

Laura Južnik Rotar

University of Novo Mesto,
Faculty of Economics and Informatics,
Department of Economics,
Novo Mesto,
Slovenia

✉ laura.juznik-rotar@guest.umes.si

What Are the Treatment Effects of a Work-First Participation Programme on Young Unemployed People in the Netherlands?

Summary: This paper evaluates the effects of the employment programme on young unemployed people in the Netherlands. The effectiveness of the programme is measured by probability of both re-employment and participation within the regular educational system. This evaluation is made in comparison to that of an individual who would continue seeking employment as an openly unemployed person. The effects of the programme are evaluated a year/two years following the start of the programme. We apply a propensity score matching method. The identification of an average treatment effect is based on the conditional independence assumption. The effects on re-employment probability and the probability of participation in the regular educational system are statistically negative, applicable to both long- and short-term scenarios.

Key words: Active labour market policy, Treatment effect, Propensity score matching, Dutch labour market.

JEL: C21, J64, J68.

Acknowledgement: The research leading to these results has received funding from the European Union's 7th Framework Programme (FP7/2007-2013) under grant agreement n° 262608, DwB - Data without Boundaries.

The Dutch labour market has suffered severe ramifications as consequence of the Great Recession. Unemployment has increased from 3.1% in 2008 to 7.2% by early 2014. Currently, over 600,000 people are unemployed in the Netherlands. One third of all unemployed have been jobless for over one year. Prior to the crisis, the Netherlands had one of the lowest unemployment rates in Europe (Marloes de Graaf-Zijl et al. 2014). The reforms of the early retirement schemes and disability insurance in the Netherlands have been effective. The results of the reforms have been evident in the increased incentives to continue working and in the decreased substitution between early retirement and disability. The empirical research provided by Rob Euwals, An-nemiek van Vuren, and Daniel van Vuuren (2011) investigates the impact of the reforms on labour market exit probabilities. They conclude that the reforms have proven to be progressive, since the participation rate of elderly workers has increased. De Graaf-Zijl et al. (2014) also point out the institutional setting connected with labour market institutions. Current disability benefits are less generous than previous welfares, whilst the option of early retirement has been made financially less viable.

In order to encourage young people to search for a job and be active in the labour market, the Dutch active labour market policy plays an important role. Some of the evaluations of active labour market programmes reveal both positive and negative effects of such programmes. There is a general consensus which indicates that there is a lack of evaluations conducted by rigorous evaluation methods (Pieter A. Gautier and Bas van der Klaauw 2009). Our empirical research is particularly relevant since it is based on the evaluation of the effects of the employment programme on young unemployed Dutch people applying a propensity score matching method. Following this section, we present a literature review in Section 1. Section 2 introduces empirical research. The findings of the empirical research are discussed in Section 3. Finally, the closing section is reserved for a conclusion.

1. Literature Review

Previously, the Netherlands introduced some major reforms and achieved a reduction of generosity of unemployment insurance. The country also reduced the number of benefit recipients and heavily financed their disability insurance programme. Implementation of social assistance was transformed from a centralised system to a completely decentralised scheme with an active role in commercial agencies. The latter resulted in high diversification of the Netherlands active labour market policy. The success of labour market programmes depends not only on the management of programmes, but also on general economic conditions as well (see, for example, Organisation for Economic Co-operation and Development - OECD 2008; Sergej Gričar and Štefan Bojnec 2013). The evaluation results of some studies found in the literature are presented hereinafter. Childcare subsidies and labour supply have been analysed by Leon J. H. Bettendorf, Egbert L. W. Jongen, and Paul Muller (2012). Between 2005 and 2009, the Dutch government increased childcare subsidies substantially, reducing the average effective parental fee by 50% and extended subsidies to so-called guest parent care. Bettendorf, Jongen, and Muller (2006) estimate the labour supply effect of this reform with a difference-in-differences strategy, using parents with older children as a control group. Bettendorf, Jongen, and Muller (2006) found that the reform had a moderately sized impact on labour supplies. Furthermore, the effects are an upper bound since there was a simultaneous increase in earned income tax credits for the same treatment group over the same period. The joint reform increased the maternal employment rate by 2.3 percentage points (3.0%). Average hours worked by mothers increased by 1.1 hours per week (6.2%). Bettendorf, Jongen, and Muller (2006) found that such increases in hours were due to the intensive margin response.

Gijs Roelofs and Van Vuuren (2011) write about fiscal decentralization, the decentralization of government expenditures to local governments, which many researchers believe enhance public sector efficiency. The Dutch Social Assistance scheme (SA) was gradually decentralized from the central government to municipalities during the early 21st century. Over the same time period, the influx to the Disability Insurance scheme for individuals with no employment history (DI) more than doubled. By 2008, an exponential share of 1.7% of the Dutch population (between 18-65 years old) were receiving benefits from this scheme. The DI scheme is fully financed by the central government. As found in Roelofs and Van Vuuren (2011) that the reform of

the SA scheme has contributed to the increased influx into the DI. Municipalities may have been tempted to redirect (potential) SA recipients to the DI scheme as a result of their increased financial responsibility for Social Assistance. Jaap De Koning, Arie Gelderblom, and Peter van Nes (2005) report negative effects of training programmes on young unemployed people. The justification of these claims are based on young unemployed people who participated in such programmes and often did not finish school, they were also deterred from returning to school. De Koning, Gelderblom, and van Nes (2005) proposed that training programmes should be offered to young unemployed people who already have a sufficient level of education. Based on Danish data, Michael Rosholm and Michael Svarer (2004) evaluated the threat effect of active labour market programmes and found a significantly positive threat effect which reduced the average duration of unemployment by approximately three weeks. Suzanne Kok et al. (2011) conducted research based on a Dutch Labour Force Survey dealing with participation of migrant women in the labour market, focusing on the culture of both their home and host country. Their results suggest that both differences in home-country female participation and the trend in native female participation affect the participation of migrant women in the labour market. Host-country participation is at least as important as home-country participation. Lex Borghans, Bas Ter Weel, and Bruce A. Weinberg (2014) demonstrated that people skills are important factors of labour market outcomes for underrepresented groups, assuming that gender differences and cultural differences may impede cross-racial and ethnic interactions. The significance of skilled/unskilled labour supplies when determining the direction of technical choices of the countries and cross-country productivity differences is provided by Tuna Dinç (2012). Dinç (2012) constructed a two sector model of productivity differences in which the level of technology is determined endogenously, depending on the aggregate capital externalities. Not only do labour skills affect labour market outcomes, they are also instrumental in determining capital productivity since capital is more productive in the advance of the skilled labour complement technologies than in the unskilled labour complement technologies. Alternatively, Kok (2013) provides research on matching quality in the Netherlands that is measured by the gap between worker skills and their job tasks. Matching between skills and tasks is better when the gap is smaller. Kok's results suggest that measured gaps are smaller in cities in comparison to the Dutch countryside. The observation of better matching qualities are explained by the location of work, but not by location of residence. Such results indicate an association with higher wages and urban wage premiums.

Applying difference-in-differences techniques and using a control group created by exact matching, Anja Deelen, De Graaf-Zijl, and Van den Berge (2014) researched labour market effects of job displacement for prime-age and mature workers. They found that involuntary job loss has a severe impact on older workers' labour market prospects. Finding a new job is relatively difficult, and wage cuts are more substantial once they find a new job. The differences between prime-age and mature workers are partly mediated by tenure and industry effects. Not only do mature workers on average have longer job tenures than prime-age workers, older workers with longer job tenures experience more negative effects of displacement as well. For prime-age workers, tenure in the job before displacement makes less of a difference

on their outcomes after displacement. Likewise, displaced mature workers are more sensitive to the situation in the local labour market in the industry from which they were displaced. Moreover, mature workers experience more prominent negative effects of changing industries after displacement on their post-displacement wages.

The effects of earned income tax credits on the labour supply responsiveness of single mothers in the Netherlands have been analysed by Bettendorf, Kees Folmer, and Jongen (2014). The share of single parents is currently increasing. In the Netherlands, statistics have increased from 360 thousand single parents in 1995 to 500 thousand in 2011. Single parents are of particular interest to policy makers, as demonstrated by the large number of subsidies and tax credits targeted at this group. In designing income support for single parents, the responsiveness of labour participation by single parents to financial incentives plays a crucial role. Until 2001, working single parents received a tax credit, the *Aanvullende Alleenstaande Ouderkorting* (Additional Credit for Single Parents), if the youngest child was below the age of twelve. In 2002, this age limit was raised to sixteen years old. The purpose was to stimulate the labour participation of single parents. The authors use this change in the age limit as a natural experiment to determine the labour supply responsiveness of single mothers to financial incentives. They use difference-in-differences (DD) and regression discontinuity (RD) to estimate the effect of the policy reform on the participation rate of single mothers. In the DD analysis, they use single mothers with a youngest child that is younger (8-11 years of age) or older (16-19 years of age) than the treatment group (12-15 years of age) as the control group. In the RD analysis, the authors focus on single mothers with a youngest child between 14-17 years of age, with the cut-off at single mothers with a youngest child that turned 16 in December of the preceding year, using data on month of birth. Both the DD and RD analysis show that the policy reform had a small effect on the participation rate of single mothers. Based on the administrative data on registered unemployed in Slovakia, Miroslav Štefánik (2014) estimates treatment effects of a training programme using propensity score matching. Štefánik (2014) estimates that the chances of getting a job during the period of fifteen months after undertaking the programme. Performance of this indicator is compared between participants and a control group, which is selected *ex post* using the propensity scores matching approach. The results reveal evidence on negative average treatment effects on the treated, whereas negative effects on the training measure on employability of participants can be assigned to mistakes in its implementation.

2. Empirical Research

2.1 Objectives of Empirical Research

The objective of empirical research is to evaluate the effects of the employment programme on young unemployed Dutch people. In conjunction, the research aims to determine the effects of employment programmes on the probability of re-employment and on the probability of participation in the regular educational system in comparison with the outcome produced in the event that an individual would continue seeking employment as an openly unemployed person. The effects of the programme will be evaluated a year/two years following the start of the programme.

Another objective of empirical research is to apply the propensity score matching method, which is a modern approach to empirical analysis for evaluating treatment effects, proposed by literature in the field of evaluating programmes (see, for example, Marco Caliendo and Steffen Künn 2012; Rajeev H. Dehejia 2013; Martin Huber, Michael Lechner, and Conny Wunsch 2013; Jeffrey H. Silber et al. 2014; Caliendo and Ricarda Schmidl 2015; Markus Frölich and Lechner 2015; Benjamin Schünemann, Lechner, and Wunsch 2015; Gerard J. Van den Berg and Bas van der Klaauw 2015; James J. Heckman, John Eric Humphries, and Gregory Veramendi 2016). The identification of an average treatment effect is based on the conditional independence assumption. This is an assumption in accordance with which participation in the programme is autonomous of after-programme outcomes and conditioned by observed external factors. Considering these assumptions, the evaluations of an average treatment effect for the treated are not bias (see, for example, Reinhard Hujer 2011; Ben Hansen, Paul R. Rosenbaum, and Dylan Small 2014; Heckman and Rodrigo Pinto 2015; Huber, Lechner, and Andreas Steinmayr 2015).

2.2 Data and Programme

The data used in this empirical research come from rich datasets provided by the Central Bureau of Statistics Netherlands. In this research we use rich microdatasets on labour and social security (Central Bureau of Statistics Netherlands - CBS 2016)¹. Since the aim is to evaluate the effects of employment programmes on young unemployed Dutch people, we concentrated on participation in work-first type of programmes for youth aged 20-24 for the year 2008. Following 2008, unfavourable economic development had serious consequences on the labour market. Among different disadvantaged groups in the labour market, young people had been seriously hindered by exponential unemployment rates as they were more vulnerable to being laid-off. Additionally, due to higher competition on the labour market during times of crisis, young people are less likely to be hired for a job when competing for a limited number of job vacancies. Work-first is the collective name for projects where young people have been given priority (although not limited exclusively to young people, this can be applied to all age groups) and they can start work as soon as possible once they have applied for social assistance. The common philosophy adapted is that young people should not be on social assistance, but should be learning and/or working. Work-first includes specific categories of young people with limited or no work experience, such as early school leavers, single parents or young people with disability. Work activities in such circumstances are often part-time and may be implemented through placement in subsidized employment within a third sector agency, a private company or public employment scheme. To deter youngsters from leaving school early the obligation to study or work has been implemented. According to this obligation all youth aged 18- to 27 who have not completed the equivalent of ISCED 3 (so called starting qualifications) are required to resume schooling or to work.

¹ **Central Bureau of Statistics Netherlands (CBS)**, 2016. Labour Market Policy Database. <https://www.cbs.nl/nl-nl/maatwerk/2013/45/labour-market-policy-database-2012> (accessed July 20, 2016).

Young people can be fined or denied (part of) their social benefits or assistance if this term is not abided by. In the Netherlands, the system was radically restructured from a strongly centralised Public Employment Service to a highly decentralised system where implementation of active labour market policy is mostly outsourced to private agencies, the information on the activation programmes is disseminated. Application for social assistance functions in correlation with the obligation to work, which indicates that the age from which young people receive social assistance is moving up (see, for example, OECD 2008; European Commission 2013). Each person used to compile the research data provided the following details: registration dates, labour market status and individual characteristics. In order to conduct the research, two groups of data have been constructed: an experimental group and a potential control group. The experimental group was constructed in a way that from the database of all individuals who have been registered with the Employment Service, we collected individuals aged 20-to-24 who registered with the Employment Services in the year 2008 as openly unemployed. From this group we collected all individuals who after being openly unemployed enrolled in an/the employment programme (participation programme). The experimental group consists of more than 200 observations. Alternatively, the potential comparison group consists of persons aged 20-to-24 who entered in the register as openly unemployed in the year 2008 and never participated in any of the programmes. All of the potential comparison group could be used as the third group of non-participants in the empirical research. However, as discussed in the evaluation literature, the pre-programme unemployment history is instrumental when explaining selection into a programme. In explaining the choice among alternative programmes, the length of the unemployment period immediately before the programme starts/ed is an important factor. In order to use this information for the estimation of the propensities, we created a hypothetical programme start date for non-participants. The procedure followed was similar to the random procedure suggested by Lechner and Wunsch (2013). First, both groups of participants, as well as the group of non-participants, were divided into sub-groups by the month of registration with the Employment Service as openly unemployed. Secondly, each of the non-participants in a sub-group was randomly assigned an observation of “length of pre-programme unemployment” from the distribution of the contemporaneous sub-group of participants. In case the non-participant’s actual unemployment period was shorter than the assigned pre-programme unemployment period, the individual was removed from the sample.

2.3 Application

The first stage in matching is to model the propensity score. First, we do probit estimation to find out how: variables, marital status, number of parents abroad, generation, age, sex, employment status before the programme and, unemployment duration before the programme influenced the participation probability in the employment (participation) programme. Table 1 displays the probit regression model of the likelihood of participation in the programme.

Table 1 Propensity Score Coefficient Estimates

| Variable | Coefficient | St. error |
|--|-------------|-----------|
| Marital status | 0.1737 | 0.1510 |
| Number of parents abroad | 0.1546 | 0.0776 |
| Generation | 0.0039 | 0.0893 |
| Age | 0.0109 | 0.0276 |
| Sex | 0.4091* | 0.0856 |
| Employment status before the programme | -1.3093* | 0.0913 |
| Unemployment duration before the programme | -0.0001 | 0.0001 |
| Constant | -1.5929* | 0.6096 |

Notes: * denotes significance at the 5% level.

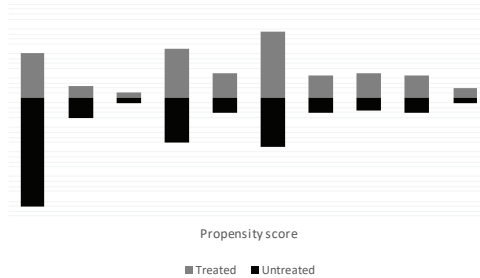
Source: Author's calculation based on CBS (2016) data.

Propensity score matching matches participants in the employment programme with the non-participants based on the propensity score and are conditional on observed characteristics X (Peter C. Austin 2011). The validity of propensity score matching depends on two conditions: (a) conditional independence assumption; and (b) assumption of common support.

Conditional independence assumption states that given a set of observable variables X that are not affected by the treatment (in our case employment programme), potential outcomes are independent of a/the treatment assignment (Christoph Rothe 2016). Conditional independence assumption is also called *unconfoundedness* (Rothe 2016), it is a steadfast assumption and is not a directly testable assumption. This assumption has to be justified by the data quality available to the researcher. Data quality should enable the researcher to identify as many observed characteristics as might influence participation in the employment programme. This process will help justify the conditional independence assumption (Alberto Abadie and Guido W. Imbens 2016). Our rich dataset contains pre-programme information containing variables, such as: employment status before the programme and unemployment duration before the programme. In order to justify the conditional independence assumption, we also follow the recommendations found in Abadie and Imbens (2016). Abadie and Imbens (2016) demonstrate that the bias in propensity score matching can be low if the same survey instrument or source of data is used for participants and non-participants. Data used for this paper is sourced from a rich datasets provided by the CBS. In this research we have used rich micro-datasets on labour and social security (CBS 2016). On the other hand, the precision of the propensity score can be improved if we deal with a representative sample survey of eligible non-participant, as well as participants. Additionally, the larger the sample of eligible non-participants, the better the matching is facilitated. In our sample we have more than 1700 eligible non-participants and more than 200 participants. Participants and non-participants should face the same economic incentives that might affect choices such as programme participation (see, for example, Martin Ravallion 2012). The dataset used in this research contains sufficient information to ensure that the conditional independence assumption holds.

Another condition for the validity of propensity score matching is the assumption of common support or overlap condition. This assumption ensures that observations with the same X values have a positive probability of being both participants and non-participants. The common support assumption implies that for each treated

individual, there is another non-treated individual who can be used as a matched comparison observation (Cristina Borra et al. 2012). While conditional independence assumption is not a directly testable assumption, the validity of the common support assumption can be tested. Figure 1 shows the propensity score histogram by treatment status. As demonstrated, for the large majority of treated individuals there is a similar control group of individuals, so that the common support assumption is satisfied. Only 0.5% of individuals have been discarded in the estimations.



Source: Author's calculation based on CBS (2016) data.

Figure 1 Distribution of Estimated Propensity Score by Treatment Status

In such types of evaluation balancing tests, researchers can also check the success rate of the matching for exogenous variables. As researchers, we aim to analyse the quality of matching between treated and non-treated individuals. We follow the work of Jose Zubizarreta, Ricardo Paredes, and Paul R. Rosenbaum (2014) who suggest checking whether significant differences between the average values of the variables for both groups exist after matching. Before matching we expect differences, but after matching the variables should be balanced in both groups and significant differences should not persist (Borra et al. 2012). Table 2 displays the mean values of variables for both treated and controls, before and after matching.

Table 2 Matching Quality

| Variable | Unmatched | | | | Matched | | | |
|--|-----------|---------|--------|---------|---------|---------|--------|--------|
| | Treated | Control | % bias | t-test | Treated | Control | % bias | t-test |
| Marrital status | 0.11 | 0.06 | 16.7 | 2.64* | 0.11 | 0.09 | 6.6 | 0.65 |
| Number of parents abroad | 1.04 | 0.76 | 30.1 | 4.31* | 1.04 | 1.02 | 1.6 | 0.17 |
| Generation | 0.82 | 0.59 | 28.5 | 4.10* | 0.82 | 0.82 | 0.1 | 0.01 |
| Age | 20.79 | 20.92 | -8.2 | -1.17 | 20.79 | 20.88 | -5.2 | -0.55 |
| Sex | 1.62 | 1.36 | 53.0 | 7.58* | 1.62 | 1.62 | -1.3 | -0.13 |
| Employment status before the programme | 0.14 | 0.71 | -142.4 | -18.53* | 0.14 | 0.14 | 0.0 | -0.00 |
| Unemployment duration before the programme | 175.18 | 170.37 | 1.4 | 0.20 | 175.18 | 151.26 | 7.0 | 0.77 |

Notes: * denotes significance at 5% level.

Source: Author's calculation based on CBS (2016) data.

According to the above matching, differences between the treated and control group are reduced considerably. An exception can be seen in the variable unemployment duration before the programme. For this variable the difference between the treated and control group is not eliminated and was minimal before the matching. With the *t*-test, conducted before and after matching, we check the null hypothesis that the mean values of the treated group and the control group do not differ after matching. This hypothesis cannot be rejected for any variable. Therefore, we conclude that with the matching it was possible to construct a control group which is statistically similar to the treated group.

3. Results

The outcome measures used for the calculation of causal effects are employment probability and probability of studies. We estimated average treatment effect on the treated (ATT) using a nearest neighbour matching estimator. Table 4 displays the estimation results of the causal effects. One year after the programme started, the effect of the programme is negative for both outcome measures considered and is statistically significant. The individuals participating in the employment programme present negative employment probability and negative probability of studies. Two years after the programme started, employment probability is still negative and is/remains statistically significant. Additionally, we also wanted to check in the literature known as lock-in effect (see, for example, Štefánek 2014). Lock-in effect refers to potential decreases in job search intensity by programme participants. To check for the potential lock-in effect in our data we use a/the log-rank test (chi square log-rank test). It is used to test the null hypothesis that there is no difference between the population survival curves (the probability of an event occurring at any time point is the same for each population). In this case, we want to test whether there are statistically significant differences in survival curves between unemployed youth who did not participate in an employment programme and those unemployed youth who did participate in an employment programme.

Table 3 Log-Rank Test

| | Chi-square | d.f. | Sig. |
|-----------------------|------------|------|-------|
| Log rank (mantel-cox) | 582.914 | 1 | 0.000 |

Source: Author's calculation based on CBS (2016) data.

Table 3 shows a log-rank test with $p = 0.000$. Taking into account $\alpha = 0.05$ we reject the null hypothesis and conclude that there are statistically significant differences between the two groups on having a shorter time to event. We can determine that unemployed youths who did participate in an employment programme have longer time to event than those unemployed youths who did not. Therefore, we can conclude, that participation in an employment programme decreases job search efforts. Participants in an employment programme are less motivated to search for a job than they would be if they did not participate in employment programme.

In contrast, two years after the programme started, the probability of studies is still negative and statistically significant. Education plays a crucial role in preparing

young people for the labour market. Young people who are low-skilled are more inclined to become unemployed, in particular on a long-term basis as this is the characteristic of the Dutch labour market. Moreover, there is a high risk that unemployed youth will become socially excluded. Furthermore, as the employment programme is a special programme targeted at disadvantaged young people, the negative effects may be due to the existence of stigmatisation. Migrant youths represented in the sample also face limited networks and discrimination in comparison to native Dutch youths. These migrant youths are unemployed more often, which may be attributed to the intensity of the job search behaviour and differences in educational attainment. Since education is a crucial factor in this transition from school to work, the education system should focus more on empowering young people with the skills and work experiences needed on the labour market, thus more likely to avoid the skill mismatches.

Table 4 Average Treatment Effect on the Treated

| Outcome measure | ATT | St. error |
|---------------------------------------|---------|-----------|
| One year after the programme started | | |
| Employment probability | -0.171* | 0.037 |
| Probability of studies | -0.076* | 0.028 |
| Two years after the programme started | | |
| Employment probability | -0.241* | 0.039 |
| Probability of studies | -0.067* | 0.028 |

Notes: * denotes significance at 5% level.

Source: Author's calculation based on CBS (2016) data.

Matching is based on the conditional independence assumption which states that the researcher should observe all variables simultaneously influencing the participation decision and outcome variables (Sascha O. Becker and Marco Caliendo 2007). If participants and non-participants differ in terms of both observed and unobserved characteristics, then the conditional independence assumption is violated and the results biased (see, for example, Borra et al. 2012). Consequentially, a strong identifying assumption and testing of this assumption with non-experimental data is not possible. Therefore, we check the sensitivity of the estimated results with respect to deviations from conditional independence assumption with the bounding approach initially suggested by Rosenbaum (2015). This approach consists of simulating an unobserved component and testing to which degree unobserved heterogeneity results are robust. If unobserved factors are present, then identical individuals with respect to observable characteristics have different probabilities of receiving treatment. To simulate an unobserved term, an artificial factor T is introduced and the influence of T is gradually increased to assess the impact of this artificial factor on the results by comparing treated and non-treated individuals (Becker and Caliendo 2007; Borra et al. 2012).

Table 5 displays the sensitivity test statistics for the average treatment effects on the treated. Since a sensitivity analysis for significant treatment effects is meaningful, a sensitivity analysis for insignificant treatment effects will be omitted. For the negative estimated treatment effects, we display the test statistics Q_- for the lower

bound, under the assumption that we underestimated the treatment effects and those who participate in the employment programme always have lower employment probability or probability of studies even in the absence of treatment. The test statistics are calculated for nearest the neighbour matching algorithm using the procedure `mhbounds.ado` provided by Becker and Caliendo (2007). For both outcome measures (employment probability and probability of studies) the starting point is a situation of unobserved heterogeneity with $\Gamma = 1.00$. The value of Γ is then gradually increased in order to assess the potential strength of unmeasured influences. The negative estimated treatment effects for the outcome measure employment probability measured one/two years after the programme started are quite robust to unobserved factors. Critical values of Γ are between 2.50 and 3.00, indicating that individuals with the same X -vector would have to differ in their odds of participation by a factor of 2.50 (3.00) to turn insignificant at the 5% significance level. Therefore, we feel confident on the strength of outcome measure employment probability one/two years after the programme started. Alternatively, for the outcome measure probability of studies one/two years after the programme started, negative estimated treatment effect results are sensitive to potential unobserved heterogeneity. With just a 5% or 15% difference in the odds of participating of individuals with the same observed characteristics, treatment effects become insignificant. But if we look for an example at the situation for $\Gamma = 2.75$ we get a significance level p of 0.006 indicating a significant effect once again. This second significant value of p indicates a significant positive treatment effect. This is due to how we assume a large negative unobserved heterogeneity which transforms our previously significant negative effect into positive.

Table 5 Sensitivity to Unobserved Heterogeneity

| One year after the programme started | | |
|---------------------------------------|------|-------|
| Gamma | Q- | p- |
| Employment probability | | |
| 1.00 | 5.66 | 0.000 |
| 1.50 | 3.57 | 0.000 |
| 2.00 | 2.14 | 0.016 |
| 2.50 | 1.04 | 0.147 |
| Probability of studies | | |
| 1.00 | 2.07 | 0.019 |
| 1.50 | 0.53 | 0.294 |
| 2.00 | 0.26 | 0.396 |
| 2.50 | 1.08 | 0.038 |
| Two years after the programme started | | |
| Gamma | Q- | p- |
| Employment probability | | |
| 1.00 | 6.03 | 0.000 |
| 1.50 | 3.78 | 0.000 |
| 2.00 | 2.23 | 0.012 |
| 2.50 | 1.04 | 0.148 |
| Probability of studies | | |
| 1.00 | 1.56 | 0.058 |
| 1.50 | 0.02 | 0.510 |
| 2.00 | 0.88 | 0.188 |
| 2.50 | 1.75 | 0.039 |

Source: Author's calculation based on CBS (2016) data.

4. Conclusion

We estimated causal effects of the Dutch employment programme (participation programme) using a propensity score matching method. The propensity score matching method is one of the modern methods of empirical analysis for evaluating treatment effects, proposed by the literature in the field of evaluating programmes. The identification of an average treatment effect is based on the conditional independence assumption. This is an assumption in accordance with which participation in the programme is independent of after-programme outcomes and conditioned by observed external factors. Considering these assumptions, the evaluations of an average treatment effect for the treated are not bias. We estimate causal effects on the young unemployed Dutch people, whereas the outcome measures are probability of re-employment and probability of participation in the regular educational system.

The results indicate that the employment programme has statistically significant negative effects in the short-term as well as in the long-term. The lock-in effect is also present amongst programme participants. Participants in an employment programme are less motivated to search for a job than they would be if they would not participate in an employment programme. The quality of the education system ensures that young people are empowered with the skills and work experiences needed on the labour market and therefore prepared for their working life. Therefore, as Dutch employers often stress that school leavers are not well prepared to enter the labour market and their skills are not suitable, there should be direct involvement by industries in the formulation of the professional competences. This latter concept would narrow the gap between worker skills and their job tasks.

In correlation, during the education process the teaching methods and strategies should be changed and combine learning and work practice as much as possible. Disadvantaged youths often yield lower self-confidence and bad work ethics, therefore there should be more effort placed into a person's centred and integrated approach, counselling and guidance, learning (and language) support to improve job sustainability and to avoid the risk of becoming socially excluded, unemployed long-term and to prevent young unemployed people from disappearing from official evidences. What is more, general economic conditions should be taken into consideration.

References

- Abadie, Alberto, and Guido W. Imbens.** 2016. "Matching on the Estimated Propensity Score." *Econometrica*, 84(2): 781-807. <http://dx.doi.org/10.3982/ECTA11293>
- Austin, Peter C.** 2011. "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies." *Multivariate Behavioral Research*, 46(3): 399-424. <http://dx.doi.org/10.1080/00273171.2011.568786>
- Becker, Sascha O., and Marco Caliendo.** 2007. "Mhbounds - Sensitivity Analysis for Average Treatment Effects." German Institute for Economic Research Discussion Paper 659.
- Bettendorf, Leon J. H., Egbert L. W. Jongen, and Paul Muller.** 2012. "Childcare Subsidies and Labour Supply: Evidence from a Large Dutch Reform." Netherlands Bureau for Economic Policy Analysis Discussion Paper 217.
- Bettendorf, Leon J. H., Kees Folmer, and Egbert L. W. Jongen.** 2014. "The Dog That Did Not Bark: The EITC for Single Mothers in the Netherlands." *Journal of Public Economics*, 119(9): 49-60. <http://dx.doi.org/10.1016/j.jpubeco.2014.07.010>
- Borghans, Lex, Bas Ter Weel, and Bruce A. Weinberg.** 2014. "People Skills and the Labor-Market Outcomes of Underrepresented Groups." *Industrial and Labor Relations Review*, 67(2): 287-334. <http://dx.doi.org/10.1177/001979391406700202>
- Borra, Cristina, Luis Palma, Carmen M. González, and Luis F. Aguado.** 2012. "Evaluation of an Active Labour Market Programme in a Context of High Unemployment." *Revista Desarrollo y Sociedad*, 70: 93-115.
- Caliendo, Marco, and Steffen Künn.** 2012. "Getting Back into the Labour Market: The Effects of Start-Up Subsidies for Unemployed Females." Institute for the Study of Labor Discussion Paper 6830.
- Caliendo, Marco, and Ricarda Schmidl.** 2015. "Youth Unemployment and Active Labor Market Policies in Europe." Institute for the Study of Labor Discussion Paper 9488.
- Deelen, Anja, Marloes de Graaf-Zijl, and Wiljan van den Berge.** 2014. "Labour Market Effects of Job Displacement for Prime-Age and Older Workers." Netherlands Bureau for Economic Policy Analysis Discussion Paper 285.
- De Graaf-Zijl, Marloes, Hugo Erken, Debby Lanser, Wiljan van den Berge, and Eric van Loon.** 2014. "Labour Market Responses to the Great Recession." In *Roads to Recovery*, ed. George Gelauff, Debby Lanser, Albert van der Horst, and Adam Elbourne, 78-104. The Hague: CPB Netherlands Bureau for Economic Policy Analysis.
- Dehejia, Rajeev H.** 2013. "The Porous Dialectic: Experimental and Non-Experimental Methods in Development Economics." World Institute for Development Economic Research Working Paper 011.
- De Koning, Jaap, Arie Gelderblom, and Peter van Nes.** 2005. "Youth Unemployment in the Netherlands." Paper presented at the Peer Review: Youth Unemployment Task Force, The Hague.
- Dinç, Tuna.** 2012. "The Two Sector Model of Learning-By Doing and Productivity Differences." *Panoeconomicus*, 59(5): 583-598. <http://dx.doi.org/10.2298/PAN1205583D>
- European Commission.** 2013. "PES Approaches for Sustainable Activation of Low Skilled Adults and Youths: Work-First or Train-First?" European Commission Employment, Social Affairs and Inclusion Analytical Paper.

- Euwals, Rob, Annemiek van Vuren, and Daniel van Vuuren.** 2011. "The Decline of Early Retirement Pathways in the Netherlands." Netherlands Bureau for Economic Policy Analysis Discussion Paper 179.
- Frölich, Markus, and Michael Lechner.** 2015. "Combining Matching and Nonparametric Instrumental Variable Estimation: Theory and an Application to the Evaluation of Active Labour Market Policies." *Journal of Applied Econometrics*, 30(5): 718-738. <http://dx.doi.org/10.1002/jae.2417>
- Gautier, Pieter A., and Bas van der Klaauw.** 2009. "Institutions and Labor Market Outcomes in the Netherlands." Insitute for Labour Market Policy Evaluation Working Paper 28.
- Gričar, Sergej, and Štefan Bojnec.** 2013. "Inflation and Hospitality Industry Prices." *Eastern European Economics*, 51(3): 91-108.
- Hansen, Ben, Paul R. Rosenbaum, and Dylan Small.** 2014. "Clustered Treatment Assignments and Sensitivity to Unmeasured Biases in Observational Studies." *Journal of the American Statistical Association*, 109(505): 133-144. <http://dx.doi.org/10.1080/01621459.2013.863157>
- Heckman, James J., and Rodrigo Pinto.** 2015. "Causal Analysis after Haavelmo." *Econometric Theory*, 31(1): 115-151. <http://dx.doi.org/10.1017/S026646661400022X>
- Heckman, James J., John Eric Humphries, and Gregory Veramendi.** 2016. "Dynamic Treatment Effects." *Journal of Econometrics*, 191(2): 276-292. <http://dx.doi.org/10.1016/j.jeconom.2015.12.001>
- Huber, Martin, Michael Lechner, and Conny Wunsch.** 2013. "The Performance of Estimators Based on the Propensity Score." *Journal of Econometrics*, 175(1): 1-21. <http://dx.doi.org/10.1016/j.jeconom.2012.11.006>
- Huber, Martin, Michael Lechner, and Andreas Steinmayr.** 2015. "Radius Matching on the Propensity Score with Bias Adjustment: Tuning Parameters and Finite Sample Behaviour." *Empirical Economics*, 49(1):1-31. <http://dx.doi.org/10.1007/s00181-014-0847-1>
- Hujer, Reinhard.** 2011. "Mikroökonomie und Kausalität: Zur Bedeutung Arbeitsmarktpolitischer Evaluationsstudien." *AStA Wirtschafts- und Sozialstatistisches Archiv*, 5(1): 5-18.
- Kok, Suzanne, Nicole Bosch, Anja Deelen, and Rob Euwals.** 2011. "Migrant Women on the Labour Market: On the Role of Home- and Host-Country Participation." Institute for the Study of Labor Discussion Paper 5817.
- Kok, Suzanne.** 2013. "Matching Worker Skills to Job Tasks in the Netherlands: Sorting into Cities for Better Careers." Netherlands Bureau for Economic Policy Analysis Discussion Paper 247.
- Lechner, Michael, and Conny Wunsch.** 2013. "Sensitivity of Matching-Based Program Evaluations to the Availability of Control Variables." *Labour Economics*, 21(C): 111-121. <http://dx.doi.org/10.1016/j.labeco.2013.01.004>
- Organisation for Economic Co-operation and Development.** 2008. *Jobs for Youth: Netherlands*. Paris: OECD. <http://dx.doi.org/10.1787/9789264041295-en>
- Ravallion, Martin.** 2012. "Can We Trust Shoestring Evaluations?" World Bank Policy Research Working Paper 5983.
- Roelofs, Gijs, and Daniel van Vuuren.** 2011. "The Decentralization of Social Assistance and the Rise of Disability Insurance Enrolment." Netherlands Bureau for Economic Policy Analysis Discussion Paper 185.

- Rosenbaum, Paul R.** 2015. "How to See More in Observational Studies: Some New Quasi-Experimental Devices." *Annual Review of Statistics and Its Application*, 2(1): 21-48. <http://dx.doi.org/10.1146/annurev-statistics-010814-020201>
- Rosholm, Michael, and Michael Svarer.** 2004. "Estimating the Threat Effect of Active Labour Market Programmes." University of Aarhus Economics Working Paper 2004-06.
- Rothe, Christoph.** 2016. "The Value of Knowing the Propensity Score for Estimating Average Treatment Effects." Institute for the Study of Labor Discussion Paper 9989.
- Schünemann, Benjamin, Michael Lechner, and Conny Wunsch.** 2015. "Do Long-Term Unemployed Workers Benefit from Targeted Wage Subsidies?" *German Economic Review*, 16(1): 43-64. <http://dx.doi.org/10.1111/geer.12040>
- Silber, Jeffrey H., Paul R. Rosenbaum, Richard N. Ross, Justin M. Ludwig, Wei Wang, Bijan A. Niknam, Nabanita Mukherjee, Philip A. Saynisch, Orit Even-Shoshan, Rachel R. Kelz, and Lee A. Fleisher.** 2014. "Template Matching for Auditing Hospital Cost and Quality." *Health Services Research*, 49(5): 1446-1474. <http://dx.doi.org/10.1111/1475-6773.12156>
- Štefánik, Miroslav.** 2014. "Estimating Treatment Effects of a Training Programme in Slovakia Using Propensity Score Matching." *Ekonomický časopis*, 62(6): 631-645.
- Van den Berg, Gerard J., and Bas van der Klaauw.** 2015. "Structural Empirical Evaluation of Job Search Monitoring." Institute for Evaluation of Labour Market and Education Policy Working Paper 16.
- Zubizarreta, Jose, Ricardo Paredes, and Paul R. Rosenbaum.** 2014. "Matching for Balance, Pairing for Heterogeneity in an Observational Study of the Effectiveness of for-Profit and Not-for-Profit High Schools in Chile." *The Annals of Applied Statistics*, 8(1): 204-231. <http://dx.doi.org/10.1214/13-AOAS713>

THIS PAGE INTENTIONALLY LEFT BLANK