period according to treated pairs.

## 5.2 Reference periods

As was described in the previous monitoring report, Act No. 5/2004 Coll. on employment services and on the amendment and supplement of various acts, traineeship was revised twice between the years 2007 and 2010. Therefore our treated and non-treated jobseekers must be divided into reference periods according to changes in intervention conditions, and criteria of eligibility.

Reference period	
1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010
16 months	24 months
<b>Criteria for eligibility of jobseekers according Act No. 5/2004 Coll.:</b>	
minimum registration period in register of jobseekers: 3 months	
jobseeker apply for intervence officially in written application form	
	jobseeker must carry out his business plan
	jobseeker must attend course devoted to the business preparation.
<b>Sustainability of self-employment:</b>	
minimum two years	minimum two years
<b>Terms of the intervention:</b>	
Ammounth of grant differ from next period	Ammounth of grant differ from previous period
Financial support can be provided several times.	Financial support can be provided several times. Since 1.4.2009 - 30.4.2010 next support can be provided not less than three monts after sustainability of previous

Source: Act No. 5/2004 Coll., § 51

In the table below are presented the numbers of treated jobseekers during the reference periods of the impact evaluation. In total, more than 40 thousand jobseekers who started their own businesses were supported. We covered a total 40 months of implementation of § 49 between the years 2007 till 2010, where more than a thousand jobseekers per month were supported. In the first reference period, almost 14 thousand jobseekers in 16 months were treated, which is 850 jobseekers per month. In the second period of 24 months, there were almost 26,500 treated jobseekers and on average 1,100 jobseekers per month.

	Reference period		total
	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	
	16 months	24 months	
No. of treated jobseekers	13 650	26 486	40136
Average per month	853	1104	1003

Source: COLSaF, authors

### 5.3 Target group of self-employment promotion

The act of employment services introduced a broad definition of the target group for support of self-employment: every jobseeker that is registered more than three months in the database of the Public employment service office.

From 1st January 2007 till 30th April 2008 it was eligible to enrol in the intervention every jobseeker:

- who was registered for at least 3 months in the database of jobseekers,
- who officially submitted an application in written form

From 1st May 2008 till 30th April 2010 the rules were changed:

Eligible for support of self-employment was any jobseeker who:

- was registered for at least 3 months in the database of jobseekers,
- officially submitted an application in written form,
- attended a training program promoted under § 46 of the Act of employment services. This specific training is focused on gaining some theoretical and practical knowledge. This training was obligatorily granted to the candidates for self-employment;
- carried out a business plan covering budget.

Controls selected in our samples are jobseekers that were eligible during the reference periods. The controls had to meet the following conditions:

- registered for at least 3 months in the database of jobseekers,
- non-treated by any intervention,
- registered in the database of SIA as self-employed in the reference period. Through that selection we could make sure of the equal motivation and ability of the treated and non-treated to start self-employment, prepare and establish a business.

## 5.4 Test of representativeness of the samples

As we mentioned before, during the process of creating the samples, some individuals were excluded from the sample because they did not have recorded all the values of the variables. For the reason of the records missing data, it was required to reduce the sample and verify the representativeness of the finally selected samples. We tested the equality of distributions of frequencies of variables for individuals which are included in the final sample with those who were excluded and did not have all variables recorded. We used the Kolmogorov – Smirnov test for this purpose.

### 5.4.1 Treated group excluded from the sample

As for the case of non-treated individuals, we tested the equality of distributions of variables in the set of treated individuals included in the final sample and those excluded because of some missing value. We used the Kolmogorov – Smirnov test too. The results are in the following table.

Hypothesis Test Summary				
Treated P49				
Variable	Null Hypothesis	Test	Sig.	Decision
Gender	The distribution of values is the same across categories of selected / non selected	Independent-Samples Kolmogorov-Smirnov Test	0.518	The null hypothesis was confirmed.
Marital status			0.893	The null hypothesis was confirmed.
Level of education (10 categories)			0.988	The null hypothesis was confirmed.
Level of education (5 categories)			0.441	The null hypothesis was confirmed.
Disadvantages			1.000	The null hypothesis was confirmed.
Evidence before 2007 (in months)			0.037*	The null hypothesis was not confirmed.
Following registration in SIA			0.964	The null hypothesis was confirmed.
Driving licence (16 categories)			0.415	The null hypothesis was confirmed.
Unemployed in months	The distribution of Unemployed in months is the same across categories of selected / non selected	Independent-Samples Kolmogorov-Smirnov Test	0.067	The null hypothesis was confirmed.
Total period of all registrations in months (COLSaF)	The distribution of Total period of all registrations in months (COLSaF) is the same across categories of selected / non selected	Independent-Samples Mann-Whitney U Test Independent-Samples Kruskal-Wallis Test	0.382	The null hypothesis was confirmed.

\* for a significance level 0,01 the null hypothesis will be retained

As is presented in the tables above, the distributions of frequencies of all listed variables are the same for dropped individuals and for those which were included into the final sample. Only in the case of the variable *Evidence before 2007*, the null hypothesis is unconfirmed. But in the case of using the significance level of 0.01 instead of 0.05, the hypothesis would have been confirmed. That means it is possible to assume that our created sample is representative for the whole population.

## 5.4.2 Distributions of frequencies of treated individuals included and excluded from the sample

In the tables below the frequencies of values of all variables are compared for included individuals and those excluded in the set of treated jobseekers.

Gender Crosstabulation				
		group		Total
		non selected	selected	
Gender	men	10424	4701	15125
	women	5886	3255	9141
	unknown	15	0	15
Total		16325	7956	24281

Marital status Crosstabulation				
		group		Total
		non selected	selected	
Marital status	unknown	30	0	30
	registered partners	1	0	1
	divorced	1187	533	1720
	single	5647	3396	9043
	widow	138	42	180
	married	9322	3985	13307
Total		16325	7956	24281

Unemployed before 2007 in months Crosstabulation				
		group		Total
		non selected	selected	
Unemployed before 2007 in	< 1 year	4109	1935	6044
	1 - 3 years	4300	2158	6458
	> 3 years	4197	1815	6012
	no evidence	3719	2048	5767
Total		16325	7956	24281

Following registration in SIA Crosstabulation				
		group		Total
		non selected	selected	
Following registration in SIA	no registration	495	1151	1646
	following registration	15830	6805	22635
Total		16325	7956	24281

Level of education_10				
		group		Total
		non selected	selected	
Level of education_10 categories	Not finished education	63	1	64
	Primary education	202	145	347
	Lower secondary professional education	45	24	69
	Secondary vocational education	3914	2282	6196
	Full secondary vocational education	5325	3673	8998
	Full secondary comprehensive education	662	489	1151
	Upper vocational education	1	4	5
	Bachelor	22	18	40
	Master	1713	1318	3031
Doctoral	3	2	5	
Total		11950	7956	19906

Driving licence_16 categories Crosstabulation				
		group		Total
		non selected	selected	
Driving licence_16 categories	Driving licence: group DE	5	1	6
	Driving licence: group D	47	23	70
	Driving licence: group D1E	5	1	6
	Driving licence: group D1	47	23	70
	Driving licence: group CE	131	76	207
	Driving licence: group C	410	221	631
	Driving licence: group C1E	131	76	207
	Driving licence: group C1	410	221	631
	Driving licence: group BE	131	76	207
	Driving licence: group B	1398	839	2237
	Driving licence: group B1	1398	839	2237
	Driving licence: group A	536	300	836
	Driving licence: group A2	0	0	0
	Driving licence: group A1	536	300	836
	Driving licence: group AM	1419	847	2266
Driving licence: group T	448	244	692	
Total		7052	4087	11139

Types of disadvantages Crosstabulation				
Count		group		Total
		non selected	selected	
Types of disadvantages	no disadvantage	15960	7729	23689
	graduate	264	209	473
	long - term unemployed	88	17	105
	age over 50 years	13	1	14
Total		16325	7956	24281

### 5.4.3 Non-treated group excluded from the sample

For testing the probability distributions of frequencies for non-treated individuals included and excluded from the sample we used the Kolmogorov – Smirnov test. As already mentioned, it compared the probability distributions of the sample of included non-treated individuals with the sample of excluded non-treated individuals. We got the following results:

Hypothesis Test Summary				
Treated P49				
Variable	Null Hypothesis	Test	Sig.	Decision
Gender	The distribution of values is the same across categories of selected / non selected	Independent-Samples Kolmogorov-Smirnov Test	0.518	The null hypothesis was confirmed.
Marital status			0.893	The null hypothesis was confirmed.
Level of education (10 categories)			0.988	The null hypothesis was confirmed.
Level of education (5 categories)			0.441	The null hypothesis was confirmed.
Disadvantages			1.000	The null hypothesis was confirmed.
Evidence before 2007 (in months)			0.037*	The null hypothesis was not confirmed.
Following registration in SIA			0.964	The null hypothesis was confirmed.
Driving licence (16 categories)			0.415	The null hypothesis was confirmed.
Unemployed in months	The distribution of Unemployed in months is the same across categories of selected / non selected	Independent-Samples Kolmogorov-Smirnov Test	0.067	The null hypothesis was confirmed.
Total period of all registrations in months (COLSaF)	The distribution of Total period of all registrations in months (COLSaF) is the same across categories of selected / non selected	Independent-Samples Mann-Whitney U Test Independent-Samples Kruskal-Wallis Test	0.382	The null hypothesis was confirmed.

\* for a significance level 0,01 the null hypothesis will be retained

As we can see in the table above, the distribution of frequencies of all variables listed in the table is the same between groups of excluded individuals and those included in the sample. Only for one variable the null hypothesis about the same distribution of the samples was not confirmed: the variable *School* (in 5 categories) and variable *Driving licence* (in 16 categories). All other variables have the same distribution. That means by excluding the individuals with a missing record, we did not have significantly different groups. So our group of non-treated individuals is representative for the whole population of non-treated jobseekers.

### 5.4.4 Distributions of frequencies of non-treated individuals included and excluded from the sample

In the tables below are presented the frequencies of values of all variables compared for included individuals and those excluded from our samples.



Gender Crosstabulation				
		group		Total
		selected	non selected	
Gender	men	17212	303892	320904
	women	5172	300643	305815
	unknown	0	267	267
Total		22384	604602	626986

Marital status Crosstabulation				
		group		Total
		selected	non selected	
Marital status	unknown	0	1761	1761
	registered partners	3	274	277
	divorced	2132	63121	65253
	single	7913	234879	242792
	widow	202	13551	13753
	married	12134	291016	303150
	Total	22384	604602	626986

Unemployed before 2007 in months Crosstabulation				
		group		Total
		selected	non selected	
Unemployed before	< 1 year	5414	473906	479320
	1 - 3 years	6467	0	6467
	> 3 years	5988	0	5988
	no evidence	0	130696	130696
Total		17869	604602	622471

Following registration in SIA Crosstabulation				
		group		Total
		selected	non selected	
Following registration in SIA	no registration	0	119244	119244
	following registration	22384	485358	507742
Total		22384	604602	626986

Level of education				
		group		Total
		selected	non selected	
Level of education_10 categories	Not finished education	1	4256	4257
	Primary education	930	49092	50022
	Lower secondary professional education	277	4559	4836
	Secondary vocational education	9841	180567	190408
	Full secondary vocational education	8429	143021	151450
	Full secondary comprehensive education	882	20131	21013
	Upper vocational education	20	363	383
	Bachelor	84	2248	2332
	Master	1907	40018	41925
	Doctoral	13	334	347
Total		22384	444589	466973

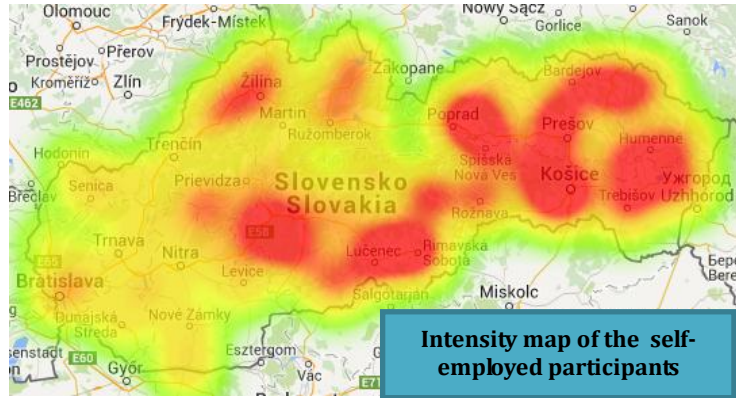
Driving licence_16 categories Crosstabulation				
		group		Total
		selected	non selected	
Driving licence_16 categories	Driving license: group DE	39	319	358
	Driving license: group D	251	2832	3083
	Driving license: group D1E	41	343	384
	Driving license: group D1	251	2832	3083
	Driving license: group CE	841	8223	9064
	Driving license: group C	2370	23189	25559
	Driving license: group C1E	841	8223	9064
	Driving license: group C1	2370	23189	25559
	Driving license: group BE	841	8223	9064
	Driving license: group B	7169	96364	103533
	Driving license: group B1	7169	96364	103533
	Driving license: group A	2599	32509	35108
	Driving license: group A2	0	7	7
	Driving license: group A1	2599	32509	35108
	Driving license: group AM	7261	98309	105570
	Driving license: group T	2579	25985	28564
Total		37221	459420	496641

Types of disadvantages Crosstabulation				
Count				
		group		Total
		selected	non selected	
Types of disadvantages	no disadvantage	20925	555077	576002
	graduate	197	8342	8539
	long - term unemployed	1026	32430	33456
	low education level	3	45	48
	organizational	3	518	521
	poor working discipline	1	78	79
	care	11	281	292
	age over 50 years	205	7171	7376
	disabled	13	660	673
	Total		22384	604602

## 5.5 Description of samples

This chapter should describe some facts about the sample at the time before the creation of pairs. This is another milestone on the path to gain matched individuals willing to establish self-employment in treated and control groups distinguished by four follow-up reference periods, which should ensure the homogeneity of intervention and validity of the counter-factual impact evaluation.

The heat or intensity map presents the number of individuals that enrolled in the program of self-employment promotion. Red areas represent the districts that were the most frequently supported. It is obvious that the majority of the participants in the samples are from the south-east parts of Slovakia, which are highly exposed to the unemployment rate, i.e. places where the intervention mostly took but growing the established business of self-employed jobseekers is very limited due to the regional purchasing power according to the lower average degree of wage in the affected areas caused by unemployment.



### 5.5.1 Permanent residence

The stated samples are composed from almost 2,400 treated jobseekers and more than 6 thousand eligible non-treated jobseekers. Individuals selected into treated and non-treated samples for both reference periods are from all regions of Slovakia. Frequencies of treated and non-treated in the regions are distributed with the biggest difference being 2%. Most of the jobseekers selected for our samples belong to Prešov region and Banská Bystrica region, where there is the highest unemployment rate in Slovakia. At first glance at the table it is obvious there is a relation between the average unemployment rate in the reference period and the number of individuals covered by the samples. Coefficients of the correlation clarify the relation between the number of treated individuals and the average unemployment

1-st reference period: 1.1.2007 – 30.4.2008							
Region	Region of permanent residence treated		Region of permanent residence non-treated		Differences between groups (%)	Average of unemployment rate in reference period (%)	
	Frequency	Percent	Frequency	Percent			
Bratislava region	175	7,4	326	7,3	-2,0	2,2	
Trnava region	181	7,6	444	8,6	-0,3	4,4	
Trenčín region	213	9,0	523	11,5	-0,3	4,7	
Nitra region	253	10,6	696	12,7	0,8	7,4	
Zilina region	265	11,2	769	16,5	1,3	6,7	
Banskobystrica region	425	17,9	1002	23,3	-1,4	15,1	
Prešov region	550	23,1	1412	14,7	0,1	13,3	
Košice region	314	13,2	893	-	1,3	12,8	
Total	2376	100,0	6065	-	-	9,2	

2-nd reference period: 1.5.2008 – 30.4.2010							
Region	Region of permanent residence treated		Region of permanent residence non-treated		Difference between groups (%)	Average unemployment rate in residence region (%)	
	Frequency	Percent	Frequency	Percent			
Bratislava region	350	5,9	825	5,1	-0,9	3,9	
Trnava region	482	8,2	1199	7,3	-0,8	7,2	
Trenčín region	651	11,0	1657	10,2	-0,9	8,2	
Nitra region	550	9,3	1653	15,0	0,8	10,4	
Zilina region	783	13,3	2455	15,4	1,8	10,3	
Banskobystrica region	965	16,3	2515	24,1	-0,9	18,4	
Prešov region	1315	22,3	3938	12,7	1,9	17,2	
Košice region	809	13,7	2077	-	-1,0	15,4	
Total	5905	100,0	16319	-	-	12,3	

1-st reference period: 1.1.2007 – 30.4.2008		
Correlation	treated	non-treated
Average of unemployment rate in reference period (%)	0,8699	0,8809
2-nd reference period: 1.5.2008 – 30.4.2010		
Correlation	treated	non-treated
Average of unemployment rate in reference period (%)	0,8777	0,8171

rate in the specific region. There is a positive correlation between these variables.

### 5.5.2 Gender

The tables indicate that intervention motivates women to establish a business or become self-employed because there is almost 16 % difference between treated and non-treated groups of women in the first reference period and almost 19 % in the second reference period. Generally, women have some barriers for making the decision to start a business. It is possible to expect a following reduction of samples after pairing according to the different share of men in treated and non-treated groups.

### 5.5.3 Marital status

Next to the text is presented a distribution of the marital status of treated and non-treated jobseekers for both reference periods. As the green bar charts shows, most of the individuals covered by all samples are married or single. Registered partners, divorcees and widows are the minority of the samples. The biggest differences are between treated and non-treated in single jobseekers for both reference periods (more than 8%). From that fact we can assume a greater willingness of single jobseekers to undergo risks without family commitments in comparison to jobseekers with another marital status.

1-st reference period: 1.1.2007 – 30.4.2008					
	Gender_treated group		Gender_non-treated group		Differences between groups (%)
	Frequency	Percent	Frequency	Percent	
men	1345	56,6	4384	72,3	15,7
women	1031	43,4	1681	27,7	-15,7
Total	2376	100,0	6065	100,0	-

2-nd reference period: 1.5.2008 – 30.4.2010					
Gender	Gender_treated		Gender_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
men	3534	59,8	12828	78,6	18,8
women	2371	40,2	3491	21,4	-18,8
Total	5905	100,0	16319	100,0	-

1-st reference period: 1.1.2007 – 30.4.2008					
	treated		non-treated		Differences between groups (%)
	Frequency	Percent	Frequency	Percent	
registered partners	0	,0	1	,0	,0
divorced	166	7,0	595	9,8	2,8
single	872	36,7	2024	33,4	-3,3
widow	15	,6	57	,9	,3
married	1323	55,7	3388	55,9	0,2
Total	2376	100,0	6065	100,0	-

2-nd reference period: 1.5.2008 – 30.4.2010					
Type of marital status	Marital status_treated		Marital status_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
registered partners	0	,0	2	,0	,0
divorced	402	6,8	1537	9,4	2,6
single	2617	44,3	5889	36,1	-8,2
widow	28	0,5	145	,9	0,4
married	2858	48,4	8746	53,6	5,2
Total	5905	100,0	16319	100,0	-

## 5.5.4 Types of disadvantage

It is suspicious that variables were not measured equally for all registered jobseekers because only about five percent of the sample admitted symptoms of a disadvantage. Most of the jobseekers in both reference periods and for treated and non-treated groups do not have any disadvantage.

According to another variable which summarizes the months of jobseekers registration, more than 80% of both groups were registered for more than one year, which indicates a long-term unemployment disadvantage.

## 5.5.5 Age

The average age of treated jobseekers in both reference periods is more than 34 years. Non-treated individuals covered in samples for both reference periods

have, on average, more than 41 years of age in the first reference period and more than 40 for the second reference period. Half of the treated samples have less than 33 years and less than forty in non-treated groups. The youngest treated jobseekers in the first reference period are 18 years old and in the second reference period 19 years old. On the other hand, the oldest treated jobseekers in the first reference period have 61 years of age and in the second reference period 73 years of age. These extreme ages show that intervention for starting a business also got jobseekers that were eligible for retirement in two years after the obligatory sustainable period.

1-st reference period: 1.1.2007 – 30.4.2008					
Type of disadvantage	Disadvantages_treated		Disadvantages_non-treated		Differences between groups (%)
	Frequency	Percent	Frequency	Percent	
no disadvantage	2297	96,7	5695	93,9	-2,8
graduate	79	3,3	58	1,0	-2,4
long - term unemployed	0	0,0	249	4,1	4,1
low education level	0	0,0	1	,0	0,0
organizational	0	0,0	2	,0	0,0
care	0	0,0	2	,0	0,0
age over 50 years	0	0,0	55	,9	0,9
disabled	0	0,0	3	,0	0,0
Total	0	0,0	6065	100,0	

2-nd reference period: 1.5.2008 – 30.4.2010					
Type of disadvantages	Disadvantages_treated		Disadvantages_treated_non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
no disadvantage	5731	97,1	15230	93,3	-3,7
graduate	156	2,6	139	,9	-1,8
long - term unemployed	17	,3	777	4,8	4,5
low education level	0	,0	2	,0	0,0
organizational	0	,0	1	,0	0,0
poor working discipline	0	,0	1	,0	0,0
care	0	,0	9	,1	0,0
age over 50 years	1	,0	150	,9	0,9
disabled	0	,0	10	,1	0,0
Total	5905	100,0	16319	100,0	-

AGE		1-st reference period: 1.1.2007 – 30.4.2008				2-nd reference period: 1.5.2008 – 30.4.2010			
		Treated		Non-treated		Descriptives_treated		Descriptives_non-treated	
		Statistic	Std. Error	Statistic	Std. Error	Statistic	Std. Error	Statistic	Std. Error
Mean		34,6351	,21170	41,5151	,13141	34,6334	,13075	40,6408	,08096
95% Confidence Interval for	Lower Bound	34,2200		41,2575		34,3771		40,4821	
	Upper Bound	35,0502		41,7727		34,8897		40,7995	
5% Trimmed Mean		34,3381		41,2440		34,1834		40,3899	
Median		33,0000		40,0000		33,0000		39,0000	
Variance		106,488		104,735		100,948		106,958	
Std. Deviation		10,31929		10,23403		10,04731		10,34204	
Minimum		18,00		19,00		18,00		19,00	
Maximum		61,00		68,00		75,76		73,49	
Range		43,00		49,00		57,76		54,49	
Interquartile Range		17,00		16,00		14,00		17,00	
Skewness		,381	,050	,338	,031	,626	,032	,357	,019
Kurtosis		-,866	,100	-,899	,063	-,150	,064	-,842	,038

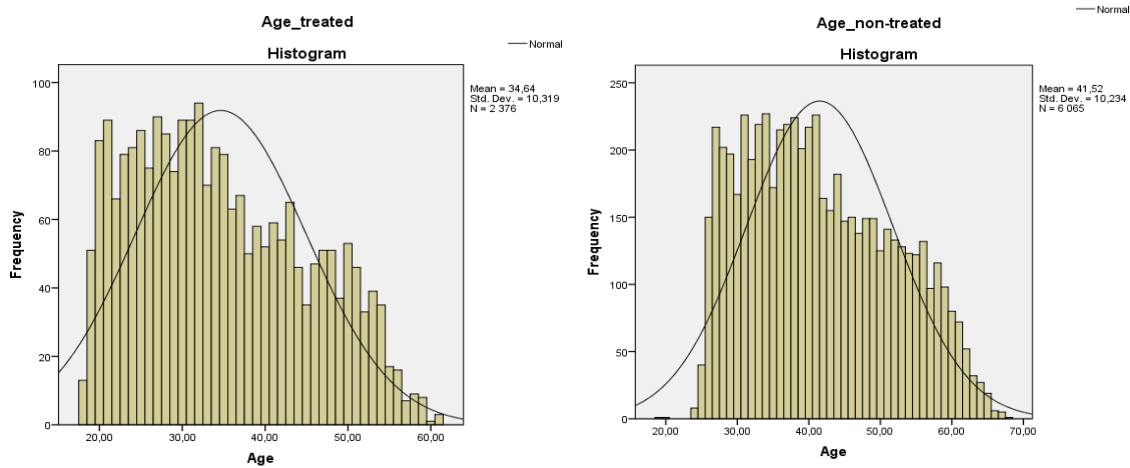
The Kolmogorov-Smirnov test of normality does not confirm the normal distribution of

age in both reference periods for treated and non-treated jobseekers in the created samples. On the other hand, histograms of distributions of age of jobseekers indicate normal distribution with right-side distribution.

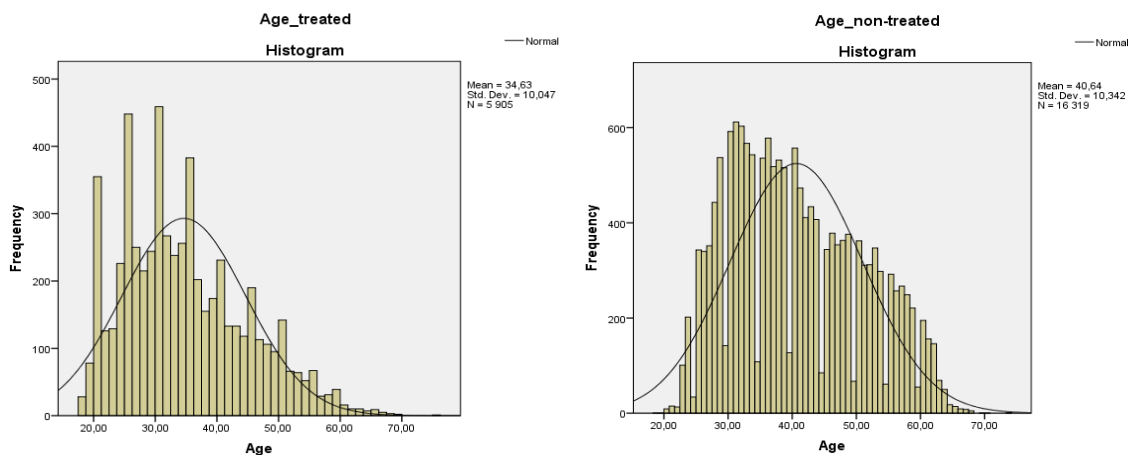
1-st reference period: 1.1.2007 – 30.4.2008					
Tests of Normality of treated			Tests of Normality of non-treated		
Kolmogorov-Smirnov <sup>a</sup>			Kolmogorov-Smirnov <sup>a</sup>		
Statistic	df	Sig.	Statistic	df	Sig.
,082	2376	,000	,076	6065	,000
2-nd reference period: 1.5.2008 – 30.4.2010					
Tests of Normality_treated			Tests of Normality_non-treated		
Kolmogorov-Smirnov <sup>a</sup>			Kolmogorov-Smirnov <sup>a</sup>		
Statistic	df	Sig.	Statistic	df	Sig.
,077	5905	,000	,076	16319	,000

a. Lilliefors Significance Correction

1-st reference period: 1.1.2007 – 30.4.2008

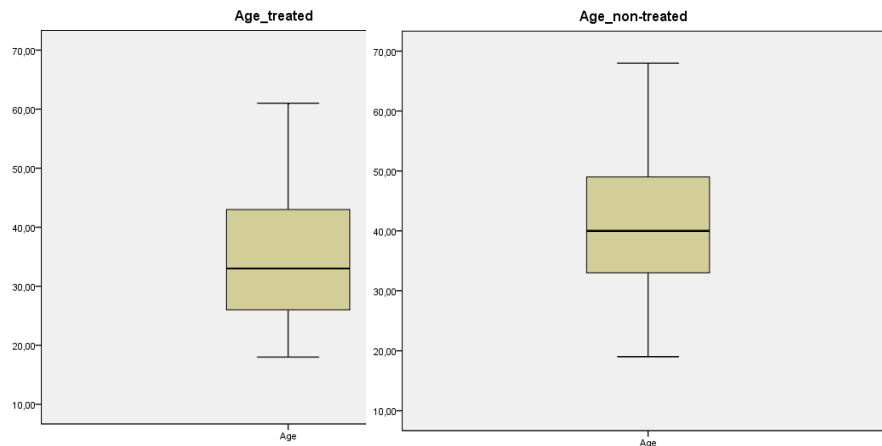


2-nd reference period: 1.5.2008 – 30.4.2010

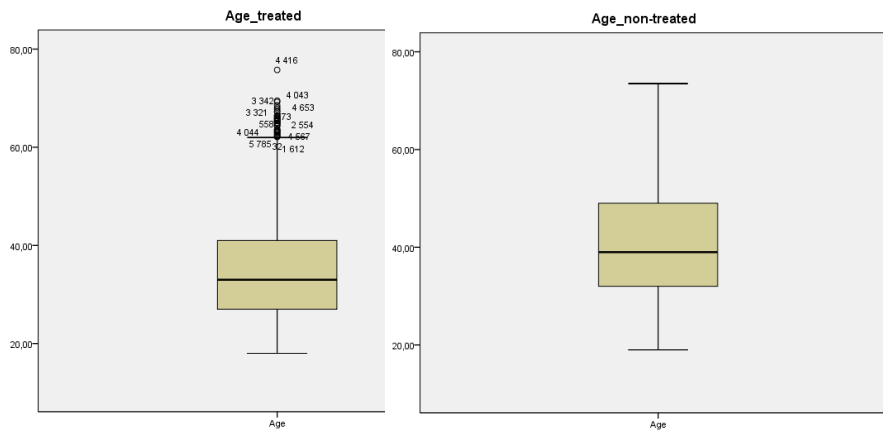


Boxplots in the charts below this text confirm a symmetric distribution of jobseekers age. The boxplot for the treated group of jobseekers in the second reference period presents the outliers of the oldest treated individuals.

1-st reference period: 1.1.2007 – 30.4.2008



2-nd reference period: 1.5.2008 – 30.4.2010



5.5.6 Level of education

Again in this case, the green bar charts in the table next to the text indicate a similar distribution in the treated and non-treated groups across the reference periods. The most frequently represented are groups of jobseekers who achieved secondary vocational school as the highest education level. The next most frequent group of highest level of education are graduates of vocational school. These groups create more than 75% in treated groups in both reference periods and more than 81% in non-treated groups. In the category of vocational school graduates, the most notable difference is between treated and non-treated groups across the periods (more than 7%).

1-st reference period: 1.1.2007 – 30.4.2008					
Level of education	Level of education_5 categories treated		Level of education_5 categories non-treated		Differences between groups (%)
	Frequency	Percent	Frequency	Percent	
primary school	81	3,4	304	5,0	1,6
secondary vocational school	1025	43,1	2474	40,8	-2,3
vocational school	790	33,2	2490	41,1	7,8
comprehensive school	139	5,9	272	4,5	-1,4
college	341	14,4	525	8,7	-5,7
Total	2376	100,0	6065	100,0	-

2-nd reference period: 1.5.2008 – 30.4.2010					
Level of education	Level of education_5 categories		Level of education_5 categories		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
primary school	90	1,5	624	3,8	2,3
secondary vocational school	2479	42,0	6629	40,6	-1,4
vocational school	1953	33,1	6964	42,7	9,6
comprehensive school	362	6,1	622	3,8	-2,3
college	1021	17,3	1480	9,1	-8,2
Total	5905	100,0	16319	100,0	-

The biggest negative difference between treated and non-treated groups is in the group of college graduates (about 7%). Those facts indicate an increased motivation of vocational school graduates

to be self-employed and a lower motivation of college graduates to establish their own business.

### 5.5.7 Registered before 2007

This variable informs us about the period of individuals' registration in the database of jobseekers before the first reference period.

From the table next to the text it is obvious that most of the treated and non-treated jobseekers are long-term unemployed. In the first reference period more than 60% of long-term unemployed jobseekers, and in the second period about half of jobseekers, are covered by our samples. The biggest difference between groups of treated and non-treated is about 6%.

1-st reference period: 1.1.2007 – 30.4.2008					
Unemployed before 2007	treated		non-treated		Differences between groups (%)
	Frequency	Percent	Frequency	Percent	
no	267	11,2	591	9,7	-1,5
< 1 year	495	20,8	1385	22,8	2,0
> 3 years	900	37,9	1949	32,1	-5,7
1 - 3 years	714	30,1	2140	35,3	5,2
Total	2376	100,0	6065	100,0	0,0

2-nd reference period: 1.5.2008 – 30.4.2010					
Unemployed before 2007	Treated		Non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent	
no	1784	30,2	3924	24,0	-6,2
< 1 year	1441	24,4	4029	24,7	0,3
> 3 years	1184	20,1	4518	27,7	7,6
1 - 3 years	1496	25,3	3848	23,6	-1,8
Total	5905	100,0	16319	100,0	-

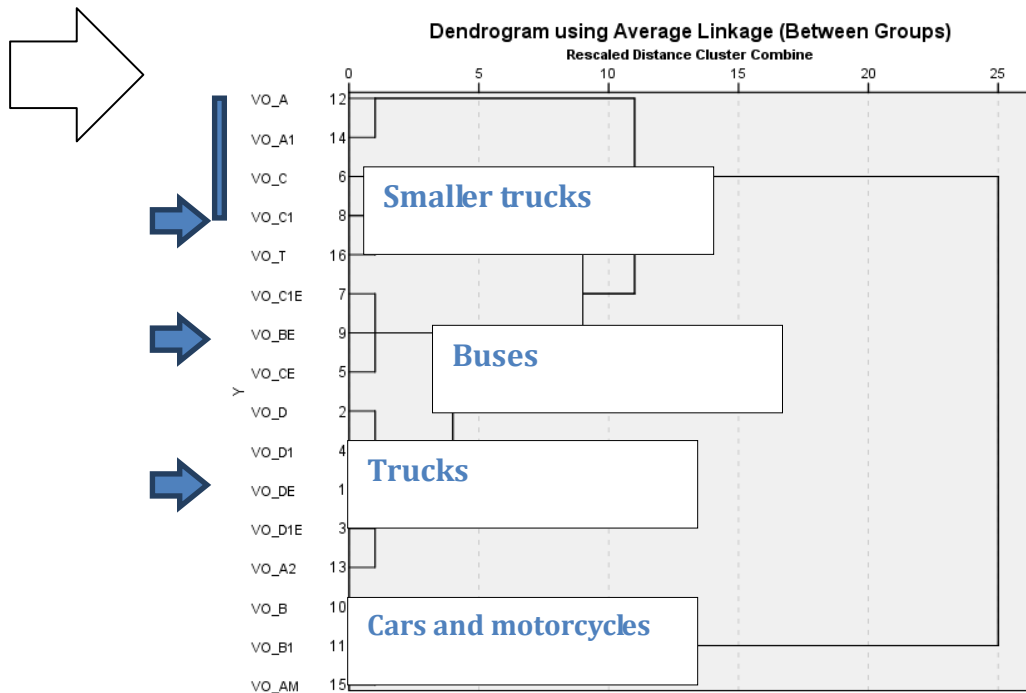
### 5.5.8 Category of driving licence

A driving licence gives the holder permission to drive with 16 types of vehicle. During realization of exact matching we found out that, due to the wide range of categories of driving licence, it is difficult to find pairs. That was the impulse for thinking about how to eliminate the wide categorization of driving licences of jobseekers. We carried out a cluster analysis which sorted permits for different categories of vehicles into groups, which then gave a maximization of homogeneity of vehicle categories.

Hierarchical clustering is based on the gradual merging of the closest pair of cases or clusters which have formed in one - each step merges one pair and the distance matrix is recalculated for the newly formed group. The algorithm is continued until all of the cases are in a cluster.

We tested the categorization in a dataset of self-employed treated and non-treated groups of jobseekers in both reference periods. In total, we tested more than 30 thousand jobseekers. The dendrogram below presents proposed clusters by vehicle type. At the fundamental level, the dendrogram shows 5 clusters, but if we assume the relation to employability there is no high contribution from the fact that a jobseeker has a driving licence for motorcycles. There are not very many types of jobs in the cultural of Slovakia which would lead to holding a driving licence for motorcycles, as there are, for instance, in Italy. That is why we used just 4 clusters of driving licences. The cluster of motorcycles was merged into the cluster of small cars and motorcycles.

There are just four types of clusters: cars and motorcycles, smaller trucks, trucks and buses.



Most of the treated and non-treated jobseekers are not holders of any driving licence (more than 88%). Just less than 12% of treated jobseekers in the samples are holders of a driving licence for the cars and motorcycles category, and less than 33% of the non-treated are holders of the same category of driving licence. The least of the jobseekers have a driving licence which could determine their placement on the labour market (trucks, buses and small trucks).

Categories of driving license	1-st reference period: 1.1.2007 – 30.4.2008					2-nd reference period: 1.5.2008 – 30.4.2010				
	Treated		Non-treated		Difference between groups (%)	Treated		Non-treated		Difference between groups (%)
	Frequency	Percent	Frequency	Percent		Frequency	Percent	Frequency	Percent	
Cars_motorcycles	291	12	1886	31	19	610	10	5375	33	23
Smaller_trucks	83	3	617	10	7	177	3	1962	12	9
Buses	29	1	179	3	2	52	1	662	4	3
Trucks	9	0	48	1	0	16	0	203	1	1
none	2084	88	4172	69	-19	5295	90	10934	67	-23

## 5.6 Analysis of variance

In the created samples of treated and non-treated individuals, it was verified by statistical hypothesis testing that the two groups significantly mutually differ in values of variables or in their probability distributions. Using one-way analysis of variance, which is an independent samples t-test, we verified the hypothesis that the means (or probability distributions) of variable frequencies are the same. Before using the independent sample t-test for two samples we always first verify whether these samples come from a normal distribution or not. In the case of non-normal distribution (which was the case for most variables), we used the non-parametric alternative to the t-test, which is the Mann-Whitney U test. We also used the Kruskal-Wallis test and the Kolmogorov-Smirnov test as non-parametric alternatives to one-way analysis of variance for two samples. The normality was verified using the Shapiro-Wilk test.



### 5.6.1 1st reference period

In the following table are the results of testing of the normal distribution of variables frequencies in the samples of treated and non-treated jobseekers in the first reference period. Based on the results from the Shapiro-Wilk test, we used the parametric or non-parametric alternative for analysis of variances.

Tests of Normality							
Variable	Treated	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	Df	Sig.
Marital status	non treated	0.299	5	0.165	0.837	5	0.157
	treated	0.264	5	.200 <sup>*</sup>	0.866	5	0.252
Level of education (10 categories)	non treated	0.345	10	0.001	0.658	10	0.000
	treated	0.301	10	0.011	0.713	10	0.001
District of School	non treated	0.164	79	0.000	0.812	79	0.000
	treated	0.161	79	0.000	0.788	79	0.000
Disadvantages	non treated	0.476	8	0.000	0.448	8	0.000
	treated	0.481	8	0.000	0.437	8	0.000
Last Occasion	non treated	0.239	41	0.000	0.72	41	0.000
	treated	0.245	41	0.000	0.774	41	0.000
Age	non treated	0.075	6065	0.000			
	treated	0.081	2376	0.000	0.959	2376	0.000
Gender	non treated	0.26	2	0.000			
	treated	0.26	2	0.000			
School (5 categories)	non treated	0.227	5	.200 <sup>*</sup>	0.895	5	0.382
	treated	0.323	5	0.096	0.738	5	0.023
Jobseeker before 2007	non treated	0.291	3	,	0.925	3	0.469
	treated	0.289	3	,	0.928	3	0.480
Driving licence	non treated	0.261	16	0.005	0.762	16	0.001
	treated	0.229	16	0.025	0.76	16	0.001

As a result of this testing, where the significance is higher than 0.05, the variable is normally distributed and vice versa. As we can see in the table above, only the variables *Marital status*, *School (5 categories)* and *Jobseeker before 2007* are normally distributed. For these three variables we used the parametric tests and, for the other variables, we used the non-parametric alternative.

In the following table, the results of testing the equality of variables or their probability distributions across the samples of treated and non-treated individuals are written. In the first table there are the results for three variables that have the normal distribution. In the second table there are the results from non-parametric testing.

Independent Samples Test			
		Levene's Test for Equality of Variances	t-test for Equality of Means

		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
<b>Marital status</b>	Equal variances assumed	6.02	0.04	-1.04	8	0.327	-737.8	706.575
	Equal variances not assumed			-1.04	5.28	0.342	-737.8	706.575
<b>Gender</b>	Equal variances assumed	470	0	-14	8439	0	-0.157	0.011
	Equal variances not assumed			-13.4	3975.81	0	-0.157	0.012
<b>School categories) (5</b>	Equal variances assumed	26.2	0	-1.34	8	0.218	-737.8	551.986
	Equal variances not assumed			-1.34	5.002	0.239	-737.8	551.986

### Level of education

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.481 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.988	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.473	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

<sup>1</sup>Exact significance is displayed for this test.

### District of school

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.000	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

### Disadvantages

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.094 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.124	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.081	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

<sup>1</sup>Exact significance is displayed for this test.

### Last occasion

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.006	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.050	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.006	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

### Age

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
2	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	.000	Reject the null hypothesis.
3	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

### Gender

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.667 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.964	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.439	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

<sup>1</sup>Exact significance is displayed for this test.

## Driving licence

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.002 <sup>1</sup>	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.001	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.002	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

<sup>1</sup>Exact significance is displayed for this test.

### Summary of all tests:

- Not significantly different values of mean or different probability distribution between groups of treated and non-treated:
  - Marital status
  - Level of education
  - School
  - Disadvantages
  - Gender
  - Jobseeker before 2007
- Significantly different values or distribution:
  - District of school
  - Last occasion
  - Age
  - Driving licence

### 5.6.2 2nd reference period

As in the first reference period, we made the verification of the normal distribution of variables frequencies and then, based on the result of this, with the Shapiro-Wilk test of normality we compared the distribution of treated and non-treated individuals.

In the following table there are the results of the normality tests.

Variable	Treated	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	Df	Sig.
Marital status	non treated	0.272	5	.200 <sup>*</sup>	0.859	5	0.226
	treated	0.307	5	0.14	0.777	5	0.052
Level of education (10 categories)	non treated	0.345	10	0.001	0.647	10	0
	treated	0.303	10	0.01	0.712	10	0.001
District of School	non treated	0.137	79	0.001	0.841	79	0
	treated	0.159	79	0	0.819	79	0
Disadvantages	non treated	0.47	9	0	0.422	9	0
	treated	0.492	9	0	0.405	9	0
Last Occasion	non treated	0.274	39	0	0.654	39	0
	treated	0.237	39	0	0.767	39	0
Age	non treated	0.076	16319	0			
	treated	0.077	5905	0			
Gender	non treated	0.485	16319	0			
	treated	0.392	5905	0			
School (5 categories)	non treated	0.309	5	0.135	0.761	5	0.038
	treated	0.189	5	.200 <sup>*</sup>	0.933	5	0.617
Jobseeker before 2007	non treated	0.39	4	.	0.754	4	0.042
	treated	0.218	4	.	0.978	4	0.887
Driving licence	non treated	0.253	16	0.007	0.749	16	0.001
	treated	0.267	16	0.003	0.777	16	0.001

Similarly to the first reference period, only 3 variables have a normal distribution of their frequencies: *Marital status*, *School (5 categories)* and *Jobseeker before 2007*. For these variables we then used an independent sample t-test to verify the hypothesis whether their means are equal or not. For all other variables we used non-parametric alternatives for this testing. The results are in the two following tables. In the first table are the results of the parametric t-test and in the second one are the results of non-parametric tests.

Independent Samples Test								
		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
<b>Marital status</b>	Equal variances assumed	9.778	0.014	1.125	8	0.293	2082.8	1850.6718
	Equal variances not assumed			1.125	5.069	0.311	2082.8	1850.6718
<b>School categories) (5</b>	Equal variances assumed	27.08	0.001	1.369	8	0.208	2082.8	1521.7272
	Equal variances not assumed			1.369	4.783	0.232	2082.8	1521.7272
<b>Jobseeker before 2007</b>	Equal variances assumed	6.554	0.043	1.546	6	0.173	1622.5	1049.8022
	Equal variances not assumed			1.546	3.084	0.218	1622.5	1049.8022

### Level of education

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.481 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.988	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.473	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

<sup>1</sup>Exact significance is displayed for this test.

### District of school

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.000	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

### Disadvantages

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.094	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.124	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.081	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

<sup>1</sup>Exact significance is displayed for this test.

### Last occasion

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	.006	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	.050	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	.006	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

## Age

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
2	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	,000	Reject the null hypothesis.
3	The distribution of Age is the same across categories of Treated/non-treated.	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

## Gender

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,667 <sup>1</sup>	Retain the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,964	Retain the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,439	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

## Driving licence

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of count is the same across categories of treated.	Independent-Samples Mann-Whitney U Test	,002 <sup>1</sup>	Reject the null hypothesis.
2	The distribution of count is the same across categories of treated.	Independent-Samples Kolmogorov-Smirnov Test	,001	Reject the null hypothesis.
3	The distribution of count is the same across categories of treated.	Independent-Samples Kruskal-Wallis Test	,002	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

<sup>1</sup>Exact significance is displayed for this test.

### Summary of all tests:

- Not significantly different values of mean or different probability distribution between groups of treated and non-treated:
  - Marital status
  - Level of education
  - School
  - Disadvantages
  - Gender
  - Jobseeker before 2007
- Significantly different values or distribution:
  - District of school
  - Last occasion
  - Age
  - Driving licence

## 5.7 Qualitative survey of self-employment

This qualitative part was carried out in the evaluation because the evaluators wanted to outline even partial motivations, aspirations, real outputs and results of the treated individuals. The main reason for this part of the research was to verify a theory about changes in the traineeship. Qualitative research was carried out through interviews over the phone. COLSaF provided a database of 48 contacts for treated individuals who were asked for interview. The database contained individuals from every region of SR (i.e. 8 regions) and three individuals for men and women, in total 48 contacts.

Finally, we carried out 17 interviews represented by 9 women and 8 men from all eight Slovak regions.

On the scheme below is described the expected theory of the changes in the intervention and prepared topics for interviews which came from three basic parts:

### A. Activities of the intervention

In the first branch of the questions which were posed to our respondents we wanted to uncover the motivation to take part in the intervention and identify activities which could lead to immediate service for the jobseeker and to increase his/her competitiveness on the open market as an entrepreneur.

During the interview we asked questions such as:

- Where did you learn about the intervention?
- How long have you planned on becoming self-employed, to start your own business?
- Have you prepared any analysis (SWOT, financial, market, competitiveness, innovation, etc.)?
- Were you self-employed in the business you worked in before or the branch from which you graduated?
- Do you have any skills or knowledge in the branch of your business?

### **B. Immediate outputs of the intervention**

Through those sorts of questions we wanted to identify provided services products with which jobseekers carried out their traineeship. We wanted to lead dialogues with jobseekers about their emotions after completing intervention.

- Has somebody helped you to prepare and carry out your business plan?
- How did training organized by PES office help you?
- What kind of information have you utilized in self-employment?
- What kind of training would help you to start your own business (soft-skills, e-business, information about electronic database of customers etc.)?

### **C. Outcomes**

This last group of questions should identify the perception of the short-term and mid-term effects of traineeship.

- Do you think the intervention helped you? Why, how?
- What would you advise changing / to do better?

#### **5.7.1 Conclusions from the interviews**

About more than one quarter of the respondents reported that they had learnt about intervention from a source other than the PES office. That information source was mainly friends, relatives or the internet. This means that most of the respondents answered that they got the initiative impulse for establishment of self-employment from an officer at the PES office. Most of the jobseekers didn't plan to do business but they took their unemployed status as the chance to become self-employed. As already presented, most of the treated jobseekers who established their self-employment were long-term unemployed before the first reference period of 1st January 2007. That is one reason why these unemployed could take this intervention as an emergency way out of their difficult living situation.

Just a few cases (i.e. 11 %) reported that they agreed with a future employer to work for the company as self-employed before they applied for the grant.

Most of those asked reported that they prepared for self-employment, but they did not want to tell how. But, in most cases, their preparation was based on skills from previous jobs. Just two respondents admitted that they wanted to start self-employment and they would have done so even if the intervention had not been granted to them. Two respondents answered that they prepared for self-employment through a specific course which they paid for on their own without any assistance from the PES office. The respondents were not able to specify how long they had prepared for intervention because they had done so a long time prior. It was obvious with many respondents that they were not willing to analyse a situation so far in the past, which is why the PES offices should have collected qualitative data immediately

after the intervention had finished.

Jobseekers did not carry out any deeper analysis of competitiveness, market, SWOT analysis, or other professional analysis. Jobseekers did not consult their business plans with any professional counsellor. Establishment of self-employment happened, in many cases, as a kind of experiment which was related to previous job skills, knowledge or contacts. When we take into account the fact that most of the treated jobseekers had finished as the highest level of education secondary school, or vocational school, it is not possible to expect that those people would be able to carry out a rigorous professional business plan according to business theory. That is the reason why intervention should be extensive in the process of counselling jobseekers in the creation of individual business plans.

Just about one quarter of asked respondents admitted that they started their business in fields they did not graduate from; the rest established themselves in the field with which they were familiar from school.

Four respondents out of five reported they had serious experience in the field of their established business. Four respondents did not have any experience in their business field from previous jobs or school, all those who were not already self-employed. That information implies the causal question: how does previous experience, or knowledge, have an influence on success in self-employment, especially in the group of secondary educated jobseekers? We can expect that a higher share of innovativeness is in the group of treated jobseekers which finished university education. Especially, the university level of education should initiate the innovative spirit of graduates.

Three from 17 replied to the answer that they had non-professional assistance during business plan preparation from family relatives, or from PES office counsellors. Most of the jobseekers prepared business plans without any help, which could be one of the key failures in the process of correctly preparing jobseekers for intervention.

Just one of the asked respondents answered that he was not satisfied with the intervention provided. We can generalize that most of the treated asked jobseekers were satisfied with the intervention and the intervention had met with the goal and promoted self-employment. The respondents report that the intervention was a starting point for them in how to escape from the evidence of unemployed jobseekers. They consider intervention a good way to start, a necessary initial impulse.

#### **What treated jobseekers would like to change?**

The vast majority of treated jobseekers would welcome some specific courses mostly based on self-representation on the market, communication strategy with clients, or customers, and information about effective communication channels used for marketing strategy. The treated missed courses based on professional advice in the fields of seeking customers, databases, information about electronic markets, etc.

Treated jobseekers would like to be informed about the law, advocacy assistance in case of bad debts, mainly in the construction sector which is a frequent profession of treated jobseekers. These self-employed have a problem earning money and that is also a reason for their failure.

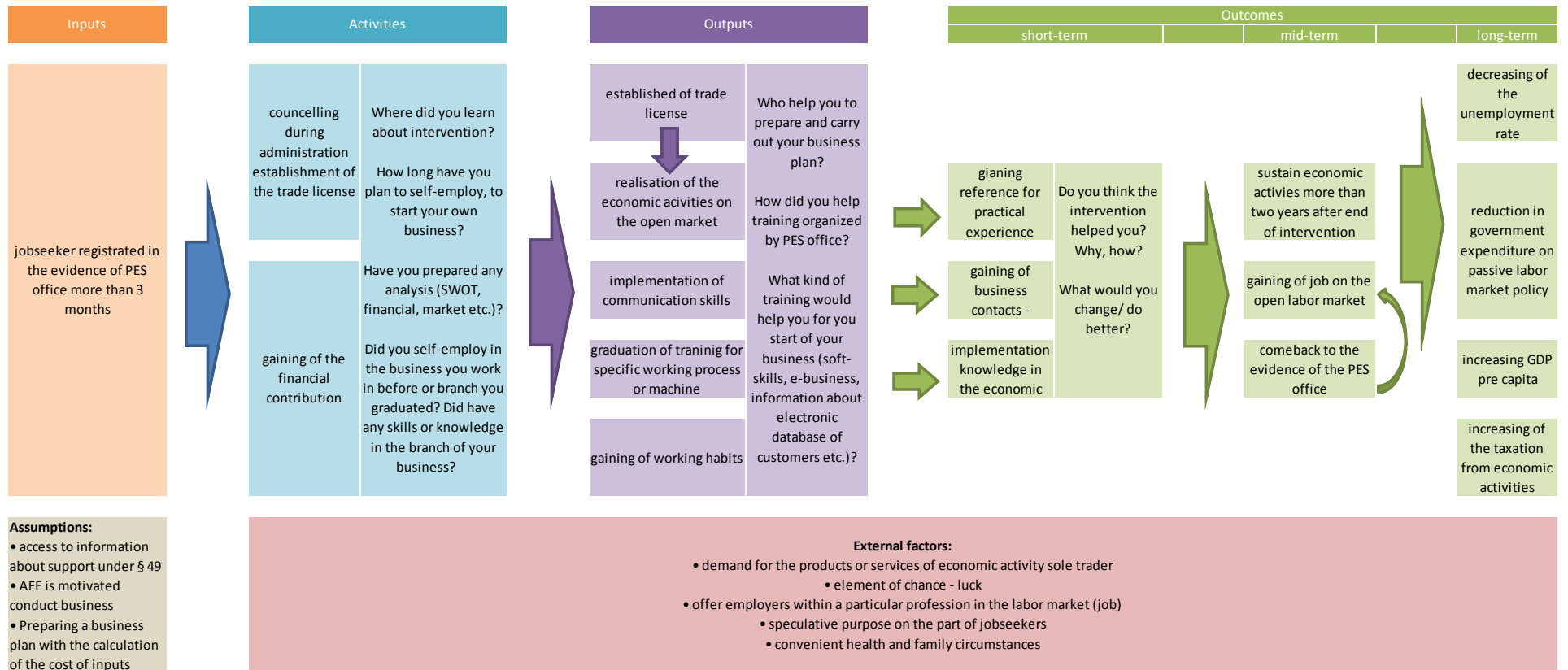
Respondents see as a limitation that they must buy exactly the same item they proposed in the approved financial plan enclosed with the business plan. Procurement of items in the financial plan is carried out with a time gap and, meanwhile, there could be an achievable product with a higher efficiency. That is why respondents would propose more flexibility in the changing types of procured items.

Some groups of respondents would propose introducing tax relief for the first two years of self-employment, which would be a reward mainly for those self-employed

jobseekers who are active and sell services or products. It is necessary to consider abuse of the tax relief.

Even obligatory preparation course concerns about preparing the jobseeker for self-employment are very positively and helpfully assessed; there are some points which could improve the effectivity of it. The asked would welcome segmentation of course participants into groups distinguished, for example, by education, because some respondents admit that they did not understand some economic categories which were familiar for the other participants who had previously dealt with accounting, etc.





## 5.8 Net effects of self-employment

### 5.8.1 Analysis of influences on self-employability

In the table next to the text are correlation coefficients and their significance on the dependent variable *Placed on LM* and Assessment base and other independent variables that are the characteristics of treated and non-treated units and their living environment. For the variable **Placed on labour market** we can see in the table of correlation coefficient, that:

- gender and age are not significant variables,
- the total period of all registrations has a negative impact on placement on LM,
- only period 2 is significant,
- if an individual is divorced or single, then they are placed on LM for a shorter period,
- primary and secondary education levels have a negative impact on placement on LM, a Master's degree has a positive impact,
- disadvantaged long term unemployed also has a negative impact.

For the variable *Self-employed*, the situation is similar. Here we can see, for example, that low education levels have a negative impact on self-employed placement on LM.

For the variable **Assessment base** we can see the following facts:

- the treated individual has a higher assessment base than the non-treated,
- women have a lower assessment base than men,
- age is not significant,
- the longer total period of all registrations has a negative impact on the assessment base,
- from marital status only single status is significant and these have a negative impact,
- primary school and comprehensive school have a positive impact on the assessment base, but college has 4 times higher impact,
- unemployment longer than 3 years has a negative impact.

Variable	placed_on_LM_pomerné		Average assessment base	
	Pearson Correlation	Sig. (2-tailed)	Pearson Correlation	Sig. (2-tailed)
placed_on_LM_pomerné			,126**	,000
Average assessment base	,126**	,000		
Gender	,091**	,000	,000	,999
Age	-,102**	,000	-,020**	,002
Unemployed in months	-,092**	,000	-,061**	,000
Total period of all registrations in months (colsaf)	-,227**	0,000	-,104**	,000
The average gross wage in the region of perm. residence	,055**	,000	,121**	,000
The proportion of women in the district of perm. residence	,025**	,000	,092**	,000
Surface of district of permanent residence	-,031**	,000	-,053**	,000
The density of population in the district of perm. residence	,022**	,000	,095**	,000
The number of municipalities in the district of perm. residence	-,007	,217	-,057**	,000
The number of cities in the district of perm. residence	,012*	,036	-,014*	,039
The registered unemployment rate in the district of perm. residence	-,068**	,000	-,081**	,000
Inhabitants density	,039**	,000	,101**	,000
Population_of_municipality_2011	,055**	,000	,056**	,000
Change_of_population:15years	,001	,869	-,008	,243
Distance_from_PESoffice	-,030**	,000	-,039**	,000
marital status=registered partners	-,013*	,025	-,004	,512
marital status=divorced	-,035**	,000	-,005	,480
marital status=single	-,002	,739	,008	,251
marital status=widow	-,019**	,001	-,002	,773
marital status=married	,025**	,000	-,004	,530
education=Not finished education	-,003	,608	-,002	,815
education=Primary education	-,109**	,000	-,005	,408
education=Lower secondary professional education	-,031**	,000	-,012	,074
education=Secondary vocational education	-,092**	,000	-,120**	,000
education=Full secondary vocational education	,052**	,000	-,007	,323
education=Full secondary comprehensive education	,034**	,000	,011	,102
education=Upper vocational education	-,003	,594	-,003	,611
education=Bachelor	-,003	,549	,003	,684
education=Master	,118**	,000	,213**	,000
education=Doctoral	-,001	,886	,009	,172
school=primary school	-,108**	,000	-,005	,426
school=secondary vocational school	,006	,283	-,040**	,000
school=vocational school	-,052**	,000	-,089**	,000
school=comprehensive school	,031**	,000	,010	,119
school=colledge	,115**	,000	,210**	,000
disadvantages=no disadvantage	,097**	,000	,036**	,000
disadvantages=graduate	,028**	,000	,003	,598
disadvantages=long - term unemployed	-,118**	,000	-,037**	,000
disadvantages=low education level	-,003	,580	-,003	,638
disadvantages=organizational	,003	,580	,007	,268
disadvantages=poor working discipline	-,002	,785	-,004	,518
disadvantages=care	-,003	,569	-,001	,876
disadvantages=age over 50 years	-,048**	,000	-,019**	,003
disadvantages=disabled	-,013*	,026	,004	,563
unemployed before 2007=< 1 year	,016**	,008	,080**	,000
unemployed before 2007=1 - 3 years	-,032**	,000	-,003	,707
unemployed before 2007=> 3 years	-,148**	,000	-,091**	,000
unemployed before 2007=no evidence	,283**	0,000	,064**	,000
period=1.0	,030**	,000	-,052**	,000
period=2.0	-,030**	,000	,052**	,000
region=Bratislavský region	,053**	,000	,116**	,000
region=Trnavský region	,024**	,000	,030**	,000
region=Trenčiansky region	,023**	,000	,012	,060
region=Nitriansky region	-,004	,495	,011	,083
region=Žilinský region	,006	,288	-,022**	,001
region=Banskobystrický region	-,023**	,000	-,024**	,000
region=Prešovský region	-,026**	,000	-,053**	,000
region=Košícký region	-,020**	,000	-,008	,240
Treated/non-treated	,583**	0,000	,047**	,000

## 5.8.2 “Post-only non-equivalent comparison design” method

There are several methodologies of how to estimate the net effect of the interventions; one the most simplistic methodologies is the difference between average treatment effects without the matching of individuals from treated and control samples. That is the reason why the method is not very robust. Another advantage of the method is its use of rather large samples.

In the table are presented two sets of reference periods; in total, more than 30 thousand individuals were used for the result, with an almost three times greater sample of controls in comparison to the treated. No missing observations were identified.

### *Measuring employability*

The frequency table below the text represents, at a glance, the average probability of the treated and controls across the set reference periods sustained on the labour market during the impact period. In the first column are situated shares of the time sustained on the labour market by the target groups in the samples, i.e. from 0 (jobseeker did not find any job in the impact period), to 1 (jobseeker remained on the labour market throughout the impact period). On the other side of the table, in the last two columns, are presented the averages for both two reference periods. According to the results, all of the treated remained on the labour market for at least one year of the impact period for a duration of 2 years. And there was an almost 10 % of probability that the treated jobseeker would remain on the labour market for the whole impact period. While controls had, on average, just up to 1 % of probability of being employed during the whole impacted period.

Yellow bar charts integrated into the table represent the tendency of the jobseekers in the different samples to be employed and sustained on the labour market in a full-time job or to be self-employed.

Simply saying, the more successful are those cumulative percentage columns that have more yellow area. In the first reference periods, the treated have more individuals that remained on the labour market mainly longer than the controls. For instance, in the first reference period it was indicated higher by almost 42 % to be employed for 70 % of the impact period for the treated while just 10 % for the non-treated.

Case Processing Summary					
reference period	Cases				
	Valid		Missing		
	N	Percent	N	Percent	
1	non treated	6 065	100,0%	0	0,0%
	treated	2 376	100,0%	0	0,0%
2	non treated	16 319	100,0%	0	0,0%
	treated	5 905	100,0%	0	0,0%
	non treated	22 384	100,0%	0	0,0%
	treated	8 281	100,0%	0	0,0%
In total		30 665			

Placed_on_L M_rounded	Reference period 1						Reference period 2						Net-effect/ differences	
	Non-treated			Treated			Non-treated			Treated			Reference period 1	Reference period 2
	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent		
0	96	1,6	1,6	0	0,0	0,0	440	2,7	2,7	0	0,0	0,0	-1,5	-2,7
0,1	312	5,1	98,4	0	0,0	100,0	1772	10,9	97,3	0	0,0	100,0	-5,1	-10,9
0,2	574	9,5	93,3	0	0,0	100,0	4562	28,0	86,4	0	0,0	100,0	-9,5	-28,0
0,3	2206	36,4	83,8	0	0,0	100,0	2363	14,5	58,5	0	0,0	100,0	-36,4	-14,5
0,4	468	7,7	47,4	823	34,6	100,0	2025	12,4	44,0	0	0,0	100,0	26,9	-12,4
0,5	456	7,5	39,7	237	10,0	65,4	2729	16,7	31,6	1891	32,0	100,0	2,5	15,3
0,6	1325	21,8	32,2	326	13,7	55,4	1027	6,3	14,9	928	15,7	68,0	-8,1	9,4
0,7	232	3,8	10,4	286	12,0	41,7	756	4,6	8,6	743	12,6	52,3	8,2	7,9
0,8	148	2,4	6,5	223	9,4	29,6	403	2,5	4,0	936	15,9	39,7	6,9	13,4
0,9	202	3,3	4,1	252	10,6	20,2	209	1,3	1,5	807	13,7	23,8	7,3	12,4
1	46	,8	0,8	229	9,6	9,6	33	,2	0,2	600	10,2	10,2	8,9	10,0
Total	6065	100,0	-	2376	100,0	-	16319	100,0	-	5905	100,0	-	-	-

On the table below the text are presented the estimated average performances of the self-employment promotion by the PES offices. There are six different dependent variables which should refer to the effects of the intervention. The first dependent variable which was measured is the average wage translated from the average assessment base in Euros, based on the records of SIA. The other effects are devoted to the placement of the jobseekers on the labour market in the form of part-time work, full-time job, or as self-employed. With those kinds of registration we can consider that the particular jobseeker was successful because he/she is out of the registration of the jobseekers and has a financial source. Even if in the registration “part-time job” isn’t a comprehensive success of employability, the jobseeker keeps in touch with labour market. Other registration refers to individual barriers for entrance to the labour market due to the needs to do personal assistance for family relatives or care for a child. The last dependent variable describes the total average months of registrations in SIA, i.e. out of the jobseeker database of the PES office.

In the next table are presented the averages of wages, and average shares of placement on the open labour market in the impact period for different types of registrations. The first row shows average assessment bases, or wages achieved in different samples in different reference periods. It is obvious that the treated ensured greater incomes than the non-treated

Dependent variable	Sample	Statistics	Reference period 1	Reference period 2
Assessment base	non treated	Mean	352,83	389,21
	treated	Mean	458,66	419,45
Full-time job	non treated	Mean	,0796	,0937
	treated	Mean	,1171	,1220
Individual barriers for entrance to LM	non treated	Mean	,0048	,0047
	treated	Mean	,0242	,0311
Placed on LM	non treated	Mean	,3884	,3434
	treated	Mean	,6137	,6915
Self-employment	non treated	Mean	,3089	,2497
	treated	Mean	,4965	,5695

but this statistical statement was rejected by the Kolmogorov-Smirnov test. There are significant differences between the assessment bases of treated and controls across selected samples. Individuals in both samples did not have an interest in being employed in part-time jobs, they preferred to find a perspective job, or source of income. In the first reference period, the treated jobseekers earned, per month, more than 100 Euros more than the non-treated and, in the second reference period, it was about 30 Euros per month.

Net-effect/ difference	Reference period 1	Reference period 2
Assessment base	105,826	30,244
Full-time job	,0375	,0284
Individual barriers for entrance to LM	,0195	,0264
Placed on LM	,2252	,3481
Self-employment	,1877	,3198

Samples of treated jobseekers had in higher frequency individual barriers to come into the labour market because of giving personal assistance to family relatives, or due to caring for a child.

Additionally, the treated remained a significantly longer time placed in full-time jobs or as self-employed than the non-treated in both reference periods. On average, the treated remained more than 60 % of the time of the impact period while the controls remained placed on the open labour market up to 40 % of the same impact periods. That is why it is possible to assume that the treated remained on the labour market longer in the first period by about more than 22 % of the impact period and in the second reference period by almost 35 %.

Also, the table below describes statements of the carried out Kolmogorov-Smirnov tests of variables which should reject or retain the null hypothesis: whether it is the distribution of the particular dependent variable which demonstrates the effect in the impact period, the same across the categories of treated/non-treated jobseekers. The statistical tests are carried out at 95 % confidence level. It is necessary to highlight inconsistency: this method is used without pairing, which is the reason why it was difficult to determine an individual impact period for controls as it was in the other methods. That is why we used the 48 months upper date of the reference period. The period of 48 months was composed of the compulsory sustaining period (24 months) and the real impact period (24 months), when the treated were not bound by any obligations.

Null Hypothesis	Test	Reference period 1		Reference period 2	
		Sig.	Decision	Sig.	Decision
The distribution of Average assessment base is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Full-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of individual barrier for entrance to LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of part-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	1,000	Retain the null hypothesis.
The distribution of Self-employment is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Placed on LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0	Reject the null hypothesis.	0	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

### Cost-benefit analysis

This paragraph is determined to show the average financial influences of provided intervention to the state budget. The numbers in the table are in three branches. The first one informs us about the performance of the treated across the reference periods, the second one about the sample of control individual jobseekers, and the last one tells us about the net effect, which is the subtraction of the treated and non-treated average performances.

Further table content items which are fundamental at the moment possibly measure the influences or flows on the state budget. Each item is divided into a situation when the

treated or non-treated jobseeker is employed. Only the items "Grant" and "Social insurance" do not distinguish between employed and non-employed statuses because the grant is paid only to the treated individuals when they are unemployed. Social insurance is not paid when a jobseeker is unemployed in the evidence of the PES office.

Net effect		Treated		Non-treated		Diff. Between treated and non-treated	
Reference period		1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010
Average time share on open market/labour market		61%	69%	39%	34%	23%	35%
Average time share out of open market/labour market		51%	54%	41%	43%	10%	11%
Unemployment allowance	employed	3 758 €	5 780 €	2 202 €	1 947 €	1 556 €	3 833 €
	unemployed	-3 124 €	-4 523 €	-2 318 €	-2 421 €	-807 €	-2 102 €
Benefit in material need	employed	1 718 €	1 779 €	1 244 €	1 100 €	475 €	680 €
	unemployed	-1 429 €	-1 392 €	-1 309 €	-1 367 €	-120 €	-25 €
Grant		-2 779 €	-2 933 €	0 €	0 €	-2 779 €	-2 933 €
Health insurance	employed	789 €	870 €	481 €	447 €	309 €	423 €
	unemployed	-656 €	-681 €	-506 €	-556 €	-151 €	-125 €
Social insurance		1 910 €	2 106 €	1 163 €	1 082 €	748 €	1 024 €
Taxes from consumption	employed	804 €	886 €	489 €	455 €	314 €	431 €
	unemployed	-19 €	0 €	0 €	0 €	-19 €	0 €
Income tax	employed	1 340 €	2 169 €	652 €	607 €	688 €	1 562 €
	unemployed	-1 114 €	-1 697 €	-686 €	-755 €	-427 €	-942 €
Total / Difference		1 198 €	2 365 €	1 412 €	538 €	-214 €	1 827 €

As we can see in the last green line of the table, both the treated and non-treated individuals brought to the state budget positive flows. Even the treated were able to return the grant back to the state budget in the way of paid taxes in the impact period. In the first reference period, one treated individual brought to the state budget almost 1,200 Euros over the cost generated due to his unemployed status in the impact period of 2 years after the intervention finished. When we switch into indicators of financial analysis, the cost effectiveness ratio shows that the invested money to one treated jobseeker by the active, or passive employment policy measures brought on average 43 % of the costs back to the national budget and in the second reference period it was already almost double, i.e. 2,400 Euros. The non-treated were, in the first reference period, merely effective, and they generated about 214 Euros greater positive flows to the state budget. On the other hand, in the second reference period, we estimated that one treated brought to the state budget about more than 1, 800 Euros more financial flow than a non-treated.

### 5.8.3 Exact matching with the application of Post-only non-equivalent comparison design

To refresh, this method is based on the creation of pairs of treated and non-treated jobseekers which are matched according to the same characteristic of independent variables, such as age, marital status, gender, number of months of jobseeker registration before the year 2007, level of education, etc.

Similar to before, for the exact matching method we used the following variables:

- reference period,
- gender,
- age (rounded to integer),
- marital status,
- region of permanent residence,
- school (5 degrees),

- length of unemployment before the year 2007 (categorized),
- driving licences categorized into 4 groups: cars and motorcycles, buses, smaller trucks, trucks.

The participants and non-participants were matched together if they had exactly the same values of these variables.

After the matching of individuals of both samples, the impact of the intervention through subtraction of the individuals' dependent variables of treated and non-treated was estimated. We measured 6 types of dependent variables, which should estimate the financial status of the individual and their employability in the impact period of 24 subsequent months:

- 1) placed on the labour market, which is total of registrations of full-time jobs and self-employment
- 2) individual barrier for entrance to LM,
- 3) part-time job,
- 4) full-time job,
- 5) self-employed,
- 6) average assessment base in Euros.

The first five variables were measured in the share of the particular type of registration in SIA during the impact period of 2 years. It was designed as a coefficient because it will be necessary to provide a comparison of results estimated based on the different types of carried out methods.

Together for both reference periods, we used almost 6400 jobseekers that created samples of treated and controls; each one of them was used just once. Every treated jobseeker was matched to individuals from the controls, which should help to estimate the net effect of self-employment promotion in different reference periods. For instance, in the first reference period, 689 treated individuals were used, and for one non-treated accounted on average 4 treated jobseekers.

Case Processing Summary					
reference period	Cases				
	Valid		Missing		
	N	Percent	N	Percent	
non treated	1	689	100,0%	0	0,0%
	2	3341	100,0%	0	0,0%
treated	1	535	100,0%	0	0,0%
	2	1821	100,0%	0	0,0%
non treated		4030	100,0%	0	0,0%
treated		2356	100,0%	0	0,0%
In total		6386			

### Measuring employability

The table presents a distribution of the samples of treated and control jobseekers across shares of sustaining time on the open labour market during the whole impact period of 24 months. The heading of the table is divided into three sections. The first two sections describe the reference periods and the second the estimated net effect for particular shares of sustaining time on the labour market. Into the cell with numbers are integrated yellow bar charts which should help to illustrate the scale of the effect provided by the specific group of samples. Simply, the more yellow highlighted in the cells, the more people who were sustained longer on the labour market as the measured desired positive effect.



Placed_on_L M_rounded	Reference period 1						Reference period 2						Net effect / difference	
	Treated			Non-treated			Treated			Non-treated			Reference period 1	Reference period 2
Share of impact period sustained on LM	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Percent	Percent
0	199	37,2	37,2	569	82,6	82,6	476	26,1	0,0	2544	76,1	76,9	-45,4	-50,0
0,1	29	5,4	62,8	11	1,6	17,4	135	7,4	73,9	56	1,7	23,9	3,8	5,7
0,2	27	5,0	57,4	6	0,9	15,8	112	6,2	66,4	55	1,6	22,2	4,2	4,5
0,3	45	8,4	52,3	14	2,0	14,9	169	9,3	60,3	102	3,1	20,5	6,4	6,2
0,4	36	6,7	43,9	10	1,5	12,9	121	6,6	51,0	46	1,4	17,5	5,3	5,3
0,5	44	8,2	37,2	13	1,9	11,5	192	10,5	44,4	119	3,6	16,1	6,3	7,0
0,6	30	5,6	29,0	2	0,3	9,6	99	5,4	33,8	61	1,8	12,5	5,3	3,6
0,7	26	4,9	23,4	6	0,9	9,3	101	5,5	28,4	60	1,8	10,7	4,0	3,8
0,8	34	6,4	18,5	15	2,2	8,4	160	8,8	22,8	93	2,8	8,9	4,2	6,0
0,9	33	6,2	12,1	16	2,3	6,2	123	6,8	14,1	71	2,1	6,1	3,8	4,6
1	32	6,0	6,0	27	3,9	3,9	133	7,3	7,3	134	4,0	4,0	2,1	3,3
Total	535	100,0		689	100,0	-	1821	100,0	-	3341	100,0	-	100,0	100,0

About every third and fourth participant of the self-employment did not find any placement during the whole impact period after the intervention finished. While three from four non-participants did not find a job in the impact period of 2 years after the matched treated finished the self-employment.

In the last section of the table are presented the net effects. It is visible that about half of the non-treated did not have any registration in SIA and were not placed on the labour market according to the available data. There could be a high number of non-treated jobseekers that didn't meet the legal conditions to be obliged to register in the database of SIA. On the other hand, the samples of treated individuals are also exposed to the same information limitations. We can only expect that this limitation is equally distributed across the treated and non-treated individuals in the samples.

Additionally, it is necessary to emphasise the fact that non-treated individuals adopted the individual impact periods of treated individuals that were matched to the non-treated into pairs. That could also be the possible reason why 80 % of non-treated jobseekers were not frequently placed on the LM.

From the yellow bar charts integrated in the table below, the treated jobseekers remained on the labour market for significantly longer than the non-treated, and the frequency table indicates extensive positive net-impacts across the reference periods.

The following tables inform us about the types of registrations in SIA of treated and non-treated jobseekers selected into samples for both reference periods.

As is presented in the first line of the table, treated jobseekers achieved on average about almost 30 Euros per month higher assessment base in the first reference period. In the second reference period, almost 20 Euros separates the treated and non-treated jobseekers.

Dependent variable	Sample	Statistics	Reference period 1	Reference period 2
Assessment base	non treated	Mean	386 €	399 €
	treated	Mean	415 €	418 €
Self-employment	non treated	Mean	0,08	0,10
	treated	Mean	0,18	0,15
Full-time job	non treated	Mean	0,03	0,04
	treated	Mean	0,16	0,25
Individual barrier for entrance to LM	non treated	Mean	0,00	0,00
	treated	Mean	0,01	0,02
Placed on LM	non treated	Mean	0,11	0,14
	treated	Mean	0,34	0,40

According to the results of the method, the treated are much more employable due to the

intervention than the controls. Non-participants were sustained for a longer time in full-time jobs; on the other hand, the treated were sustained for a longer time as self-employed. Treated and non-treated groups did not have any interest in finding part-time jobs.

This method is also limited due to the exclusion of a large part of the samples which were not matched between treated and non-treated groups.

Net-effect/ difference	Reference period 1	Reference period 2
Assessment base	29,843	19,345
Self-employment	0,10	0,05
Full-time job	0,13	0,21
Individual barrier for entrance to LM	0,01	0,02
Placed on LM	0,23	0,26

Also, the table below describes statements from the carried out Kolmogorov-Smirnov tests of variables which should reject or retain the null hypothesis: whether it is the distribution of the particular dependent variable which demonstrates the effect in the impact period, the same across categories of treated/non-treated jobseekers. The statistical tests are carried out at 95 % confidence level.

The average assessment base was significantly different in the first reference period between treated and controls; in the second reference period the differences were not significant. The distribution of individual barriers for entrance to LM and part-time jobs were the same across the categories of the variables between treated and controls. The result of the other dependent variables significantly differs between treated and controls.

Null Hypothesis	Test	Reference period 1		Reference period 2	
		Sig.	Decision	Sig.	Decision
The distribution of Average assessment base is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,578	Retain the null hypothesis.
The distribution of Self-employment is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Full-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Individual barrier for entrance to LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,544	Retain the null hypothesis.	0,544	Retain the null hypothesis.
The distribution of Part-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	1,000	Retain the null hypothesis.
The distribution of Placed on LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

### Cost-benefit analysis

As was done in the previous method, cost-benefit analyses were provided representatively for one jobseeker treated and non-treated for both reference periods with the adoption of the probability of being employed in the set impact periods. One treated was able to repay the grant and also generated, on average, more than 1 thousand Euros for the state budget. And, in the second reference period, it was on average more than 2,500 Euros. The net-effect estimated through subtraction of the controls' average financial effect is up to 6,500 Euros, according to the reference period.

Net effect		Treated		Non-treated		Diff. Between treated and non-treated	
		1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010
Reference period							
Average time share on open market/labour market		60%	70%	7%	7%	54%	63%
Average time share out of open market/labour market		51%	54%	41%	43%	10%	11%
Unemployment allowance	employed	3 701 €	5 848 €	371 €	400 €	3 330 €	5 448 €
	unemployed	-3 124 €	-4 523 €	-2 318 €	-2 421 €	607 €	-2 102 €
Benefit in material need	employed	1 692 €	1 800 €	209 €	226 €	1 483 €	1 575 €
	unemployed	-1 429 €	-1 392 €	-1 309 €	-1 367 €	120 €	-25 €
Grant		-2 779 €	-2 933 €	0 €	0 €	-2 779 €	-2 933 €
Health insurance	employed	777 €	881 €	81 €	92 €	697 €	789 €
	unemployed	-656 €	-681 €	-506 €	-556 €	151 €	-125 €
Social insurance		1 881 €	2 131 €	196 €	222 €	1 686 €	1 909 €
Taxes from consumption	employed	791 €	896 €	82 €	94 €	709 €	803 €
	unemployed	-19 €	0 €	0 €	0 €	19 €	0 €
Income tax	employed	1 319 €	2 194 €	110 €	125 €	1 210 €	2 070 €
	unemployed	-1 114 €	-1 697 €	-686 €	-755 €	427 €	942 €
Total / Difference		1 042 €	2 526 €	-3 770 €	-3 941 €	4 812 €	6 467 €

### 5.8.4 Propensity score exact matching

The procedure of application of this method consists of:

- estimation of the logistics model with its application on individuals on the samples of treated and control individuals,
- matching only those individuals who have the same value of propensity score,
- individual non-treated adopted impact periods of the treated individual which was matched with the non-treated,
- enforcement of post-only comparison design,
- tests of differences between the treated and non-treated results of dependent variables.

Case Processing Summary					
reference period		Cases			
		Valid		Missing	
		N	Percent	N	Percent
1	non treated	956	100,0%	0	0,0%
	treated	514	100,0%	0	0,0%
2	non treated	6 968	100,0%	0	0,0%
	treated	3 432	100,0%	0	0,0%
In total	non treated	7 924	100,0%	0	0,0%
	treated	3 946	100,0%	0	0,0%
In total		11 870			

In the table next to the text are presented sample sizes. In total, pairs were created from almost 12 thousand eligible jobseekers in two reference periods. As can be seen in the table, the samples do not contain any missing data. The first reference period is represented by a smaller number of treated and non-treated of individuals in comparison with the second reference period.

For the logistic model we used all independent variables, similar to before, with categorical variables coding as written in the table:

- 1) Gender
- 2) Age
- 3) Marital status – used as a categorical variable
- 4) Level of education\_10 categories
- 5) Level of education\_5 categories – used as a categorical variable
- 6) Types of disadvantages
- 7) Unemployed in months
- 8) Total period of all registrations in months (COLSaF)
- 9) Unemployed before 2007 in months – used as a categorical variable
- 10) The average gross wage in the region of permanent residence
- 11) The proportion of women in the district of permanent residence
- 12) Surface area of district of permanent residence
- 13) The density of population in the district of permanent residence
- 14) The number of municipalities in the district of permanent residence
- 15) The number of cities in the district of permanent residence
- 16) The registered unemployment rate in the district of permanent residence
- 17) Inhabitants density
- 18) Population of municipality in 2011
- 19) Change of population: 15 years
- 20) Distance from PES office

Categorical Variables Codings	
Level of education_5 categories	primary school
	secondary vocational school
	vocational school
	comprehensive school
	colledge
Marital status	registered partners
	divorced
	single
	widow
Unemployed before 2007 in months	married
	< 1 year
	1 - 3 years
	> 3 years
	no evidence

- 21) District of permanent residence
- 22) Region of permanent residence
- 23) Driving licence: cars and motorcycles, buses, trucks, small trucks

The dependent variable in the logistic regression was the variable *Treated / non-treated*, with values 1 for participants and for non-participants. In the logistic regression procedure we used the Backward conditional stepwise method, with the condition of entry probability 0.01 and removal probability 0.05. Using this method we get the final best logistic regression for modelling the probability (or odds, score) of participating in the programme with the given independent variables. This model was created separately for every reference period. In the following tables the results of the final logistic models are presented.

The results are very similar to before. As we can see, according to the values of odds  $\text{Exp}(B)$ , *Age*, *Disadvantages*, *Total period of all registrations*, *Distance from PES office*, *District of permanent residence*, *Marital status* category 3 and *Driving licence* category *motorcycles* have odds smaller than 1. That means if their value changes by 1 and all the other variables stay the same, the probability of being treated will decrease. For example for *Age*, if the individual is 1 year older, this changes the probability of being treated 0.883 times. For a categorical variable this is true compared to the reference category (the last category for all categorical variables). All variables have odds  $\text{Exp}(B)$  higher than 1, so their change (in case other variables stay the same) will cause an increase in the probability of being in a treatment group by a multiple of  $\text{Exp}(B)$ .

In the second reference period, the variables with odds  $\text{Exp}(B)$  smaller than 1 cause a decrease in the probability of being in the treatment group, in case they change by 1 and the other variables stay the same. Other variables with odds  $\text{Exp}(B)$  greater than 1 increase the probability of being in the treatment group with a change in these variables of 1.

These two logistic regression models

Variables in the Equation Period 1						
	B	S.E.	Wald	df	Sig.	Exp(B)
Gender	,387	,070	30,388	1	,000	1,473
Age (rounded)	-,125	,005	671,846	1	,000	,883
Education level	,211	,021	100,035	1	,000	1,235
Disadvantages	-,502	,155	10,453	1	,001	,605
Unemployed in months	,115	,006	398,460	1	,000	1,122
Total period of all registrations in months (colsaf)	-,084	,005	280,536	1	,000	,919
The density of population in the district of perm. residence	,000	,000	5,029	1	,025	1,000
The registered unemployment rate in the district of perm. residence	,030	,008	14,280	1	,000	1,031
Population_of_municipality_2011	,000	,000	8,772	1	,003	1,000
Distance_from_PESoffice	-,008	,004	4,023	1	,045	,992
District of permanent residence	-,092	,021	19,013	1	,000	,912
motorcycles	-,707	,104	46,027	1	,000	,493
trucks	1,476	,453	10,630	1	,001	4,377
Marital status_category_3	-1,349	,087	238,655	1	,000	,259
Marital status_category_4	1,209	,401	9,083	1	,003	3,350
Constant	3,649	,283	166,428	1	,000	38,447

Variables in the Equation Period 2						
	B	S.E.	Wald	df	Sig.	Exp(B)
Gender	,757	,040	357,821	1	,000	2,133
Age (rounded)	-,082	,002	1197,217	1	,000	,921
Education level	,344	,023	223,954	1	,000	1,410
Disadvantages	-,764	,079	94,511	1	,000	,466
Unemployed in months	,024	,003	69,832	1	,000	1,025
Total period of all registrations in months (colsaf)	-,018	,002	61,858	1	,000	,982
The average gross wage in the region of perm. residence	,001	,000	11,559	1	,001	1,001
Surface of district of permanent residence	-,001	,000	33,713	1	,000	,999
The density of population in the district of perm. residence	,000	,000	14,641	1	,000	1,000
The number of cities in the district of perm. residence	,008	,001	53,442	1	,000	1,008
The number of municipalities in the district of perm. residence	,071	,023	9,968	1	,002	1,074
Inhabitants density	,000	,000	8,331	1	,004	1,000
Population_of_municipality_2011	,000	,000	20,461	1	,000	1,000
motorcycles	-1,303	,058	503,522	1	,000	,272
Marital status_category_2	-,217	,072	9,055	1	,003	,805
Marital status_category_3	-,713	,046	237,122	1	,000	,490
School_category_1	,749	,221	11,462	1	,001	2,115
School_category_2	,792	,100	62,089	1	,000	2,208
School_category_3	,783	,106	54,739	1	,000	2,187
School_category_4	,620	,107	33,855	1	,000	1,859

were created with a significance level of 0.05; all coefficients are statistically significant, tested with the Wald test. The classification result is correct in more than 90 % of cases. Nagelkerke R-square is more than 80 % in both reference periods.

In the table next to the text the sample sizes are presented. In total, pairs were created from almost 12 thousand eligible jobseekers in two reference periods. As can be seen in the table, the samples do not contain any missing data. The first reference period is represented by a smaller number of treated and non-treated of individuals in comparison with the second reference period.

### Measuring employability

Another table below the text represents the share of sustained jobseekers in the impact period on the labour market in the first column. Then the table refers to values for the first and second reference periods for treated and non-treated groups of samples; finally, in the last two columns are presented the net effect of the interventions for the concrete share of remaining on the labour market in the impact period. While about 10 % of treated jobseekers could not be placed on LM in the impact period, more than 40 % of the controls were not employed during the whole first impact period. From the shape created by the yellow bar chart it is obvious that the treated lose placement on LM much more easily than the controls. Just more than 6 or more than 8 % of the treated ensured placement on LM for the whole measured impact period, while almost every second non-treated jobseeker who was placed on LM sustained employment, or self-employment, for the whole measured impact period. From the frequency table it is possible to deduce (last two columns) that there is about a 30 % higher probability for non-participants that they will not find any placement during the impact period, which is the main reason why intervention has been estimated as having a positive effect. Almost every second non-participant was in the evidence of jobseekers.

Placed on LM rounded	Reference period 1						Reference period 2						Net-effect/ differences	
	Non-treated			Treated			Non-treated			Treated			Reference period 1	Reference period 2
Share of impact period sustained on LM	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent		
	0	395	41,3	41,3	50	9,7	9,7	3469	49,8	16,8	649	18,9	18,8	-31,6
0,1	44	4,6	58,7	41	8,0	90,3	245	3,5	50,2	257	7,5	81,1	3,4	4,0
0,2	55	5,8	54,1	30	5,8	82,3	268	3,8	46,7	243	7,1	73,6	0,1	3,2
0,3	81	8,5	48,3	68	13,2	76,5	348	5,0	42,9	388	11,3	66,6	4,8	6,3
0,4	28	2,9	39,9	41	8,0	63,2	199	2,9	37,9	241	7,0	55,3	5,0	4,2
0,5	42	4,4	36,9	68	13,2	55,3	233	3,3	35,0	418	12,2	48,2	8,8	8,8
0,6	20	2,1	32,5	37	7,2	42,0	150	2,2	31,7	191	5,6	36,1	5,1	3,4
0,7	23	2,4	30,4	38	7,4	34,8	114	1,6	29,5	215	6,3	30,5	5,0	4,6
0,8	29	3,0	28,0	58	11,3	27,4	192	2,8	27,9	316	9,2	24,3	8,3	6,4
0,9	17	1,8	25,0	50	9,7	16,1	105	1,5	25,2	226	6,6	15,1	7,9	5,1
1	222	23,2	23,2	33	6,4	6,4	1620	23,6	23,7	291	8,5	8,5	-16,8	-15,1
Total	956	100,0	-	514	100,0	-	6968	100,0	-	3432	100,0	-	-	-

The next table presents the types of registration in SIA during the impact periods. The first rows describe the average assessment base; in the first reference period the treated achieved about more than 90 Euros per month higher than the controls; in the second reference period the situation changed and the treated achieved about 25

Dependent variable	Sample	Statistics	Reference period 1	Reference period 2
Assessment base / wage per month	non treated	Mean	371,19	430,95
	treated	Mean	465,00	405,61
Self-employment	non treated	Mean	,3463	,2760
	treated	Mean	,2089	,1682
Full-time job	non treated	Mean	,0297	,0739
	treated	Mean	,2826	,2670
Individual barrier for entrance to LM	non treated	Mean	,0051	,0109
	treated	Mean	,0349	,0519
Place on LM	non treated	Mean	,3760	,3499
	treated	Mean	,4915	,4352

Euros per month less but statistical tests stated that the difference is insignificant. Then, from the table, it is visible that the treated jobseeker has greater interest in being employed in a full-time job than non-treated jobseekers. This statement is confirmed by the table below which presents a test of differences between the treated and non-treated.

In the first reference period there was estimated a higher than 11 % employability of the treated in the impact period, while in the second reference period the effectivity of the intervention decreased and the treated were employed for a shorter share of the impact period – 8.5 %. Finally, we may state that the program of self-employment promotion in both reference periods had a positive net effect on the probability of placement on the open labour market.

We identified that the treated and non-treated significantly differed from each other in the assessment base in the first period, in self-employability, in the tendency to find a full-time job and as well in placement on the open LM. In the table above it is also obvious that treated jobseekers are significantly more successful in sustaining full-time jobs than controls.

Net-effect/difference	Reference period 1	Reference period 2
Assessment base	93,817	25,340
Self-employment	-1374	-1078
Full-time job	,2529	,1931
Individual barrier for entrance to LM	,0298	,0410
Placed on LM	,1155	,0853

Net-effect/difference	Reference period 1	Reference period 2
Assessment base	80,326	40,968
Self-employment	-,822	-,1186
Full-time job	,2140	,1934
Individual barrier for entrance to LM	,0269	,0496
Placed on LM	,0818	,0748

Null Hypothesis	Test	Reference period 1		Reference period 2	
		Sig.	Decision	Sig.	Decision
The distribution of Assessment base is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,011	Reject the null hypothesis.	0,116	Retain the null hypothesis.
The distribution of Self-employment is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Full-time jobs is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Part-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	1,000	Retain the null hypothesis.
The distribution of Individual barrier for entrance to LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,132	Retain the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Placed on LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

### Cost-benefit analysis

In the next table there are again presented numbers uncovering the financial influences of the intervention on the state budget per jobseeker for the set impact period. The last green line shows that the treated in the first reference period were able to repay about 2/3rds of the grant back during the impact period while in the second reference period the treated were able to return on average just less than 10 % of the grant. The net effect of the intervention had an estimated negative average influence on the state budget (from 2 thousand up to 3,300 Euros per jobseeker). If the cost-benefit analysis didn't calculate the amount of the grant the participants of the program received, the net-effect would be positive. In the first reference period on average the treated earned for the public budget about 700 Euros more than the controls.

Net effect	Treated		Non-treated		Diff. Between treated and non-treated		
	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	
Reference period							
Average time share on open market/labour market	49%	44%	38%	35%	12%	9%	
Average time share out of open market/labour market	51%	54%	41%	43%	10%	11%	
Unemployment allowance	employed	3 010 €	3 638 €	2 131 €	1 984 €	879 €	1 654 €
	unemployed	-3 124 €	-4 523 €	-2 318 €	-2 421 €	-807 €	-2 102 €
Benefit in material need	employed	1 376 €	1 120 €	1 204 €	1 120 €	172 €	-1 €
	unemployed	-1 429 €	-1 392 €	-1 309 €	-1 367 €	-120 €	-25 €
Grant		-2 779 €	-2 933 €	0 €	0 €	-2 779 €	-2 933 €
Health insurance	employed	632 €	548 €	465 €	456 €	167 €	92 €
	unemployed	-656 €	-681 €	-506 €	-556 €	-151 €	-125 €
Social insurance		1 530 €	1 326 €	1 126 €	1 102 €	405 €	223 €
Taxes from consumption	employed	644 €	558 €	473 €	464 €	170 €	94 €
	unemployed	-19 €	0 €	0 €	0 €	-19 €	0 €
Income tax	employed	1 073 €	1 365 €	631 €	618 €	442 €	747 €
	unemployed	-1 114 €	-1 697 €	-686 €	-755 €	-427 €	-942 €
Total / Difference		-856 €	-2 672 €	1 212 €	645 €	-2 068 €	-3 317 €

### 5.8.5 Propensity score nearest neighbour matching

This method is very similar to the previous one. The difference is based on the rule of pairing treated and non-treated individuals, where each treated unit is matched to the control unit with the closest propensity score. The method was applied without replacement, i.e. one participant and non-participants can be used as a match only once and for every participant we used the 5 nearest neighbours in propensity score.

In the samples, in total across the reference period, more than 13 thousand jobseekers were matched from the treated and control group of samples. No missing data was identified. Every non-treated and non-treated individual was used just once and, in every group of treated and his 5 nearest neighbours, there had to be in addition to treated individual also at least one non-treated individual. That is the reason why we have 887 pairs in the first reference period and 3,129 pairs in the second reference period.

Case Processing Summary					
Group of sample	Reference period	Cases			
		Valid		Missing	
		N	Percent	N	Percent
non treated	1	2153	100,0%	0	0,0%
	2	7095	100,0%	0	0,0%
treated	1	887	100,0%	0	0,0%
	2	3129	100,0%	0	0,0%
In total		13264			

#### Measuring employability

The frequency table below again presents the shares of sustained time on the open labour market during the impact period of 24 months for treated and control units. The results are very similar to the previous one. Even a high percentage of controls were not all placed on LM during the impact period; every second one who found a place on the labour market remained employed for the whole impact period. On the other hand, a rather big part of the treated sample placed for at least for 10 % of the impact period but just every ninth or tenth remained placed on the labour market for the whole impact period. From that point of view, again the stability of placement seems to be in the group of non-treated.

The last two columns in the first line show that there is a higher than 24 % and lower than 30 % probability that the treated will be employed for at least 10 % of the impact



period, i.e. up to almost 2 and half months. On the bottom of the table, in the last two columns are presented values that show that the non-treated sustained placement on LM for the whole period with 16 % higher probability than the participants of the intervention.

Placed on LM rounded	Non-treated						Treated						Net-effect/ differences	
	Reference period 1			Reference period 2			Reference period 1			Reference period 2			Reference period 1	Reference period 2
	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent		
0	746	34,6	34,6	3492	49,2	49,2	96	10,8	10,8	595	9,0	19,0	-23,8	-30,2
0,1	104	4,8	65,4	267	3,8	50,8	71	8,0	89,2	251	8,0	81,0	3,2	4,3
0,2	123	5,7	60,5	266	3,7	47,0	58	6,5	81,2	208	6,6	73,0	0,8	2,9
0,3	161	7,5	54,8	305	4,3	43,3	121	13,6	74,6	342	10,9	66,3	6,2	6,6
0,4	60	2,8	47,3	220	3,1	39,0	80	9,0	61,0	239	7,6	55,4	6,2	4,5
0,5	83	3,9	44,5	250	3,5	35,9	108	12,2	52,0	375	12,0	47,7	8,3	8,5
0,6	62	2,9	40,7	143	2,0	32,3	49	5,5	39,8	173	5,5	35,8	2,6	3,5
0,7	45	2,1	37,8	111	1,6	30,3	60	6,8	34,3	196	6,3	30,2	4,7	4,7
0,8	54	2,5	35,7	171	2,4	28,8	95	10,7	27,5	284	9,1	24,0	8,2	6,7
0,9	40	1,9	33,2	99	1,4	26,4	75	8,5	16,8	203	6,5	14,9	6,6	5,1
1	675	31,4	31,4	1771	25,0	25,0	74	8,3	8,3	263	8,4	8,4	-23,0	-16,6
Total	2153	100,0		7095	100,0		887	100,0		3129	100,0			

The output next to the text presents the types of registrations across the impact periods the average assessment base earned during the impact period. From the table it is obvious that the treated and non-treated achieved the assessment base, in the whole period, of about 400 Euros per month. Treated units achieved, in the first reference period, 80 Euros monthly more than controls. In the next reference period, the situation changed and the controls were more successful because they were able to achieve about 40 Euros per month more than treated individuals, but this difference was established as being non-significant by the Kolmogorov-Smirnov test. As presented in the previous results, the treated had identified more individual barriers for entrance to the open LM. While non-treated individuals remained mostly self-employed, treated units were mostly placed in full-time jobs. Participants and non-participants did not have any interest in part-time jobs. Generally, the treated remained on the labour market about 3% longer than the non-treated in the first reference period, which represents about 21 days of the impact period. In the second reference period, the treated remained placed on LM longer by about 7.5% of the whole impact period of 24 months. In other words, treated individuals were more successful in placement by about 54 days than controls in average numbers.

Dependent variable	Sample	Statistics	Reference period 1	Reference period 2
Assessment base	non treated	Mean	384	444
	treated	Mean	464	403
Self-employment	non treated	Mean	,40	,28
	treated	Mean	,21	,17
Full-time job	non treated	Mean	,05	,07
	treated	Mean	,27	,27
Individual barrier for entrance to LM	non treated	Mean	,01	,01
	treated	Mean	,04	,06
Placed on LM	non treated	Mean	,45	,36
	treated	Mean	,48	,43

Net-effect/ difference	Reference period 1	Reference period 2
Assessment base	80,326	-40,968
Self-employment	-,1822	-,1186
Full-time job	,2140	,1934
Individual barrier for entrance to LM	,0269	,0496
Placed on LM	,0318	,0748

The next table presents the results of the carried out tests of differences between participants and non-participants. They significantly differ from each other in self-employment, full-time jobs, placement on LM and individual barriers for entrance to the LM.

Null Hypothesis	Test	Reference period 1	Decision	Reference period 2	Decision
		Sig.		Sig.	
The distribution of Assessment base is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,167	Retain the null hypothesis.
The distribution of Self-employment is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Full-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Part-time job is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	1,000	Retain the null hypothesis.	1,000	Retain the null hypothesis.
The distribution of Individual barrier for entrance to LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,008	Reject the null hypothesis.	0,000	Reject the null hypothesis.
The distribution of Placed on LM is the same across categories of Treated/non-treated.	Independent-Samples Kolmogorov-Smirnov Test	0,000	Reject the null hypothesis.	0,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

### Cost-benefit analysis

The financial effect on the state budget of treated and non-treated units are again estimated through cost-benefit analysis in the context of the results from the propensity score nearest to neighbour matching.

The table presents in the last green line the estimated average influences on the state budget per one individual from samples of the treated and non-treated. On average,

Net effect	Treated		Non-treated		Diff. Between treated and non-treated		
	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	
Average time share on open market/labour market	48%	43%	45%	36%	3%	7%	
Average time share out of open market/labour market	51%	54%	41%	43%	10%	11%	
Unemployment allowance	employed	2 953 €	3 617 €	2 553 €	2 030 €	▲ 400 €	▲ 1 588 €
	unemployed	- 3 124 €	- 4 523 €	- 2 318 €	- 2 421 €	→ - 807 €	▼ - 2 102 €
Benefit in material need	employed	1 350 €	1 114 €	1 442 €	1 146 €	→ - 92 €	→ - 33 €
	unemployed	- 1 429 €	- 1 392 €	- 1 309 €	- 1 367 €	→ - 120 €	→ - 25 €
Grant	- 2 779 €	- 2 933 €	- €	- €	▼ - 2 779 €	▼ - 2 933 €	
Health insurance	employed	620 €	545 €	557 €	466 €	→ 63 €	→ 79 €
	unemployed	- 656 €	- 681 €	- 506 €	- 556 €	→ - 151 €	→ - 125 €
Social insurance	1 501 €	1 318 €	1 348 €	1 128 €	▲ 153 €	▲ 190 €	
Taxes from consumption	employed	631 €	554 €	567 €	474 €	→ 64 €	→ 80 €
	unemployed	- 19 €	- €	- €	- €	→ - 19 €	→ - €
Income tax	employed	1 053 €	1 357 €	756 €	633 €	▲ 297 €	▲ 725 €
	unemployed	- 1 114 €	- 1 697 €	- 686 €	- 755 €	→ - 427 €	→ - 942 €
Total / Difference	- 1 012 €	- 2 720 €	- 2 406 €	- 778 €	- 3 418 €	- 3 498 €	

participants were able to return to the state budget in the first impact period more than 2/3rds of the provided grant; in the impact period of the second reference period it was just less than 10 % of the grant. That is why the treated have an estimated negative net effect on the state budget in the amount of almost 3,500 Euros per participant.

### 5.8.6 Comparison of the method results

This subchapter should provide a view on the outcomes of the four carried out methods that estimated the net effect of the self-employment promotion. As mentioned before, it was a 3 and a half year long evaluation period during which the intervention was distributed to the eligible jobseekers that applied for the grant. That period was divided into two separated, so called reference, periods when the intervention rules were changed.

In the table are presented in the first row the minimal size of representative samples, that estimated are at being at a confidence level of 95 %, i.e. about 380 individuals. All the methods used bigger samples, which should ensure the accuracy of the estimated outcomes across the methods. In total for both periods, more than 40 thousand eligible jobseekers received a grant from COLSaF. The Post-only non-comparison design is the method that was carried out with the assistance of all available data, which is the reason in the table the bar charts show the highest frequency of concerned samples. For the first reference period just 17 % of all treated jobseekers were used due to the availability of correct data. And, in the second reference period, we used 62 % of the program participants.

		Reference period		total
		1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	
		16 months	24 months	
Frequencies	Minimal estimated size of samples (confidence level 95 %)	374	379	381
	No. of treated jobseekers	13 650	26 486	40136
	post-only non-equivalent comparison	2376	16319	18695
	exact matching	535	1821	2356
	propensity exact score matching	514	3432	3946
	propensity score nearest neighbour matching	887	3129	4016
Share on treated jobseekers	post-only non-equivalent comparison	17%	62%	47%
	exact matching	4%	7%	6%
	propensity exact score matching	4%	13%	10%
	propensity score nearest neighbour matching	6%	12%	10%

The other performed methods counted with lower scopes of samples and representativeness due to the rules of the matching, which substantially limited samples.

The other table presents five dependent variables whose role is estimation of the net effect from some points. The first one is the **assessment base** achieved by jobseekers. The values show the differences of averages between treated and non-treated units. In the first reference period, the result is obvious because all the methods confirmed that the net financial impact on the height of the assessment base per month of treated individual was positive from 30 to 106 Euros more than the controls earned in the impact period. In the second reference period, exact matching and only-past non-equal comparison design established a positive effect of intervention on the participants' assessment base. But more rigorous methods estimated a negative net impact on the height of the assessment bases of treated units. Even the statistical test in the propensity score nearest neighbour matching method stated that negative differences between the treated and non-treated were insignificant. It is possible to make the conclusion that the assessment bases in the second reference period of treated and non-treated were similar.

For the dependent variable **full-time job**, the notion that every difference between treated and non-treated is significant was tested. The values in the table indicate that the treated were much more determined to find a job because even for the one propensity score exact matching design was estimated a positive difference between treated and controls. That method estimated the negative net effect on placement of the treated on the labour market. From the values it is obvious that the self-employment sustainability of controls is significantly higher.

The other values indicate that participants of the program are significantly more exposed to **individual barriers for entrance onto the labour market**. Even barriers were not long-term parts of the impact period but some participants were recipients of accident benefit, care allowance, or they were personal assistants for relatives during the impact periods in both reference periods of the intervention.

It is possible state that, overall, the **placement of the treated** on the open **labour market**

was more frequent and sustainable than with non-treated individuals who were eligible and also established self-employment during the impact period. In the first reference period, participants of the intervention remained on the labour market longer by up to five months. In the second reference period, three of the carried out methods confirmed a positive effect as well. Robust methods propensity nearest neighbour and propensity exact matching estimated the lowest difference, which showed that participants of the self-employment promotion remained on the open labour market on average about 50 days longer than their nearest controls. That result was tested as being the significant difference between both groups of samples.

The last dependent variable reveals that the treated were not so successful in self-employment during the impact periods. Even jobseekers supported by grants were able to remain for a longer time, on average, on the open market. They intended to find a full-time job. Mainly due to placement in full-time jobs, participants were more successful than non-participants.

Net effect	CIE design	Differences		Sig.	
		1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	1.1.2007 - 30.4.2009	1.5.2008 - 30.4.2011
Assessment base	Exact matching	30 €	19 €	0,000	0,578
	Post-only non-equal comparison design	106 €	30 €	0,000	0,000
	Propensity score exact matching	94 €	25 €	0,011	0,116
	Propensity score nearest neighbour matching	80 €	41 €	0,000	0,167
Full-time job	Exact matching	0,13	0,21	0,000	0,000
	Post-only non-equal comparison design	0,04	0,03	0,000	0,000
	Propensity score exact matching	0,25	0,19	0,000	0,000
	Propensity score nearest neighbour matching	0,21	0,19	0,000	0,000
Individual barriers for entrance to LM	Exact matching	0,01	0,05	0,544	0,544
	Post-only non-equal comparison design	0,02	0,03	0,000	0,000
	Propensity score exact matching	0,03	0,04	0,132	0,000
	Propensity score nearest neighbour matching	0,03	0,05	0,008	0,000
Placed on LM	Exact matching	0,23	0,26	0,000	0,000
	Post-only non-equal comparison design	0,23	0,35	0,000	0,000
	Propensity score exact matching	0,12	0,09	0,000	0,000
	Propensity score nearest neighbour matching	0,03	0,07	0,000	0,000
Self-employment	Exact matching	0,10	0,05	0,000	0,000
	Post-only non-equal comparison design	0,19	0,32	0,000	0,000
	Propensity score exact matching	0,14	0,11	0,000	0,000
	Propensity score nearest neighbour matching	0,18	0,12	0,000	0,000

Another outcome informs us about the estimated values of the carried out cost benefit analysis as one method which should uncover the impact of the intervention on public finance. The values were analysed for all three performed methods of the impact evaluation. Cost benefit analysis of the self-employed counted with 48 months of impact period. We assumed the term of sustainability of self-employment (24 months – a condition of the intervention) as well as the impact (24 months after the conditions of the sustainability of the self-employment compliance).

The values in the table differ according to estimated placement on the labour market across the methods. In the previous table were presented the net effects of placement on the labour market. While the first two provided methods are rather optimistic and post only-non-comparison design is not very accurate, taking into account the features of the individuals, we again advise assessing the financial influence of the evaluation by the last carried out method – propensity nearest neighbour matching.

According to those outcomes, the intervention had a negative effect on the national

budget. Both reference periods estimated a very similar net financial impact on public finance. The provided values show that one treated can generate for the national budget almost 3500 Euros less than the nearest control individual. On the other hand, in the cost benefit analysis, the provided grant was taken into account. If the grant were not counted, then the net impact would be lower (the average grant was more than 2900 Euros).

CIA design	Treated		Non-treated		Diff. Between treated and non-treated	
	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	1.1.2007 - 30.4.2008	1.5.2008 - 31.12.2010	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010
Post-only non-comparison design	1 198 €	2 365 €	1 412 €	538 €	214 €	1 827 €
Exact matching	1 042 €	2 526 €	3 770 €	3 941 €	4 812 €	6 467 €
Propensityscore exact matching	- 856 €	262 €	1 212 €	645 €	- 2 068 €	- 3 317 €
Propensity score nearest neighbour matching	- 1 012 €	- 2 720 €	2 406 €	778 €	- 3 418 €	- 3 498 €

### 5.8.7 Identification of the successful target group for self-employment

This sub-chapter should interpret the successful target and eligible group of the intervention. In the tables below the text are presented the values that are the outcomes of the analysis. Positive values represent a higher net-effect of the treated in comparison to the controls; for easier orientation, blue and red bar charts were added into the cells. Also, the tables contained on the right side results of the statistical test the null hypothesis: the means of treated and non-treated individuals is the same.

From a gender point of view, we identified across the reference periods insignificant differences between the performance of treated and non-treated units. In the first reference period, women were more successful in placement on the labour market, in the second reference period it was men. As stated in one of the previous sub-chapters, age and gender were tested as being insignificant characteristics of the jobseekers in relation to placement on the labour or open market.

Gender	Ref. period 1	Ref. period 2	Ref. period 1	Ref. period 2
	Difference of means of Placement on LM between treated and non-treated individuals		Test of the difference across categories of Treated and Non-treated	
men	0,11	0,10	reject	reject
women	0,13	0,06	reject	reject

In first reference period, widows were the most successful category of **marital status**, but this category is not created on average in about 1 % of all samples and, in the second reference period, the difference between treated and non-treated widows is insignificant. In both reference periods, divorced treated individuals remained about 15 % longer employed than their controls.

Marital status	Ref. period 1	Ref. period 2	Ref. period 1	Ref. period 2
	Difference of means of Placement on LM between		Test of the difference across categories of	
divorced	0,16	0,15	reject	reject
single	0,06	0,07	reject	reject
widow	0,39	0,03	reject	retain
married	0,14	0,09	reject	reject

The most successful category of treated jobseekers were graduates of lower secondary professional education. Overall, the longest time sustained employed on average jobseekers were those with the highest secondary level of education. They remained about 15 % longer than non-treated jobseekers. Just remember that the most effective groups in the traineeship evaluation were jobseekers with achieved higher, tertiary, education.

Level of education	Ref. period 1	Ref. period 2	Ref. period 1	Ref. period 2
	Difference of means of Placement on LM between treated and non-treated		Test of the difference across categories of Treated and Non-treated	
primary school	0,10	0,11	retain	reject
secondary vocational school	0,11	0,07	reject	reject
vocational school	0,13	0,10	reject	reject
comprehensive school	0,14	0,11	retain	reject
colledge	0,02	0,03	retain	reject

Probably, that relates to the category of economic activity of self-employment. Almost 70 % of self-employed jobseekers established a business in construction, services in repair of vehicles or manufacturing. More than every 10th treated jobseeker started to work as real estate agents. Especially, this economic activity has been identified in the survey as the occasion which was offered to jobseekers during job interviews with big real estate agencies. Jobseekers agreed that they would take the grant for self-employment establishment and would work for these real estate agencies.

SK NACE 2 digits	SK NACE 2 digits	Mean	N	Cumulative percent
Stavebníctvo	Construction	,43	980	29%
Veľkoobchod a maloobchod; oprava motorových vozidiel a motocyklov	Wholesale and retail trade; repair of motor vehicles and motorcycles	,43	703	49%
Priemyselná výroba	Manufacturing	,47	649	68%
Činnosti v oblasti nehnuteľností	Real estate activities	,44	405	80%
Ostatné činnosti	Other activities	,38	202	86%
Administratívne a podporné služby	Administrative and support services	,44	148	90%
Ubytovacie a stravovacie služby	Accommodation and food services	,37	103	93%
Informácie a komunikácia	Information and communication	,51	99	96%
Doprava a skladovanie	Transport and Storage	,36	46	97%
Vzdelávanie	education	,37	37	98%
Finančné a poisťovacie činnosti	Financial and insurance activities	,37	29	99%
Umenie, zábava a rekreácia	Arts, entertainment and recreation	,40	25	100%
Zdravotníctvo a sociálna pomoc	Health care and social assistance	,25	4	100%
Dodávka vody; čistenie a odvod odpadových vôd, odpady a služby odstraňovania odpadov	Water supply; cleaning and waste-water treatment, waste management and remediation activities	,90	2	100%
Average		,44		

The highest net-impact was achieved in the group of individuals that were unemployed for more than 3 years and in the central and eastern regions of Slovakia. Bratislava region had the lowest level of net-effect in placement on LM. In Bratislava, no significant differences between treated and controls were identified that could be related to a kind of non-quantification variable as being the motivation of jobseekers in the region with the lowest unemployment rate and highest living standard.

Unemployed before 2007	Ref. period 1	Ref. period 2	Ref. period 1	Ref. period 2
	Difference of means of Placement on LM between treated and non-treated		Test of the difference across categories of Treated and Non-treated	
no evidence	-	-	-	-
< 1 year	0,06	0,09	retain	reject
1 - 3 years	0,10	0,07	reject	reject
> 3 years	0,15	0,13	reject	reject

Region of permanent residence	Ref. period 1	Ref. period 2	Ref. period 1	Ref. period 2
	Difference of means of Placement on LM between treated and		Test of the difference across categories of Treated and Non-	
Bratislava region	0,10	0,00	retain	retain
Trnava region	0,11	0,01	retain	reject
Trenčín region	-0,03	0,10	retain	reject
Nitra region	0,07	0,09	retain	reject
Žilina region	0,05	0,12	retain	reject
Banská Bystrica region	0,20	0,12	reject	reject
Prešov region	0,17	0,06	reject	reject
Košice region	0,10	0,11	reject	reject

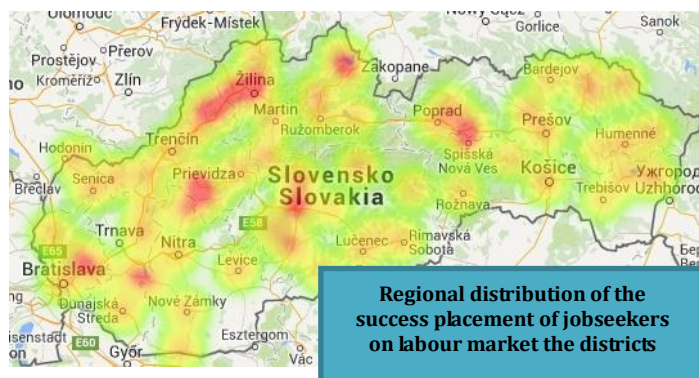
Age	Difference of means of Placement on LM between treated and non-treated		Test of the difference across categories of Treated and Non-treated	
	Ref. period 1	Ref. period 2	Ref. period 1	Ref. period 2
<= 25	0,31	0,08	reject	reject
25 - 35	0,05	0,09	reject	reject
35 - 45	0,13	0,07	reject	reject
45 - 55	0,14	0,12	reject	reject
> 55	0,12	0,03	reject	reject

Reference period		placed_on_L M_pomerné	Gender	Age	Marital status	Level of education_10 categories	Level of education_5 categories	Unemployed before 2007 in months	Region of permanent residence
1	non treated	Pearson Correlation Sig. (2-tailed)	-.006	-.031	-.045	0,147	0,13	-0,158	-0,118
	treated	Pearson Correlation Sig. (2-tailed)	,860	,337	,166	,000	,000	,000	,000
2	non treated	Pearson Correlation Sig. (2-tailed)	,018	,062	,019	,076	0,092	-.043	-.066
	treated	Pearson Correlation Sig. (2-tailed)	,690	,162	,675	,085	,037	,331	,133
	non treated	Pearson Correlation Sig. (2-tailed)	-.021	,002	,044	0,095	0,055	-.1	-0,068
	treated	Pearson Correlation Sig. (2-tailed)	,079	,891	,000	,000	,000	,000	,000
			-.084	,016	,039	0,045	0,042	-0,019	-0,041
			,000	,354	,022	,009	,013	,264	,017

. Correlation is significant at the 0.05 level (2-tailed).

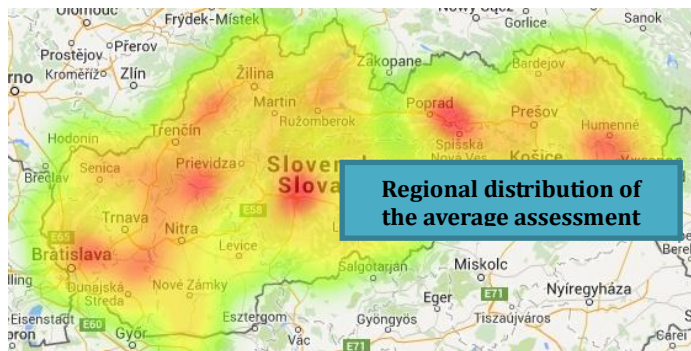
The following figures provide a spatial orientation of the two fundamental outcomes of performances achieved in the impact period in particular LM districts of Slovakia.

The map presents the averages of the achieved assessment base in the impact periods, the red areas show the above averages of the assessment base earned by treated jobseekers. It is possible to see the highest assessment bases were achieved in Zvolen, upper Považie, Senec district, Levoča, Spišská Nová Ves, generally western parts of Slovakia.



The next map also presents districts with the highest sustainability of participants of the programme in the impact period on the open market. As was the case with the average assessment base, sustainability relates to the average assessment base, except in one area in the east of Slovakia, around Humenné. There are more than average sustainability treated jobseekers on the labour market but they earn below average money.

### 5.8.8 Impact of the self-employment promotion



This part of the evaluation report

describes the estimated average influence of the intervention on the unemployment rate decreasing, or the number of registered jobseekers. Impact is calculated year by year according to average estimated placement on LM as one of the outcome variables. Particularly used for estimation of the impact were the shares of placement on LM of Propensity exact matching period method that were applied to the number of treated jobseekers during these years.

We measured 4 years of impact, which is the reason why the impact is also cumulative and estimated just for the number of treated jobseekers in the years from 2007 to half of 2010. In other words, it means that we calculate with the same jobseekers for four consecutive years. Four years because grants were distributed with the condition that self-employment must be sustained at least two years after the intervention provided.

To emphasise the distortion which occurs without the use of a counter-factual impact evaluation approach, we decided to calculate impact as the gross effect and net effect. The net effect or impact informs us about the real estimated percentage of influence due to the traineeship, i.e. with subtraction of the effect which would occur if the intervention did not exist.

At least the provided grant focused on establishing self-employment decreased the number of registered jobseekers from 0.8 to 8.3 %.

During the years the impact evaluation observed, about 3 – 4 % of the unemployed registered jobseekers and about 0.4 % of the Slovak labour force<sup>6</sup> were treated. The difference between gross and net effect in this case is multiple and differs year by year according to the number of the treated jobseekers in previous years. That is reason we can assume that, without the counter-factual impact evaluation method, impacts would be also multiply overestimated and the method would not make sense.

Additionally, we estimated the annual impact on decreasing the number of all registered jobseekers. The gross effect of the self-employed is from almost 1 to 8.3%, depending on the commutation of the previously treated jobseekers. The net impact on the number of registered jobseekers is lower and achieved values from 0.5 to 1%.

	2007/2008	2008/2009	2009/2010	2010/2011	2011/2012	2012/2013	2013/2014
No. of registered jobseekers (total SR)	248 556	379 553	381 209	399 800	425 858	398 876	373 754
No. of treated jobseekers	10 000	12 000	13 000	4 000	-	-	-
Estimated number of jobseekers placed on LM: gross effect	10 000	22 000	31 137	31 127	20 980	11 756	2 766
Estimated number of jobseekers placed on LM: net effect	1 155	2 299	3 408	3 749	2 593	1 450	341
Gross effect on decreasing no. of registered jobseekers (total SR)	4,0%	5,8%	8,2%	7,8%	4,9%	2,9%	0,7%
Net effect on decreasing no. of registered jobseekers (total SR)	0,5%	0,6%	0,9%	0,9%	0,6%	0,4%	0,1%
Gross effect on decreasing of unemployment rate (total SR)	0,37%	0,81%	1,15%	1,15%	0,78%	0,44%	0,10%
Net effect on decreasing of unemployment rate (total SR)	0,04%	0,09%	0,13%	0,14%	0,10%	0,05%	0,01%

Source: Statistics office of Slovak Republic, authors

6 i.e. denominator of the unemployment rate equation.



Finally we can conclude that the intervention had an annual net effect on the unemployment rate decreasing from 0.04 % up to 0.14 % during the impact period. These numbers may appear to be low but we must assume that, annually, the policy covered about 0.4 % of the population and we measure net impact, which is the difference between average performance of the participants and non-participants of the programme for self-employment promotion. The intervention had an influence on the decrease in the unemployment rate and made sense for unemployed jobseekers.



### Financial impact of the self-employment promotion

The intention of this part of the evaluation was to estimate the overall financial impact of the ALMP measure, taking into account all the participating individuals. We counted with the numbers from the performed cost-benefit analysis.

The table below, composed from the two parts, first tells us about the financial effect of the intervention according to gross effects, and the second part refers to the financial impact, which considers the net effects. We estimated that treated individuals were able to bring to the national budget about 75 mil. Euros across the reference periods.

If we consider the estimated net effect of the intervention, the participants of the self-employment promotion generated for the national budget about 2 times less money than the same eligible jobseekers. This means that the treated jobseekers brought to the national budget about 140 mil. Euros more than non-treated jobseekers in total for all reference periods.

	Treated_gross effect in two years of impact period			Net-effect in two years of impact period		
	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	In total	1.1.2007 - 30.4.2008	1.5.2008 - 30.4.2010	In total
Propensity score nearest neighbour matching	- 1 012 EUR	- 2 720 EUR	- 1 866 EUR	- 3 418 EUR	- 3 498 EUR	- 3 458 EUR
No. of treated jobseekers in ref. period	13 650	26 486	40 136	13 650	26 486	40 136
Total effect on national budget	- 14 000 000 EUR	- 72 000 000 EUR	- 75 000 000 EUR	- 47 000 000 EUR	- 93 000 000 EUR	- 139 000 000 EUR
Estimated annual financial effect	- 7 000 000 EUR	- 36 000 000 EUR	-	- 23 500 000 EUR	- 46 500 000 EUR	-

## 6 Strengths and weaknesses of the evaluation

Every type of research has its strengths and weaknesses in the conditions in which it was implemented and carried out. The reason for identifying the strengths and weaknesses is to ensure the highest values of notice and documentary objective and reliable information in relation to the evaluation issues.

### 6.1 Strengths

- Before drawing up this evaluation report, a MLSAF SR pilot impact evaluation was prepared that has helped to identify the control sample, particularly for the self-employment. Lessons learned have greatly contributed to the design of the evaluation itself in the field of data sources as well as the identification of results, impacts and the very feasibility of the quasi-experimental designs.
- The most important strength of the evaluation is the availability of the database of the jobseekers from COLSaF and SIA in a time series of 7.5 years. That is credible secondary administrative data about the performance of the ALMP.
- The evaluation was carried out with the highest possible size of samples (the traineeship had more than 130 thousand jobseekers and the self-employment promotion more than 30 thousand individuals).
- For the data, four different methods were used, which ensured a higher degree of certainty of the estimated effects.
- Outcome variables estimated the influences of the individual barriers of individuals for entrance onto the labour market.
- The outcomes of the method indicate the same impacts, which show the higher credibility of the carried out evaluation.
- This report presents the first counterfactual impact evaluation of traineeship and support of self-employment which has by more methods identified the net effect of jobseekers placement in the labour market as well as the impact on the national budget and the overall unemployment rate.
- The impact evaluation is a combination of two basic approaches: theory based impact evaluation and counterfactual impact evaluation. The first approach has provided a space to identify the causes of the intervention's lowered efficiency and led to recommendations for the policy makers. The second approach is mainly based on quantitative analysis, confirming the effectiveness and efficiency of evaluated interventions.
- The results and knowledge gained from this project could be used in the new programming period 2014-2020 to evaluate the operational programme Human Resources 2014-2020 impacts at the end of the programming period.

## 6.2 Weaknesses

- Through the database of SIA it is not possible to be completely assured that the supported people were successfully placed and maintained jobs in the labour market or continued in self-employment. It is therefore likely that all the results referring to success are slightly underpowered and, in fact, perhaps the percentage of referred persons was slightly higher than in reality. In other words, this fact has led to slight underestimation of the intervention's impact and the net effect of self-employment.
- The availability of very relevant data such as SK NACE of controls and identification numbers of organisations would be a key matching variable for databases that would test the financial and economic conditions of the jobseekers. Vplyvom absencie tejto informácie nebolo možné identifikovať, v ktorých oblastiach SK NACE by bolo vhodné klásť dôraz na podporu SZČ. Due to the absence of this information, it was not possible to identify the SK NACE fields in which it would be appropriate to stress the self-employment support within the next interventions. In the implementation of the forthcoming evaluations we recommend to focus on data from the Financial Administration of the SR or to perform surveys on control group.
- The provided surveys were carried out with assistance of non-representative samples. The samples are relatively small but gender-stratified. A smaller sample could have caused certain unquantifiable misinterpretation. Ideally, it would be appropriate to perform a qualitative research on 380 jobseekers.
- The data does not cover self-employed units that were not obliged to register with the Social Insurance Agency, because they did not reach the conditions stated by the Act about the minimal assessment base. Again, the fact has led to a slight underestimation of the intervention's impact and the net effect of self-employment.

### 6.3 Lessons learned

Although the CIE has in time of its realisation discovered more difficulties in data set selection, final database creation, it had a positive effect leading to creation of primary cooperation between employers responsible for databases of the Social Insurance Agency and COLSAF. For the project it is obvious that for such analyzes in the Slovak Republic it will be necessary to ensure uniform registration systems in different organizations. For example COLSAF has its own information system that is not connected and compatible with the Social Insurance Agency schemes, tax systems, state aid, etc.

Further it was found out as follows:

**Graduate work experience:** The ability to decrease unemployment rate is the undeniable fact of the graduate work experience. We tried to measure this important effect of the intervention and estimated that graduate work experience was annually able to decrease unemployment rate in average by approximately 0.2 - 0.3%. We must take into account that if the intervention was successful for 100% and every participant would be placed on the labour market after graduate work experience, the decrease in unemployment rate would be double.

To realize this fact, the graduate work experience is relevant for graduates and it was one of the most important measures of the ALMP aimed at activating young jobseekers as a part of the perspective work force. Another argument could be the overall financial impact that has been estimated to 540 million Euros over the reference period. In other words, the contribution of the programme participants to the national budget was approximately 540 million Euros higher than the possible contribution of similar eligible jobseekers not supported by any ALMP measure.

**Self-employment:** Outcomes of the evaluation clarified that the intervention is rather a measure of ALMP ensuring placement of participants out of the jobseekers evidence than a real tool to promote self-employment as a way of activating work force of wide group of jobseekers. Participants considerably prefer placement in full-time jobs to continuation of their self-employment business. It is obvious that the intervention should be designed for much deeper support for the programme participants.

The evaluation of the self-employment and graduate experience indicates that both interventions significantly affect the placement on the open labour market.

The outcomes of the evaluation uncovered a motivation force for women. Slovak women frequently have barriers to entry into the open market by establishing their own business, mainly due to their families and limited time they can dedicate to work. Women are generally not motivated to become self-employed. Significant differences were identified in the sample analysis between women who participated in the programme and those who did not participate. While there is one woman not participating in the programme standing for three men; there is one participating woman for one participating man. Indicatively, the grant is one of the forces that can change the attitude of women in the decision making process to start a self-employment business.

The process of the evaluation required the participation of relevant institutions such as the COLSAF and the Social Insurance Agency. The evaluators identified relevant data with these institutions that would help to estimate the net effect of the intervention. Qualitative research was made with respondents that passed the intervention in the

relevant time period. Generalizing the opinion of all respondents, it can be assumed that both interventions raised positive feedback and emotion in participants' minds.

The results of the CIE were disseminated during the workshop organised on the 28. May 2015 with a participation of Managing Authorities. They were informed about CIE methods used and evaluation results. At the workshop were also deputies of the Central Coordinating Body and Slovak evaluation experts.

## 7 Conclusions and recommendations

This chapter presents the final conclusions and recommendations of the entire evaluation report; it contains the most important and most interesting findings, consequences, conclusions and recommendations that should be topical for policy makers and implementation bodies of active labour market policy measures.

### 7.1 Conclusions

#### 7.1.1 Traineeship

Evaluation of the traineeship was carried out with the size of sample which represents more than half of the participants. The most robust method of counter-factual impact evaluation estimated the net effect with the assistance of 16 % of all the participants that were enrolled and intervened upon. In total, we used the registration of more than 131 thousand young eligible jobseekers that were supported and not. The evaluation considered more than 6 years of implementation of this measure of ALMP. Answers gained from the interviewed participants of the program confirmed that their aspirations of traineeship met with the objective stated in the act.

Non-participants of the program are eligible jobseekers that were not treated before and during the evaluation period by any other ALMP measure to eliminate the effects of other interventions.

One of the most important lessons which it was possible to learn due to this evaluation report was the real net effect of the traineeship, which was distributed to young unemployed jobseekers up to 26 years of age. The net effect represents an answer for the fundamental counter-factual evaluation question: Does the traineeship affect employability and sustainability on the open labour market? Or, simply: what would have happened if the intervention had not existed?

The aspiration of the policy makers was to help unemployed young graduates to improve their status on the labour market due to them obtaining the relevant professional skills and practical experience that would be valuable and attractive for employers.

We had the opportunity to learn that placement of young participants was sustained on the open labour market for a significantly longer time than non-participants that did not receive any other intervention of the active labour policy measures. We measured that the placement on the labour market during the impact period of 2 years after the intervention had been correctly complied with. Depending on the method which was used for estimation, participants of the traineeship on average stayed up to half a year longer employed than those eligible jobseekers that did not want to attend the traineeship. Non-participants of the programme were able to stabilize their position on the open labour market better than participants, because every second non-participant that was even once employed in the 2 years impact period remained employed for both these two years. But, on average, every 8th participant of the traineeship sustained employment during the entire impact period of 2 years. The overall effect in placement of participants is significantly higher because three participants from four were employed even for one month in the impact period, while in the non-participants group, three from five individuals were not employed.

Mostly graduates were placed in full-time jobs; very rarely did they have interest in

becoming self-employed, which was considered as another type of placement on the labour market. Registration in the Slovak Insurance Agency of part-time jobs was considered as a not fully placed jobseeker on the labour market. According to the results, in most cases and methods, the participants of the traineeship were more strenuous and they were able to find part-time jobs on average for a longer period than their peers. The independence tests confirmed in the last three reference periods a significant positive treatment effect on participants' placement in part-time jobs due to the intervention.

Also in the study, the probability that a barrier occurred in the individual units which would create barriers for entrance into the labour market was measured. These could be, for example, care for a child, receiving a disability pension, being a personal assistant, etc. These types of registrations indicate to us that granted jobseekers were forced by a life event to stay out of the labour market on average for a very similar time in the impact period to non-participants. On average, there exists a 4 % probability that a participant/non-participant will be exposed to an individual barrier preventing entrance into the labour market.

Jobseekers that attended traineeship earned, on average during the 2 years long period after intervention, from 430 up to 500 Euros per month, depending on the specific year. Graduates that were participants of the program for traineeship earned on average a bit more than half of the average gross nominal wage in Slovakia during the first two years of working. However, from the values, this was obviously connected to the average wage being slightly increased over the years. The evaluation uncovered generally significant negative differences between the participants and non-participants of the program. Just to simplify, those unemployed and registered graduates that attended the traineeship earned on average from 30 to 80 Euros per month more.

The overall financial influence of one individual participant was measured through cost-benefit analysis. The analysis considered items such as paid unemployment allowance benefit in material need, grant, health and social insurance, taxes paid from consumption, or income. In the first reference period from 1st January 2007 until 30th April 2008, when the financial balance was very positive, on average one participant was able to return the provided grant and also bring in some extra money (about two thousand Euros over 2 years) due to the saved allowances and paid taxes. In the next three reference periods, the financial balance of participants became negative also due to the weaker power of placement on the labour market. Correlation confirmed that graduates that were registered as jobseekers for a shorter time earned, in the impact periods, a significantly higher wage. Another aspect which influenced the negative financial balance of the participants in the cost-benefit analysis was the change in the average amount of the provided grant, which increased three times from the first reference period in 2007, from a value of almost 350 Euros for the whole traineeship period. In the last three reference periods, which started 1st May 2008 until 30th April 2012, the novelization of Act No. 5/2004 Coll. stated that the provided grant would be calculated based on the level of the living wage. But, in general, treated participants of the traineeship program brought in to the public budget more money than non-participants. This means that even though the grant was not returned back to the budget, the counter-factual situation that, if the intervention had not existed, it would influence the public budget much more dramatically on average. Overall, it is possible to quantify that one participant brought to the budget about 5,000 Euros more than one non-participant in the 2 years after the traineeship finished.

Last but not least, the goal of the evaluation was try to identify the characteristics of the groups that achieved the most significant positive net effect. We decided to identify these

characteristics through their success of sustainability on the labour market in the impact period. The most successful participants of traineeship were women, about 19 or 23 – 24 years of age, with single or married marital status, with the highest level of education (graduates of college), registered less than three years before the year 2007, and inhabitants of the west Slovakian regions. Additionally, outcomes indicate that those jobseekers that carried out traineeship at private companies had a slightly higher chance of being employed for a longer period over the following 2 years after the interventions finished.

The evaluation provides identification of the participants that were the most successful in sustaining a placement on the open labour market, or open market as self-employed units. The evaluation confirms that age and gender do not influence the placement of jobseekers that intend to start with self-employment. More than average were the successful divorced jobseekers that achieved the highest education level – secondary, individuals that started self-employment in water supply, waste management, IT sector or manufacturing. Almost three participants from four established self-employment in an economic activity such as construction, repair of motor vehicles, manufacturing, or real estate. The named categories of economic activities brought average or more than average results in keeping individuals employed. Very effective were individuals that had been for the long-term outside of the labour market, i.e. unemployed for more than 3 years before the year 2007.

One undeniable fact of the traineeship is its ability to decrease the unemployment rate. We tried to measure as well this important effect of the intervention and we estimated that, annually, on average, the traineeship was able to decrease the unemployment rate by about 0.2 – 0.3 %. We must take into account that, if the intervention was 100% successful and every participant were to be placed on the labour market after traineeship, the unemployment rate would decrease twice as much again. To realize this fact the traineeship has a sense for graduates, and it was one of the most important measures of ALMP that was targeting young jobseekers to activate them as a perspective work force. The other argument could be the overall financial impact, which has been estimated on the level of 540 mil. Euros over the reference period. In other words, the participants of the program brought to the national budget about 540 mil. Euros more than would have been brought by the same eligible jobseekers had they not been supported by any ALMP measure.

### **7.1.2 Self-employment**

Also for estimation of the net effect of the self-employment promotion, a large sample of jobseekers was used. We used evidence of less than every second participant of the jobseeker programme that were encouraged by financial grant to establish self-employment. The evaluation covered the period from the start of 2007 to the end of April 2010, in total 40 months of distribution of disposable grants for jobseekers that applied for intervention, carried out a financial business plan and were registered in the evidence of jobseekers at the PES office for more than 3 months. For the most rigorous method, we used on average every 10th participant of the program and in total for evaluation methods we used samples of more than 30 thousand individuals that were participating in the program and units that did not take the grant and were eligible, but self-employed. Estimation of the net effect of the intervention is the fundamental objective of the carried out evaluation of self-employment. The net effect of the intervention should be sustainable placement of a participant of the programme on the open market or labour



market. Conditions of the intervention state that every participant must be self-employed minimally for two years after the grant is provided. For two years after compliance of this condition was the fixed impact period. The impact period is characterized by the participants being out of intervention duties. It is a period when participants can decide to be self-employed, find a job or return back to the registration of jobseekers.

The evaluation report estimated whether participation in the programme made sense for eligible jobseekers. Due to the evaluation, it is possible say what the treatment effect of jobseekers that have an interest in being self-employed would be. In the evaluation were compared the performances of the participants and non-participants of the ALMP programme focused on self-employment promotion. To ensure the highest level of comparability of both these groups, non-participants were just individuals that were eligible in the particular reference period and data from SIA confirmed they started self-employment in the impact period of the particular reference period.

The most desired effect of this active labour policy measure is sustainable self-employment of the participants on the open market or placement on the open labour market as an employee in a full-time job, i.e. out of the evidence of jobseekers. That effect is represented in the variable "Placement on the labour market." According to the carried out counter-factual impact evaluation methods, we are able to estimate that, on average, participants remained out of the jobseekers evidence about less than 20 % of the impact period (2 years after sustainability of self-employment). In other words, one participant of the program would be employed about 50 days less if the financial intervention had not been granted.

The most rigorous methods that were performed for estimation of the net-effects show that the programme had a negative effect on the self-employment sustainability of participants. Participants prefer full-time jobs. Non-participants of the programme remained self-employed about one month longer than participants in the 2 years long impact period. Generally, participants as well as non-participants of the programme do not prefer to be placed in a part-time job. That would probably be due to the higher average age of both samples. The previous intervention focusing on traineeship proved interesting to young jobseekers, about the same as part-time jobs.

The situation is very similar to the traineeship: non-participants of the programme were able to find their stable position on the market more easily than participants, because every second non-participant that was even once employed in the 2 years impact period sustained employment for the entire two years. But, on average, every 8th participant of the self-employment sustained employment throughout the 2 year impact period. The overall effect in placement of participants is significantly higher because seven participants from ten were employed even for one month in the impact period, while in the non-participants group it was just about every second.

The evaluation also concentrated on measurement of the probability that barriers occur in individual units which could create barriers for entrance into the labour market. These could be, for example, caring for a child, receiving a disability pension, being a personal assistant, etc. These types of registrations indicate that participants of the programme were forced to stay out of the labour market for a longer time on average than non-participants. An individual barrier occurred in the group of participants much more frequently than in the group of non-participants. While in the group of participants there exists about 10 % probability that individuals will be exposed to a barrier of entrance into the market, while in the sample of non-participants it is just about 0.3 %.

On average, about  $\frac{3}{4}$  of a month in the impact period was a longer period when individual barriers for entrance to the labour market occurred to the participants of the

programme.

Part of the evaluation was devoted to the estimation of the average financial effect that occurred due to the distribution of grants for self-employment promotion. Provided cost-benefit analysis compared the financial flows of one participant and one non-participant according to the average time when they were employed and unemployed. According to the final outcomes, the intervention had a negative effect on the national budget. Both reference periods have estimated a very similar net financial impact on public finance. Provided values estimate that one participant of the programme can generate for the national budget almost 3500 Euros less than a non-participant. On the other hand, in the cost benefit analysis, the provided grant was also calculated. If the grant were not counted, the net impact would be significantly lower (the average provided grant was more than 2900 Euros). In the first reference period it was estimated that one participant earned about 80 Euros per month more than those jobseekers that started self-employment without a grant from COLSaF. In the second reference period one non-participant of the programme earned 20 more than the same participant, but this difference was tested as being insignificant.

The overall estimated financial impact of the intervention shows that the negative influence on public finance to be at the level of almost 140 mil. Euros over the evaluated reference periods as the difference between participants and non-participants, i.e. net financial impact.

We learned from the survey that many of the interviewed participants of the self-employment promotion program of COLSaF welcomed and appreciated the provided training as a preparatory course for self-employment. They answered that the course was rather encouraging for their orientation among offices or in business vocabulary, awareness about business finance and accounting. That supplementary intervention could be evaluated as being very requested. Only the orientation of the courses could be improved (see Recommendation for Self-employment).

The outcomes of the evaluation uncovered a motivation force for women. Slovak women frequently have barriers for entrance into the open market through the establishment of their own business, mainly due to family and the limited time they can devote to the work. Women are not generally motivated to become self-employed. In the analysis of samples were identified significant differences between women who participated in the program and those who did not. While for one non-participating woman in the program there were three men, for one participating women there was one participating man. Indicatively grand is one of the powers which can change the attitude of women in the decision making process to start with self-employment. In the programming period 2014-2020, the operation program Research and Innovation is incorporating into the structure of the activities "social innovation", which operates by encouraging women into the business establishment. Participants remained employed for a longer time mostly in the western districts, but not in Bratislava, Trnava or Trenčín districts, where there is a lower unemployment rate. That could occur due to the influence of a weaker motivation to employ jobseekers that live in the environment of a higher living standard.

Finally, the most important effect of the self-employment promotion is the impact on the decrease of the unemployment rate. The analysis of the impact on the unemployment rate proved the justification of the intervention, which decreased the unemployment rate annually from about 0.1 % up to 0.14 %. These values represent impacts that occurred mainly due to the intervention. The numbers clearly show the effect which would have occurred if the intervention had not existed. The impact might seem too low but we must consider that, annually, only about 0.4 % of total labour force in Slovakia were exposed to

the intervention.

Self-employment could be a perspective measure of the ALMP that must be reformed into a more complex tool which could provide the participants with more than just basic general information about business, but instead additional services that would aid the competitiveness of the self-employed units.

Outcomes of the evaluation clarified that the intervention is more of a measure ensuring the placement of the participants out of the evidence of the jobseekers register than a real tool for the promotion of self-employment as a means how to activate the work force of a wide group of jobseekers. The participants significantly prefer to be placed in a full-time job than continue in business as self-employed.

## 7.2 Recommendations

1. The Ministry of Labour, Social Affairs and Family of the SR should begin to carry out regular surveys of active labour policy measures. COLSaF could distribute forms to all participants after the provided intervention. Every participant should evaluate the whole process and time period of the intervention and activities which were provided to him/her. Information could be collected through an on-line form. That would be a unique source of valuable information. There should be simple questions focused on the topics in the carried out Qualitative research of both active measures. And the form should contain an open space for the statements of participants. The information should be electronically recorded and there should be prepared a *modus operandi* for analyses which should be provided to the policy makers and experts for methodology. Additionally, it would be welcome to analyse difficulties which appeared during the activities of the intervention or in the sustaining period after the intervention (2 years obligatory sustainability for self-employment after the grant was provided). It is important to emphasise that very valuable information and lessons could be provided through analysis of the reasons for why the treated failed; for instance, why did the self-employed close their business after the minimal claimed 2 years sustainability period?

### 7.2.1 Traineeship

2. In the performed survey, about 10 % of program participants admitted that they worked in the business which fit with the type and specialization of the education they had completed. Most of the participants carried out their traineeship in public sector organizations (mainly in public offices, education, health or social organisations), the rest of the participants carried out their traineeship in private companies with a slightly higher propensity to be placed on the labour market with greater sustainability in the impact period after the intervention finished. COLSaF should actively search for companies and organizations that would be a better match for the participant's profession. Graduates should have experience in the branch in which they studied and graduated. That would be ensured through transparent and clear categorisation. COLSaF should be encouraged to create such an electronic system which would identify the economic nomenclature of the organisation for a particular group of professions.
3. Four-hours working time appears to be insufficient according to the multiple opinions of the program participants. They claim that the working time was insufficient to manifest their capabilities. The policy maker could start a pilot

- programme with a prolongation of working days.
4. Participants identified the need to gain something tangible through the carried out traineeship, some **recommendation, or certificate**, which could enforce their positions as seeking jobseekers in a job interview and which would upgrade the intervention to a more serious level. Treated jobseekers would be more likely to seriously make an effort to gain the chance of a job by preparing as much as possible.

### 7.2.2 Self-employment

5. Self-employment is a rather wide topic which is exposed to a number of influences which determine the success of self-employment. There are some aspects from the open market which decide whether the established business crosses “death valley” as one of the most important initiative stages from the business cycle of any start-up. That is the reason why it is necessary to provide participants of the intervention with the relevant information which would ensure a healthy start and sustainability of their self-employment, i.e. desired status.
6. It would be helpful to collect and analyse the problems of the self-employed and create a FAQ or account on a social network site which would represent the place where some information would be published about the support for the self-employed, or start cooperation with the Slovak Business Agency, which is the body responsible for development of micro, small and medium sized enterprises, with the National Business Centre currently in the process of preparing a network of regional affiliations to be closer to regions. According to the responses of interviewed participants of the intervention, they would welcome some soft support, some of the entrepreneurs would like to receive support such as expert counselling, legal counselling, marketing counselling, market experts, accounting counselling, graphics ensuring transmission information about additional funding of the business plans through grants or non-grant schemes, etc.
7. The survey showed that participants of the programme would welcome some legal assistance in case of bad debts, mainly in the construction sector, which is a frequent profession of treated jobseekers. These self-employed have a problem to gain money and that is also a reason for their failure. The policy maker could build self-employment promotion as a stronger measure of ALMP. The measure should be really active and should reflect the actual needs of the programme participants.
8. Policy makers could pilot an introduction of the selective intervention for jobseekers that have not had any experience with self-employment or with another form of entrepreneurship (by using a limitation of the retrospective assessment of the distinguishing criteria). The treatment should be much more complex, mainly for first time participants of the programme.
9. Respondents see as a limitation that they must buy exactly the same items they proposed in the approved financial plan enclosed with the business plan. The procurement of items in the financial plan is carried out with a time gap and meanwhile a more achievable product with higher efficiency might appear. That is why respondents propose more flexibility in the changing types of procured items.
10. Even the obligatory preparation course concerns on preparing the jobseeker for self-employment are very positive and helpful. However, there are some points which could still be improved. The individuals could be segmented according to achieved highest level of education, or type of education. Participants who are for the first time encountering some economic categories are mixed together on

courses with other participants to whom these subjects are very familiar. Interviewed participants felt that mixing was very limiting. It would be desirable to create at least two types of courses.

11. The characteristics of the programme participants revealed that intervention was distributed to jobseekers in retirement (in some cases more than 70 years of age). It is not obvious what exactly the aspiration of the policy is for these jobseekers who are still eligible. We propose limiting the age of eligible jobseekers for self-employment.

### 7.2.3 Monitoring of the relevant data

Relevant and correct data is a fundamental part of the counter-factual impact evaluation and a building piece of the policy-based evidence which it is desired to build. That is the reason why the implementation body and policy maker should devote intensive effort to form databases which would be useful for analysing the effects which occur due to the distributed intervention.

12. It would be helpful for COLSaF to collect data which would logically complement each other on different levels. There were identified many inconsistencies among the level of education of jobseekers, types of schools and fields of specialisation. There should be implemented a controlling mechanism which ensures that one jobseeker that has recorded primary school as their highest achieved education and the last graduated school is an university. We advise using one official nomenclature to unify the recording of data.
13. In the database of COLSaF there were identified numerous groups of records incomplete in some basic characteristics such as age, or gender. Other provided variables were also without records, which occur in quite a robust selection of the samples.
14. It is worth considering initiating the creation of direct linkages between COLSaF and SIA, to supply data which has already been recorded in SIA. This could ensure a simplification of the work at regional PES offices and overlapping in the work of managers and officers. This data should be unified via a common methodological manner.
15. SIA does not register the identification number of the business (IČO) of the self-employed, which is important for the identification of the jobseeker in the other official database of the Financial Directorate of SR, which could provide exhaustive information about the financial and economic condition of the firms.
16. It would be very helpful to the data if they were supplied with the address and contact details of jobseekers to enable the creation of a focused group for a qualitative survey. For instance, we could describe why treated graduates remained on the open labour market mostly for almost the entire impact period with a higher frequency than the controls. And why in the group of treated there is a higher frequency and probability of being unemployed for the entire impact period of 2 years.
17. SIA should ensure uniform recording of the data according to official nomenclatures and prevent the use of unauthorized characters in the names of municipalities, such as: ´, @, ®.



## References

- (1.) Act No. 222/2004 Coll. on value added tax, and amendments and supplements to various acts
- (2.) Act No. 461/2003 Coll. on social insurance, and amendments and supplements to various acts
- (3.) Act No. 5/2004 Coll. on employment services, and amendments and supplements to various acts
- (4.) Act No. 580 / 2004 Coll on health insurance, and amendments and supplements to various acts
- (5.) Act No. 595/2003 Coll. on income tax, and amendments and supplements to various acts
- (6.) Act No. 599/2003 Coll on assistance in material need, and amendments and supplements to various acts
- (7.) Active measures of labour market policy – statistics: [http://www.upsvar.sk/statistiky/aktivne-opatrenia-tp-statistiky.html?page\\_id=1248](http://www.upsvar.sk/statistiky/aktivne-opatrenia-tp-statistiky.html?page_id=1248)
- (8.) BORIK – CABAN: Pilot assessment of the impact of selected measures of the active labour market policy ALMP. In Bratislava: MPSVR SR, ÚPSVR; 2013.
- (9.) Dagmar Gombitová: Úvod do monitorovania a evaluácie. In Bratislava, 2007, ISBN978-80-96960-9-5.
- (10.) European Commission: “Think Small First”: A “Small Business Act” for Europe, 2008.
- (11.) European Commission: Design and commissioning of counter-factual impact evaluations: A practical guide for ESF managing authorities. Luxembourg: Publication Office of the European Union, 2013, ISBN 978-92-79-28238-6.
- (12.) Harvan Peter a Machlica Gabriel: Trh práce v krízovom roku 2009. In Bratislava: Inštitút finančnej politiky, Ministerstvo financií SR, 2010.
- (13.) The Nonequivalent Groups Design: <http://www.socialresearchmethods.net/kb/quasnegd.php>
- (14.) Kulhavý. V. – Sirovátka. T.: Hodnocení efektivity programů APZ a doprovodných nástrojů a projektů politiky zaměstnanosti v roce 2007. In Praha: Výzkumný ústav práce a sociálních věcí, v.v.i., 2008; ISBN 978-80-7416-012-7.
- (15.) Linda G. Morra Imas and Ray Rist: The Road to Results: Designing and Conducting Effective Development Evaluation. In Washington, The World Bank. June 2009, ISBN: 978-0-8213-7891-5.
- (16.) Matrix of distances among Slovak towns and villages: <http://frdsa.fri.uniza.sk/~janosik/>
- (17.) Shahidur R. , Khandker, Gayatri B, Koolwal, and Hussain A. Samad: Handbook on impact evaluation: quantitative methods and practices. In Washington: The World Bank, 2010, ISBN 978-0-8213.8028-4.
- (18.) The Canadian Journal of Program Evaluation Vol. 24 No. 2 Pages 31–56 ISSN 0834-1516 Copyright © 2010 Canadian Evaluation Society Decree of the Ministry of Labour, Social Affairs and Family No. 181/2012 Z.z. Decree of the Ministry of Labour, Social Affairs and Family No. 194/2011 Z.z. Decree of the Ministry of Labour, Social Affairs and Family No. 300/2010 Z.z. Decree of the Ministry of Labour, Social Affairs and Family No. 252/2009 Coll. Decree of the Ministry of Labour, Social Affairs and Family No. 225/2008 Coll.
- (19.) Řeháková, B.: Nebojte se logistické regrese. [http://sreview.soc.cas.cz/uploads/5f6961faa17dd98a67cfb71a5205469d297369f5\\_372\\_475REHAK.pdf](http://sreview.soc.cas.cz/uploads/5f6961faa17dd98a67cfb71a5205469d297369f5_372_475REHAK.pdf)
- (20.) Šoltés, E.: Regresná a korelačná analýza s aplikáciami. In Bratislava: Iura Edition, s.r.o., 2008, ISBN 978-80-8078-163-7.
- (21.) Stankovičová, I., Vojtková, M.: Viacrozmerné štatistické metódy s aplikáciami. In Bratislava: Iura Edition, s.r.o., 2007, ISBN 978-80-8078-152-1.

## List of abbreviations

ALMP	Active Labour Market Policy
COLSaF	Central Office of Labour and Social Affairs
FAQ	Frequently Asked Questions
IČO	Personal Identification Number
ISCO	International Standard Classification of Occupations
No	Number
PES	Public Employment Services
SIA	Social Insurance Agency
Sig.	Significance
NUTS	Nomenclature of Units for Territorial Statistics
SK NACE	Classification of Economic Activities