I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11870

Vocational Training for Unemployed Youth in Latvia: Evidence from a Regression Discontinuity Design

Massimiliano Bratti
Corinna Ghirelli
Enkelejda Havari
Giulia Santangelo

OCTOBER 2018

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11870

Vocational Training for Unemployed Youth in Latvia: Evidence from a Regression Discontinuity Design

Massimiliano Bratti
European Commission JRC, University of Milan, IZA and LdA

Corinna Ghirelli
Bank of Spain

Enkelejda Havari
European Commission JRC

Giulia Santangelo
European Commission JRC

OCTOBER 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Vocational Training for Unemployed Youth in Latvia: Evidence from a Regression Discontinuity Design*

This paper evaluates the impact of a vocational training programme on the labour market outcomes of unemployed youth in Latvia. The programme is part of the Youth Guarantee scheme for the period 2014-2020, the largest action launched by the European Union to reduce youth unemployment rate and to support young people aged between 15 and 29 who are not in education, employment or training (NEETs). To estimate the causal effect of participating in the programme on the employment outcomes, we exploit a rule that gives priority for participation to unemployed people under the age of 25 using a fuzzy Regression Discontinuity Design setup. The estimated effects of the programme on the probability of being employed and monthly income up to 3.5 years after registering in the programme are positive but not statistically significant, although we find a strong positive effect of the priority rule on programme participation. This is the first evidence on the impact of a programme within the current Youth Guarantee scheme in Europe and our findings are in line with those from the literature on the evaluation of active labour market policies targeting youth.

JEL Classification: J01, J08, J18, J24

Keywords: youth guarantee, Latvia, vocational training, employment, policy evaluation, administrative data, regression discontinuity design

Corresponding author:
Enkelejda Havari
European Commission Joint Research Centre
Via E. Fermi 2749
Ispra (VA), 21027
Italy
E-mail: enkelejda.havari@ec.europa.eu

* We thank the Evaluation Division of the EU Funds Strategy Department of Latvia, the State Employment Agency and the State Revenue Service. We are indebted to Normunds Strautmanis and Janis Leikucs from the Latvian Ministry of Finance for granting access to and collecting the data from the responsible authorities. We also thank Andrea Albanese, Marco Bertoni, Audinga Baltrunaite, Lorenzo Cappellari, Antonio di Paolo, Florian Mayneris, Fabrizia Mealli, Giovanni Mellace and Matteo Picchio for their helpful discussions, as well as conference participants at the 30th Annual Conference of the European Association of Labour Economists (Lyon, 2018), the Counterfactual Methods for Policy Impact Evaluation conference (Berlin, 2018), the IZA World of Labour conference (Berlin, 2018), the European Society for Population Economics conference (Antwerp, 2018), the RWI Youth employment conference (Berlin, 2017), XII Jornadas de Economia Laboral (Vallodolid, 2017), the 32nd annual Conference of the Italian Association of Labour Economists (Cosenza, 2017), and seminar participants at the IAB-Institute for Employment Research (Nuremberg, 2017). Any views expressed in this paper are those of the authors and do not reflect those of the European Commission or the Bank of Spain.

1 Introduction

The global financial crisis was followed by an increase in unemployment rates across Europe, particularly among young people. In 2013, the youth unemployment rate peaked at 23.9%, exceeding 50% in countries such as Greece (58.3%) and Spain (55.5%).[1] In response, the European Union (EU) launched two actions in 2013, aiming to tackle youth unemployment: the Youth Employment Initiative (YEI) in February and Youth Guarantee (YG) in April. Under the YG initiative, all EU Member States are committed to implementing policies aimed at reducing youth unemployment. This commitment is reflected in the large number of YG implementation plans that identify specific measures in each country.[2] Typically, these plans involve a number of active labour market policies (ALMPs) — e.g., apprenticeship, traineeship, job placement or further education leading to a qualification — targeted at young people who are not in employment, education or training (NEETs). The YG plans are financed by a mix of EU funds and national resources. The YEI is one of the main financial instrument supporting the implementation of YG schemes with a total budget of 8.8 billion euros for the period 2014-2020. It is mainly targeted at regions with particularly high levels of youth unemployment.[3]

Although YG plans and national experiences are well documented (Cabasés Piqué et al., 2016; Pastore, 2015; Escudero and Mourelo, 2015, among others), evidence of the effectiveness of ALMPs financed through this package is almost non-existent. The main reason is that most of these measures were introduced very recently and are still running, therefore the data collection process for performing rigorous impact evaluations is ongoing. Since YG is the EU's response to high unemployment rates among young people, it is important to investigate whether or not this action has been effective in achieving this goal. The aim of this study is to fill this gap by providing evidence from the recent implementation of the YG scheme in Latvia.

Latvia provides an interesting case study as it was one of the European countries most affected by the financial crisis. As stated by Paul Krugman, "The most acute problems are on Europe's periphery, where many smaller economies are experiencing crises strongly reminiscent of past crises in Latin America and Asia: Latvia is the new Argentina." (The New York Times, 2008). Nevertheless, the country later showed the fastest recovery in the post-recession period and Latvia's gross domestic product (GDP) has shown stable

[1] Eurostat (2017), Employment and unemployment: Labour Force Survey, 2017.

[2] National YG schemes are presented to the European Council for approval and the European Commission monitors their implementation. Updates on YG in each country are regularly published in the official pages of the European Commission's Directorate General for Employment and Social Affairs.

[3] Countries that had at least one region with a youth unemployment rate higher than 25% in 2012 were eligible to use funds from YEI. It was topped up in 2017 for regions with youth unemployment higher than 25% in 2016. The initial budget of 2014-2015 was 6.4 billion euros; however, in September 2016, given the still high levels of youth unemployment, the European Commission proposed to increase this budget. In June 2017, the Council and the European Parliament agreed to an increase of the YEI by 2.4 billion euros for eligible Member States for the period 2017-2020. See http://ec.europa.eu/social/main.jsp?catId=1176.

growth since 2011. The unemployment rate has not yet fallen below pre-crisis levels however.[4] Indeed, Latvia is one of the countries that met the conditions for accessing YEI funds to finance the national YG plan.[5]

In this analysis, we focus on the evaluation of a large vocational training (VT) programme targeted at NEETs aged between 15 and 29.[6] We use a fuzzy regression discontinuity design (FRDD), taking into account a specific eligibility criterion used by the Latvian government that gives a higher priority for participation to unemployed people under the age of 25. The priority rule is used as an instrument for participation, to estimate the causal effect of participation in the VT programme on labour market outcomes later in life (up to two years after participation), allowing us to address the issue of potential self-selection into the VT programme.

We use rich administrative data from the Latvian State Employment Agency (SEA), which provides information on the population registered as unemployed on specific dates (including both participants and non-participants in the VT programme), and match them with data from the State Revenue Service (SRS), which gathers information on an individual's income on specific dates before and after the programme.

Our results show that the priority rule strongly predicts participation in the programme. Our first stage results are strong and statistically significant. On the one hand, we find that being younger than 25 increases the likelihood of participation by 7 percentage points (pp). On the other hand, we find a positive but statistically non-significant effect of the training programme on the probability of being employed and on the level of one's monthly income up to two years after the training. The FRDD makes it possible to achieve local identification, i.e., to estimate the effect of the VT programme on individuals who received training as a result of the priority rule (i.e., *compliers*). Our analysis of this sub-population shows that, on average, compliers tend to have a lower level of education and be less employable than the average individual in the sample.

We contribute to two strands of the literature: the general and rich literature on the evaluation of ALMPs and the more specific and thin literature that looks at NEETs and disadvantaged target groups. The former is already quite extensive, yet results are mixed. Our paper contributes to the extant literature in two ways. First, to our knowledge, our paper is one of the first evaluations of the YG scheme in Europe for the pro-

[4] Until 2007, the Latvian economy was growing rapidly and the national unemployment rate was 6.1% (youth unemployment was 13%). Latvia was severely hit by the financial crisis, with total unemployment reaching 17.5% by 2009 (33% for youth unemployment) — the highest in the EU after Spain — and the country's real GDP falling by 13% with respect to the previous year. Latvia's recovery was also among the fastest in Europe: in 2011, its real GDP grew by 6.4% (one of the highest growth rates in Europe, second only to Estonia). Despite this, by 2016 the GDP remained below pre-crisis levels and the national youth unemployment rate was still 17% (Source: Eurostat).

[5] The Latvian YG plan is financed by the YEI, the European Social Fund (ESF) and the Latvian State Budget. It is managed by the State Employment Agency (SEA) of Latvia (the Latvian public employment service).

[6] In Latvian, the programme is called "JG Profesionalas apmacibu programmas" and is one of the professional training programmes provided by the Latvian public employment service.

gramme period 2014–2020. Second, we provide evidence on the effects of the VT training in the short and medium-term, as individuals are followed up to two years after their participation in the programme. This is an important added value, as most of the literature that analyses large ALMPs on youth focus on short-term effects.

The paper is organised as follows. Section 2 summarises the relevant literature. Section 3 describes the programme and the phases of its implementation. Section 4 presents the data and some descriptive statistics. Section 5 explains the empirical strategy. Section 6 discusses the main results and Section 8 presents the conclusion.

2 Literature review

The existing literature focuses on programmes very similar to YG that were implemented in the Nordic countries and the UK in the 1980s and 1990s, such as the British New Deal for Young People (NDYP), the Danish Youth Unemployment Program (YUP), and the German *Jugend mit Perspektive* (JUMP).[7] A common feature of these programmes is a strict age limit on participation. In the case of Sweden, Carling and Larsson (2005) showed that the reform passed in 1998 had a positive effect on youth employability in the short term but no impact in the long term.[8] A recent paper by Hämäläinen et al. (2014) examined the YG programme introduced in Finland in 2005. Taking into account the age-eligibility threshold, the authors found that the programme moderately increased unsubsidised employment among young people aged 23–24. The reduction in the unemployment rate was otherwise negligible. Furthermore, estimates based on level of education showed that the programme did not improve the labour market prospects of unskilled young people. Overall, evaluation studies of similar programmes have found small, positive effects on employment prospects in the short term, however these tend to become negligible in the long term. An exception is the paper by Blundell et al. (2004), which demonstrates positive effects of participating in the UK's NDYP programme on youth employment in both the short and long term. However, unlike other programmes, the NDYP mostly involved activation measures such as job-search assistance and counselling services. The programme introduced extensive job assistance and wage subsidies to employers; it was piloted in certain areas and then extended to others. Exploiting variation in age-eligibility criteria and geographical area, the authors found that the programme significantly increased employment among young people aged 18–24 years.

A key finding of the ALMP literature is that effects seem to differ depending on when the impact is

[7] Sweden was the first country to introduce its own programme for young people in 1984, followed by Norway in 1993 and Denmark and Finland in 1996 (Mascherini, 2012; Escudero and Mourelo, 2015). Similarly, the UK implemented the NDYP in 1998 to target unemployed young people aged 18–24.

[8] This reform was similar in spirit to YG: it targeted unemployed individuals aged 20-24 and aimed to prevent long-term unemployment by guaranteeing an assignment to one of a number of ALMPs within the first 100 days of unemployment.

measured (Card et al., 2015), regardless of the age group being targeted. In most cases, the estimated effects of training programmes in the short term, i.e. a maximum of one year after it has been completed, are often negative or not significant. A possible reading of this evidence lies in the so-called "lock-in effect", i.e. the fact that participants, compared with non-participants, do not have enough time or do not effectively spend time searching for a job while attending the training programme. The lock-in effect theory does not rule out the possibility that such training can increase participants' employment prospects however. Indeed, the programme could prove effective if evaluated in the medium or long term, when lock-in effects have faded away or have been outweighed by the beneficial effects of the programme.[9]

Evidence from the US,[10] Germany,[11] and more recently, Latin America (Ibarrarán et al., 2015), confirms that results are sensitive to the timeframe used for measuring outcomes. In the best-case scenario, outcome data should be collected over a long enough period to enable the dynamics of the estimated effects to be documented and to uncover the potentially positive effects of the programme.

To the same extent, the literature on training programmes targeted specifically at young people does not show clear-cut results (Caliendo and Schmidl, 2016).[12] Card et al. (2010) carried out a meta-analysis of the effectiveness of ALMPs in continental Europe, northern Europe, and Anglophone countries, concluding that programmes targeted at young people are less effective at improving employment prospects than programmes targeted at adults. Kluve (2010) performed a similar analysis for several European countries, confirming the tendency of training programmes (at least in developed countries) to show low-to-modest effects on employment rates.[13]

Among developing countries, the results are more encouraging. For instance, an evaluation of a large randomised training programme in Colombia in the early 2000s found positive effects on employment and earnings (Attanasio et al., 2017, 2011).[14] Similar findings emerged for another large-scale experimental training programme in the Dominican Republic.[15] The evidence from this country suggests that the results were heterogeneous, with the programmes proving particularly effective for men and for young people living

[9] The literature defines medium-term effects as those measured within two years of the end of a programme, whereas long-term effects are measured over longer periods.

[10] This literature evaluates the effects of US job training programmes based on large-scale experiments: the National Supported Work experiment (Couch, 1992), the JOBSTART demonstration (Cave et al., 1993) and Job Corps (Schochet et al., 2008; Flores-Lagunes et al., 2010).

[11] This literature relies on quasi-experimental designs and administrative data: see Fitzenberger and Völter (2007) for East Germany and Lechner et al. (2011) and Biewen et al. (2014) for West Germany.

[12] For reviews of evaluations of ALMPs targeted at young people in developed countries, see Grubb (1999) for the US and Martin and Grubb (2001) and Heckman et al. (1999) for the US and Europe.

[13] Kluve (2010) considered Austria, Denmark, Germany, Finland, France, Ireland, the Netherlands, Norway, Switzerland, and the UK. Card et al. (2010) focused on Austria, Germany and Switzerland (continental Europe); Denmark, Finland, Norway and Sweden (northern Europe); and Australia, Canada, New Zealand, the UK and the US (Anglophone countries).

[14] In both studies, the programme targeted young people who were assigned to the training course at random.

[15] The effectiveness of such a programme on employment and earnings has been evaluated over time: Card et al. (2011) for short-term effects; Ibarrarán et al. (2014) and Ibarrarán et al. (2015) for long-term effects.

in the capital (Ibarrarán et al., 2015). These effects tend to grow over time, which confirms the conjecture that measuring outcomes over a long enough timespan is crucial.

A major difference between the recent programmes targeted at young people in developing countries and most of the programmes implemented in developed countries in the 2000s (including YG) is that the former are designed to be "demand-driven". This offers a possible explanation for the different results obtained. Typically, training programmes are held in two phases in developing countries: in the first phase, participants attend a training course in the classroom and prepare for a specific occupation, while the second phase consists of internships in the private sector. The rationale behind "demand-driven training" is twofold. First, having the proper skills may not be enough for an unemployed person to find a job, as their offered skills must match demand in the labour market. Second, it is important that the training provides skills that are in demand. The combination of these phases is designed to make it easier for participants to enter employment. One of the few papers that tackles this issue is the one by Alfonsi et al. (2017), who design a labour market experiment (a randomised control trial) that considers both the demand and supply side. The experiment tracks 1700 workers and 1500 firms in Uganda over four years and looks at the effect on employment of offering VT to workers compared to offering wage subsidies to train the workers on-the-job (firm-trained workers). They find that both treatments lead to skill accumulation but whilst VT workers learn sector-specific skills, firm-trained workers learn more firm-specific skills. The employment rate is larger for VT workers which is explained by the certifiability and transferability of their skills (the authors estimate a structural model to uncover the mechanisms). This evidence is also confirmed by looking at the firm-side, where the higher employment rate for VT workers is explained by a better match of these workers to more productive firms.

3 Youth Guarantee and the VT programme

YG is a recent EU initiative aimed at helping unemployed young people enter employment or re-enter the education system. Under the YG umbrella, EU Member States commit to implementing measures to ensure that unemployed young people receive good offers of employment, continued education, apprenticeships or traineeships within four months of leaving school or becoming unemployed.

In Latvia, YG consists of a series of ALMPs targeted at young people who meet the following eligibility criteria: (i) being aged between 15 and 29 and (ii) being registered as unemployed at the Latvian SEA. After registering at the SEA, unemployed young people undergo a profiling phase during which caseworkers assess their key competencies. Furthermore, caseworkers assist them in their job search by providing tailored career guidance. Within four months of registration, unemployed young people should be offered one of the following options, depending on their needs: a satisfactory employment opportunity, the opportunity to continue their education or the opportunity to participate in an ALMP (e.g., an apprenticeship or a vocational traineeship).

In this study, we analyse the effect of a VT programme[16] aimed at improving or providing a vocational qualification in line with labour demand. Based on their assessments however, caseworkers can offer other types of programmes to YG recipients: (i) first work experience measures and workshops offered to young people with sufficient skills but without work experience; (ii) subsidised work, which is offered to disadvantaged young people such as those with disabilities or the long-term unemployed; and (iii) support for entrepreneurship, which is offered to young people with the appropriate education or entrepreneurial skills. In our analysis, we selected as a control group potential YG recipients who are eligible for YG but did not take part in any ALMP.[17]

The VT programme we focused on started in January 2014 and was expected to continue until 2018.[18] As with all YG programmes, this VT programme targets unemployed young people who registered at the SEA and who were aged 15–29 at the time of registration. In addition, it targets potential YG recipients who (i) have a vocational qualification that is not in demand in the labour market, (ii) have lost their vocational skills or (iii) have not yet achieved a vocational qualification. Finally, unlike other YG programmes, the VT programme in question has an additional eligibility criterion that gives a higher priority for participation to young people under 25 years of age. In our identification strategy, we used this priority rule to address concerns of potential self-selection into the VT programme and to identify the programme's causal effect on later employment and income.

Training courses are managed through a voucher system[19]: unemployed young people receive a voucher that can be spent in one of the country's vocational education institutions.[20] Training courses run for a period of three to nine months, and participants' start and end dates vary. After passing a final examination, participants receive a certification of their professional qualification.

After a participant registers at the SEA, her eligibility for participating in the VT programme is checked. We use the registration date to monitor participation in the VT programme over a specific time (we observe participation and define treatment if one participated in the VT programme within 12 months of registration). We match these monitoring data with data from individuals' tax records, which provide information

[16] For a detailed description of the VT programme, see Appendix A. In this section we sketch the features of the programme and the phases of the implementation that are relevant for the empirical strategy used in the study.

[17] Data have been extracted by the Latvian SEA such that people in the control group had not participated in any type of ALMP.

[18] This paper considered young people who registered between January and December 2014 but participated in the training between 2014 and 2015, in order to have a long enough timespan after the training to observe employment outcomes and minimise the impact of "lock-in effects".

[19] The provision of ALMPs through vouchers has recently been discussed in the literature referring mainly to the German and US active labor market policies. One issue with this type of provision is related to the fact that individuals may decide not to redeem their vouchers, inducing efficiency losses. Strittmatter (2016), Doerr et al. (2017) and Martin et al. (2018) discuss both the advantages and disadvantages of using vouchers for providing ALMPs within the German labour market.

[20] During the training programme, participants receive a monthly allowance of 100 €. In addition, if they wish to attend a course that is not available locally, their commuting costs are later reimbursed.

on their employment status and income on specific dates. The earliest date is June 2016. Hence, to ensure that we observe individuals' outcomes for long enough after the completion of their training programme, we consider unemployed young people who registered at the SEA between January and December 2014. We provide additional evidence to show that there was no manipulation, i.e. that potential YG recipients did not register as unemployed at the SEA purely in order to participate in the programme (see Section 6.1). Since data on income were provided at fixed dates (June 2016, December 2017, June 2017), we can observe income for a period going between 1.5 and 3.5 years after registration. This means that our estimates capture both short- and medium-term effects on employment.

4 Data and descriptive statistics

In this study, we use individual administrative records obtained by merging data from the Latvian SEA (i.e. the public employment service) with data from the SRS (i.e. the state tax authority).

The SEA gathers data on persons who are registered as unemployed in Latvia. From the day of registration at the SEA, unemployed people receive job search assistance, which can last up to four months. During this time, SEA caseworkers screen the profiles of those who are registered as unemployed, assess their needs, and check whether or not they are eligible to attend training courses.

For all who are registered as unemployed, the SEA database contains the following information: gender, date of birth, residence, nationality, highest level of education attained, and exact date of registration, which is the starting date of the period of unemployment. In addition, for all VT participants, the database has information on the start and end dates of the training, the type of course attended, and whether the participant completed or dropped out of the course. It is also possible to observe if the individual had participated in another programme in the YG scheme before taking part in the VT training in question.

The administrative data from the SRS provide information on the labour market performance of each individual on specific dates.[21] This allows us to define an indicator of formal employment at specific times. For individuals who are formally registered as employed in the SRS database, it is also possible to observe information on their earnings, and the firm's size and industry. This information was extracted for January 2012, June 2012, December 2012, June 2013, December 2013, June 2014, December 2014, June 2015, December 2015, June 2016, December 2016 and June 2017. Because the intervention began in January 2014, the information collected between January 2012 and December 2013 was used to construct pre-intervention

[21] The SRS database collects monthly reports from employers, which declare employees' monthly incomes, insurance, working hours, the quality of the job (based on the International Standard Classification of Occupations, ISCO-88), the firm's sector (Statistical Classification of Economic Activities in the European Community, i.e. NACE category) and the firm's size (number of employees).

measures of individuals' labour market careers (e.g. employment status, monthly income, hours worked and social contributions). Data collected in June 2016, December 2016 and June 2017 serve as outcome variables for evaluating the labour market performance of individuals from 1 to 3 years after the VT programme.

The SEA database provides access to the list of all those who were on the SEA unemployment register from June 2013 to December 2015. Our initial sample is composed of 1,890 individuals who participated in the VT programme and 38,567 individuals who did not. As the programme officially started on January 1st 2014 and because individuals could enroll in the programme at any time (provided that they met the eligibility criteria), we had to impose some sample selection criteria. The research design is as follows.

First, we select an inflow sample of individuals who registered at the SEA from January 1st 2014 to December 31st 2014, and who, on the registration date, were aged between 15 and 29 years.[22] All of these individuals were eligible for the YG package and all of them received job search assistance after registration at the SEA.

Second, for each individual, we set a timeframe of one year from the date of registration to assess his or her participation in the training programme. The treatment status is defined as a binary indicator that equals 1 if the individual participated in the VT programme *within one year of the registration date*, and 0 otherwise. Hence, someone who registered at the SEA on December 31st 2014 and started the VT by the end of December 2015 is included in the treated group.[23] By contrast, control individuals are those who did not participate in the VT under evaluation or in any other training programme managed by the SEA. The final sample comprises 11,615 individuals. Of these, 898 participated in the VT (treated group). The remaining 10,717 individuals constitute the control group. In our baseline estimates, the outcome variables are measured for all individuals in June 2016, i.e. at least one year after participation. Analyses on outcomes measured later are reported in the Appendix and confirm the baseline results.

As shown by the results of the t-tests on the difference in the means, the treated and control groups are balanced in terms of nationality, since the proportion who were not Latvian is not statistically different between the groups. In terms of gender, the proportion of women is higher in the treated group than in the control group. The two groups also differ in terms of level of education and area of residence. First, those in the treated group are, on average, less educated: the proportion of unemployed people with primary education or lower and general secondary education is higher in the treated group than in the control group, whereas the proportion with professional secondary or higher education is higher in the control group. Second, the proportion of individuals living in the capital or in other cities is higher in the control group

[22] This is done to prevent potentially dynamic sample selectivity issues. Indeed, including persons considered long-term unemployed in the sample, e.g. those registered in 2012 or 2013, would have introduced potentially substantial unobserved individual heterogeneity into the analysis. For instance, those who first registered at the SEA in 2012 and remained unemployed in 2014 could be the least employable due to a lack of skills or effort in searching for a job. By selecting those registered at the SEA in 2014, we sought to limit these potential concerns.

[23] We excluded 40 people who participated in the training programme more than one year after the registration date.

than in the treated group, with more people in the treated group living in national development centres, regional centres, or rural areas. Finally, average income and number of years with positive income in the pre-treatment period (before 2014), a proxy of labour market experience, are higher in the control group than in the treated group. All in all, these statistics suggest that individuals participating in VT may be the least employable in terms of observable characteristics (past work experience, education, etc.) and perhaps also in terms of unobservable characteristics (motivation, job-search effort, etc.), and are consistent with VT programmes being primarily targeted to those lacking adequate vocational qualifications.

Table 1 reports descriptive statistics for the variables used in the analysis, for both the treated and control groups.

Table 1: Descriptive statistics.

	(A) Controls			(B) Treated			t-test	
Variable	Mean	St.Dev.	N.	Mean	St.Dev.	N.	Diff. (A)-(B)	t-stat
Outcome variables								
Employed June 2016	0.452	0.498	10705	0.413	0.493	898	0.0388*	(2.25)
Income June 2016	303.621	462.255	10705	204.864	307.492	898	98.76***	(6.29)
Employed December 2016	0.420	0.494	10705	0.398	0.490	898	0.0222	(1.29)
Income December 2016	307.948	513.522	10705	219.927	336.546	898	88.02***	(5.05)
Employed June 2017	0.434	0.496	10705	0.428	0.495	898	0.00657	(0.38)
Income June 2017	326.470	515.588	10705	254.501	367.991	898	71.97***	(4.10)
Control variables								
Female	0.486	0.500	10705	0.569	0.495	898	-0.0834***	(-4.81)
Foreign nationality	0.374	0.484	10705	0.352	0.478	898	0.0225	(1.34)
Primary or lower	0.314	0.464	10705	0.392	0.488	898	-0.0776***	(-4.79)
General secondary	0.291	0.454	10705	0.366	0.482	898	-0.0749***	(-4.72)
Professional secondary	0.229	0.420	10705	0.175	0.380	898	0.0545***	(3.76)
College or higher	0.165	0.371	10705	0.067	0.250	898	0.0981***	(7.77)
Rural area	0.545	0.498	10705	0.647	0.478	898	-0.102***	(-5.91)
Region: Kurzeme	0.176	0.381	10705	0.194	0.395	898	-0.0181	(-1.37)
Region: Latgale	0.175	0.380	10705	0.324	0.468	898	-0.149***	(-11.03)
Region: Pieriga	0.168	0.374	10705	0.116	0.320	898	0.0521***	(4.05)
Region: Riga	0.234	0.423	10705	0.109	0.312	898	0.125***	(8.63)
Region: Vidzeme	0.105	0.306	10705	0.112	0.316	898	-0.00785	(-0.74)
Region: Zemgale	0.143	0.350	10705	0.145	0.352	898	-0.00203	(-0.17)
Average income before 2014	281.419	350.715	10705	180.947	233.529	898	100.5***	(8.43)
# yrs with income>0 before T	1.903	1.793	10705	1.355	1.621	898	0.548***	(8.85)
Fraction yrs with income>0 before T	0.381	0.359	10705	0.271	0.324	898	0.110***	(8.85)
Observations							11603	

5 Empirical strategy

The main identification issue that needs to be tackled when assessing the causal effect of the VT programme on subsequent employment is that the unemployed are not randomly allocated to the VT programme. Indeed, they are screened according to their vocational qualifications, and after being offered VT, they can choose whether to participate or not. While the selection of potential beneficiaries of the programme by SEA workers according to observable characteristics (e.g. the level of individual qualifications and latent employability) is addressed by including observable variables such as the level of education or past incomes in the regressions, there may still be a self-selection problem for individuals accepting to participate in the course, driven by factors such as individual ability or motivation. Participants in the VT programme may differ from non-participants in terms of unobservable characteristics and this could have a direct impact on their employment status after the treatment period, invalidating the analysis.

In order to overcome this identification issue, we exploit a specific feature of the programme, namely the fact that the SEA gave a higher priority for participation in the VT programme to unemployed people under the age of 25, even though the YG targets all individuals in the age range of 15–29. These features of the programme make it ideal for the use of a fuzzy regression discontinuity design (FRDD), a methodological approach used when there is a threshold in a given individual attribute that determines assignment into a treatment (sharp RDD) or the probability of being treated (FRDD). In our case, the priority rule and the fact that participation ultimately remains voluntary makes the design fuzzy. That is, being subject to the priority rule increases the probability of participating in the programme but does not exactly determine programme participation.

The idea behind this identification strategy is that, conditional on the observable variables, individuals under the age of 25 were more likely to participate in the VT programme since they had priority. As individuals have no control over their age, around the cut-off age of 25, allocation to the VT programme is "as good as randomly assigned". This means that on average, treated and control units around the age of 25 have identical observable and unobservable characteristics. This relies on the assumption of "no manipulation", meaning that controls and treated units around the age of 25 cannot completely determine their position with respect to the threshold. If this was not the case, individuals with a particular interest in participating may disproportionately lie to one side of the threshold, e.g. by registering at the SEA with the specific purpose of participating in YG's VT programmes. The comparison across the threshold would then be biased because individuals to each side of the threshold would be different in terms of unobservables.

In our analysis, age is the individual attribute (or the running variable) that determines the probability of participating in the VT programme, with the age of 25 representing the threshold, or cut-off point, below which participation in the programme drastically increases due to the priority rule. Since we do not know the exact date on which the profiling phase took place, we assume that for each individual in the sample, the eligibility conditions for participation in the VT programme were assessed on the day of registration at the SEA. Hence, we measure age on the date of registration at the SEA as a continuous variable (in days). According to the priority rule, an individual who is older than 25 on the day she registers at the SEA has a very low probability of participating in the programme. By contrast, an individual who is younger than 25 on the day she registers at the SEA has a higher probability of participating in the programme, thanks to the priority rule. Since participation is voluntary, the probability of participation for those who are less than 25 years old is strictly less than one.

The running variable is centred at the cut-off point, so it is equal to 0 for someone who registers at the SEA on the day of their 25^{th} birthday and takes a negative (positive) value for those who at the time of the registration are younger (older) than 25. As such, the running variable represents the difference between age at the time of registration and the cut-off value: larger negative values correspond to individuals who are exposed to the priority rule for a longer period. Note that, since the running variable is measured for each individual on the starting date of the unemployment spell (i.e. the date of registration at the SEA, which is assumed to coincide with the date on which profiling takes place), the probability of participating in the programme for each individual in the sample jumps from zero to a positive value at the cut-off point.

We estimate the causal effect of participation in the VT programme on subsequent labour market outcomes via Instrumental Variables (IV), in particular two-stages least squares (2SLS). In our setting, the discontinuity in the probability of participation in the programme given by the age-specific rules can be used as an instrument for participation status (Angrist and Pischke, 2009).

Our equation of interest is:

$$Y_i = \beta_1 T_i + f(\tilde{x_i}) + \beta^\top \mathbf{W}_i + \epsilon_i, \tag{1}$$

where Y_i is the employment status or monthly wage of individual i at a certain point in time after the training is completed (June 2016, December 2016 and June 2017), T_i is the treatment variable, which takes a value of 1 if the individual is registered at the SEA as unemployed for the period of January – December 2014 and completes the VT programme by December 2015. $\tilde{x_i}$, is the running variable, i.e. age measured on the date of registration at the SEA and centred at the cut-off point. $f(\cdot)$ is a polynomial in the running variable (either linear or quadratic), which represents the relationship between the running variable and the outcome. $\mathbf{W}_i$ is a vector of individual covariates such as gender, nationality, level of education, residence area, number of years worked and average income in the pre-treatment period, including the intercept term. ϵ_i is an individual specific error term.

As explained above, we estimate equation (1) via 2SLS. The excluded instrument used in 2SLS is a dichotomous indicator for an individual being younger than 25, i.e. the programme priority rule $Z = \mathbf{1}(x < 25)$.

The corresponding first stage equation in the case of a linear polynomial reads as follows:

$$T_i = \gamma_1 Z_i + \gamma_2 \tilde{x}_i + \gamma^\top \mathbf{W} + \eta_i, \tag{2}$$

where Z_i is the instrument for T_i and η_i is an individual error term. In the analysis, we specify the polynomial in the running variable as either linear or quadratic.

That is, the main equation in the case of a linear polynomial becomes:

$$Y_i = \beta_1 T_i + \beta_2 \tilde{x}_i + \beta^\top \mathbf{W}_i + \epsilon_i, \tag{3}$$

where $\beta_2 \tilde{x}_i$ is a parametric function in the running variable that can be linear or quadratic.[24]

To check the sensitivity of the estimates, we estimate these models on increasingly narrower bandwidths defined in terms of the following age groups: 21–29, 22–28, 23–27 and 24–26. We discuss the results in Section 6.

6 Baseline results

In this section we present and discuss our baseline results. We first show some descriptive evidence on the first stage (i.e the effect of the priority rule on participation in the VT programme) and the reduced form equation (i.e. the direct effect of the priority rule on the outcome variable) (see Figure 1).

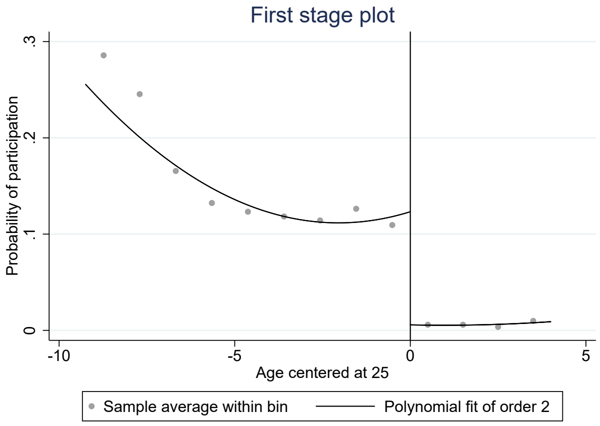
Figure 1a is a plot of the first stage, showing the probability of participating in the VT programme as a function of age at registration at the SEA.[25] We use a quadratic specification in age and do not condition on other covariates such as gender, nationality, region of residence, etc. We notice a clear jump in the probability of participating in the VT at the cut-off (age 25), due to the priority rule. The size of the jump is about 10 percentage points (pp), and we later show that the results are statistically significant at the 1% level.

Figure 1b shows the probability of being employed in June 2016 as a function of age at SEA registration, normalised at the cut-off point and using a quadratic specification in age, as before. Here, we do not observe a clear jump at the discontinuity point, although we still see that the probability of being employed in June 2016 (about 1 year after the completion of the training) is higher for youths under 25 years of age compared to those aged 25 or above. We now go beyond the descriptive evidence and discuss our regression models in more detail, starting from the first stage regressions.

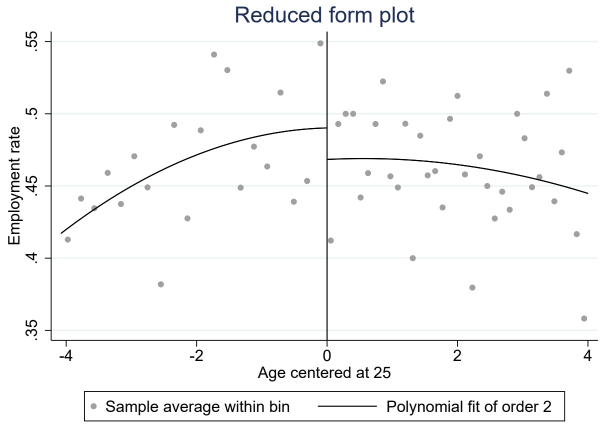
As mentioned in Section 4, we consider as treated an individual who enrols in the VT programme (under the umbrella of the YG) within 12 months of registration at the SEA. As a sensitivity check, we run

[24] As a sensitivity analysis we estimate the same model by allowing the parametric function in the running variable to have different slopes on both sides of the cut-off, as explained in Angrist and Pischke (2009) (pages 197–198). Results do not change (available upon request).

[25] Age is measured in days, hence it is defined on a continuous scale.



(a) Probability of participating in VT.



(b) Probability of being employed in June 2016.

Figure 1: Fuzzy RDD: first stage and reduced form regressions.

the same analysis, changing the enrolment window in the VT programme from 12 months to 6 months.[26] The main idea is that early-stage and late-stage enrollees in the VT programme (with respect to their date of registration at the SEA) may be two different sub-populations for whom the treatment may produce different effects. Moreover, differences between these groups may also be produced by differences in lock-in effects, according to the timing of the VT programmes. However, results do not change.[27]

Tables 2 and 3 show the results of the first stage equation for the linear specification and quadratic specification, using different age bandwidths, namely 21–29, 22–28, 23–27, 24–26. Table 2 shows that being subject to the priority rule (i.e. being younger than 25 years of age) increases the probability of participating in the programme by about 10 pp when considering individuals between the ages of 21–29, which drops to 7 pp when considering cohorts that are very close to the threshold (age 24–26). Thus, we observe a strong effect of the priority rule on participation in the programme. The F-statistics are way above the threshold of 10, and range between 135 when considering the 21–29 age bandwidth, to 13 when considering the smallest bandwidth (age 24–26). Results are very similar when we use a quadratic specification in age (Table 3). Note that the larger the bandwidth around the cut-off, the higher the precision of the estimates, as we are using more data points to fit our model. Wider age bandwidths imply that the estimates tend to be less accurate (higher bias), due to the fact that we are using data points that are far away from the cut-off. For this reason we report the estimates for different age bandwidths.

[26] Among all individuals who registered at the SEA in 2014 and participated in the training (938), 234 (24%) participated more than 6 months after their registration date. Estimation results are not reported here but are available upon request.

[27] Results are available upon request.

Table 2: First stage results, linear specification.

Window width	21–29	22–28	23–27	24–26
Age<25	0.111 ***	0.105 ***	0.089 ***	0.067 ***
	(0.010)	(0.011)	(0.014)	(0.019)
Age centred at 25	0.002	-0.002	-0.011 *	-0.034 **
	(0.002)	(0.003)	(0.006)	(0.015)
Female	0.030 ***	0.032 ***	0.039 ***	0.036 ***
	(0.005)	(0.006)	(0.007)	(0.010)
Foreign nationality	0.002	0.002	0.008	0.000
	(0.006)	(0.006)	(0.008)	(0.010)
General secondary	0.010	0.007	0.003	0.008
	(0.007)	(0.008)	(0.009)	(0.013)
Professional secondary	-0.018 ***	-0.020 **	-0.016 *	-0.005
	(0.007)	(0.008)	(0.010)	(0.013)
College or higher	-0.027 ***	-0.032 ***	-0.039 ***	-0.026 *
	(0.008)	(0.009)	(0.010)	(0.014)
Rural area	0.004	0.005	0.006	0.006
	(0.007)	(0.008)	(0.009)	(0.012)
Region: Kurzeme	0.026 ***	0.027 ***	0.025 **	0.019
	(0.009)	(0.010)	(0.012)	(0.016)
Region: Latgale	0.061 ***	0.060 ***	0.058 ***	0.042 **
	(0.009)	(0.010)	(0.012)	(0.017)
Region: Pieriga	0.001	0.002	0.002	-0.009
	(0.010)	(0.011)	(0.013)	(0.018)
Region: Vidzeme	0.021 **	0.014	0.014	-0.004
	(0.011)	(0.012)	(0.015)	(0.020)
Region: Zemgale	0.021 **	0.021 *	0.020	0.014
	(0.009)	(0.011)	(0.013)	(0.018)
Fraction yrs with income>0 before T	0.003	0.001	0.005	0.006
	(0.008)	(0.010)	(0.012)	(0.016)
Average income before T	-0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Cons	-0.042 ***	-0.041 ***	-0.046 ***	-0.028
	(0.011)	(0.012)	(0.015)	(0.021)
F-stat	135.30	91.81	43.08	13.07
N.obs	9623	7491	5141	2593

*, **, *** statistically significant at the 10%, 5% and 1% level, respectively.

Note. This table shows the estimated coefficients from the first stage regressions using a linear specification in age. The columns show the results for different age windows: 21–29, 22–28, 23–27 and 24–26. Standard errors are in parentheses.

Table 3: First stage results, quadratic specification.

Window width	21–29	22–28	23–27	24–26
Age$<$ 25	0.111 ***	0.105 ***	0.089 ***	0.068 ***
	(0.010)	(0.011)	(0.014)	(0.019)
Age centred at 25	0.002	-0.002	-0.011 *	-0.032 **
	(0.002)	(0.003)	(0.006)	(0.015)
Age_sq centred at 25	0.000	-0.000	0.005 *	0.027 *
	(0.001)	(0.001)	(0.003)	(0.014)
Female	0.030 ***	0.032 ***	0.039 ***	0.036 ***
	(0.005)	(0.006)	(0.007)	(0.010)
Foreign nationality	0.002	0.002	0.008	0.000
	(0.006)	(0.006)	(0.008)	(0.010)
General secondary	0.010	0.007	0.003	0.008
	(0.007)	(0.008)	(0.009)	(0.013)
Professional secondary	-0.018 ***	-0.020 **	-0.016 *	-0.005
	(0.007)	(0.008)	(0.010)	(0.013)
College or higher	-0.027 ***	-0.033 ***	-0.039 ***	-0.026 *
	(0.008)	(0.009)	(0.010)	(0.014)
Rural area	0.004	0.005	0.006	0.006
	(0.007)	(0.008)	(0.009)	(0.012)
Region: Kurzeme	0.026 ***	0.027 ***	0.026 **	0.018
	(0.009)	(0.010)	(0.012)	(0.016)
Region: Latgale	0.061 ***	0.060 ***	0.059 ***	0.042 **
	(0.009)	(0.010)	(0.012)	(0.017)
Region: Pieriga	0.001	0.002	0.002	-0.009
	(0.010)	(0.011)	(0.013)	(0.018)
Region: Vidzeme	0.021 **	0.013	0.015	-0.004
	(0.011)	(0.012)	(0.015)	(0.020)
Region: Zemgale	0.021 **	0.021 *	0.021	0.014
	(0.009)	(0.011)	(0.013)	(0.018)
Fraction yrs with income$>$ 0 before T	0.003	0.001	0.006	0.007
	(0.008)	(0.010)	(0.012)	(0.016)
Average income before T	-0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Cons	-0.043 ***	-0.040 ***	-0.053 ***	-0.039 *
	(0.011)	(0.013)	(0.016)	(0.022)
F-stat	135.29	91.52	43.76	13.63
N.obs	9623	7491	5141	2593

*, **, *** statistically significant at the 10%, 5% and 1% level, respectively.

Note. This table shows the estimated coefficients from the first stage regressions, using a quadratic specification in age. The columns show the results for different age windows: 21–29, 22–28, 23–27 and 24–26. Standard errors are in parentheses.

We then proceed by showing the results from the naive ordinary least squares (OLS), the reduced form, and the 2SLS regression models, to offer a complete overview of the training being evaluated. In all specifications, we use an indicator for being employed in June 2016 and one's gross monthly income as outcomes of interest (if the individual is not working, we impute a monthly income equal to 0).[28]

Tables are organised into three different panels, respectively showing the estimates from the ordinary least square (OLS), the reduced form (RF), and instrumental variable (IV) regressions reported for different age bandwidths around the cut-off and using both linear and quadratic specifications.

OLS estimates from Table 4 show that participating in the VT programme has a positive but statistically non-significant effect on the probability of being employed in June 2016, independent of the functional form chosen for the polynomial in age. However, as discussed in the previous section, OLS estimates could be biased as individuals may self-select into participation in the VT and unobserved factors such as ability or motivation can dilute the effect. For this reason, we rely on a fuzzy RDD design and show both the reduced form and the IV estimates for each specification. We see that participating in the VT programme increases the probability of being employed in June 2016 by 15.2 pp when considering the 21–29 age bandwidth, although this result is not statistically different from 0 and remains similar when we narrow the bandwith. Finally, the IV estimates also show no effect of participating in the VT programme in terms of employability, which is in line with the findings from the literature analysing youth participation in ALMPs in developed countries.[29] This result is consistent with the fact that despite having a strong first stage, the estimate from the reduced form equation shows a zero effect of the priority rule on employability.

We now discuss Table 5, where we report the estimates for gross monthly income as of June 2016 (income is coded as zero for unemployed individuals). As before, the OLS estimates and the reduced form estimates both show a positive but statistically non-significant association between participation in the VT programme and monthly income, independent of the age specification and bandwidth chosen. As expected, the IV estimates are positive but statistically non-significant. We find similar results for employment outcomes observed in December 2016 and June 2017 (see Tables C.1- C.4 in Appendix C.)

In general, our findings are in line with the conclusions of similar ALMP studies from other European countries, as reviewed in Section 2. Most do not find any effect of training programmes in either the short or medium term. One of the few studies that finds a moderate effect in the long term is by Blundell et al. (2004), but in their study, the intervention in question is related to job-search assistance and does not involve training in the classroom.

[28] As an extension, we also show the results for employment outcomes measured in December 2016 and July 2017, up to two years after the completion of the training. In this case, we rely on smaller samples as we have to restrict the definition of treated and control groups due to the administrative burden of extracting the data from the tax records and matching them with the SEA register of unemployed persons. See Tables C.1- C.4 and the related discussion in Appendix C.

[29] The IV estimates are somehow sensitive to the choice of bandwidth but are never statistically significant at conventional levels.

Table 4: Results for the probability of being employed in June 2016.

Window width	21–29	22–28	23–27	24–26
		Linear		
		OLS		
Treated	0.001	0.005	0.008	0.037
	(0.021)	(0.023)	(0.028)	(0.041)
		RF		
Age<25	0.017	0.019	0.008	0.028
	(0.019)	(0.022)	(0.027)	(0.039)
		IV		
Treated	0.152	0.179	0.094	0.414
	(0.176)	(0.212)	(0.307)	(0.586)
F-stat	135	91.8	43.1	13.1
N.obs	9623	7491	5141	2593
		Quadratic		
		OLS		
Treated	0.001	0.005	0.007	0.037
	(0.021)	(0.023)	(0.028)	(0.041)
		RF		
Age<25	0.016	0.018	0.009	0.029
	(0.019)	(0.022)	(0.027)	(0.039)
		IV		
Treated	0.147	0.175	0.103	0.418
	(0.176)	(0.213)	(0.305)	(0.574)
F-stat	135	91.5	43.8	13.6
N.obs	9623	7491	5141	2593

*, **, *** statistically significant at the 10%, 5% and 1% level, respectively.
Note. This table shows the OLS, reduced form, and IV results, using an indicator for being employed in June 2016 as the outcome. We use two different specifications for age: linear (top panel) and quadratic (bottom panel). The columns show the results for different age windows: 21–29, 22–28, 23–27 and 24–26. The control variables are the same as reported in the first stage regressions. Standard errors in parenthesis.

A general concern with the estimates based on the FRDD is that the magnitude of the IV estimates is bigger in absolute value compared to the OLS estimates. This could either be due to a lack of precision or the presence of weak instruments. However, strong positive effects from the first stage equation and the lack of statistical significance in the reduced form regressions seem to exclude weak instruments as the primary reason for finding no effect of the VT programme on employment and monthly income. A further consideration to be made is that from a parametric point of view, estimating a FRDD is equivalent to estimating an IV. In a heterogeneous-effect setting, this implies that the IV estimators provide local estimates on the compliers, i.e. the individuals whose participation status is affected by the priority rule. Therefore, it

Table 5: Results for monthly income in June 2016.

Window width	21–29	22–28	23–27	24–26
		Linear		
		OLS		
Treated	-44.486 **	-55.746 ***	-47.940 *	-8.794
	(18.612)	(21.214)	(25.685)	(37.339)
		RF		
Age< 25	10.618	3.773	-2.308	46.694
	(17.543)	(20.307)	(24.954)	(35.079)
		IV		
Treated	95.561	35.768	-26.003	697.955
	(158.084)	(192.343)	(280.314)	(557.019)
F-stat	135	91.8	43.1	13.1
N.obs	9623	7491	5141	2593
		Quadratic		
		OLS		
Treated	-44.519 **	-55.900 ***	-47.769 *	-9.809
	(18.610)	(21.215)	(25.693)	(37.365)
		RF		
Age< 25	10.024	3.321	-2.584	47.866
	(17.546)	(20.314)	(24.968)	(35.110)
		IV		
Treated	90.188	31.523	-28.878	700.298
	(158.023)	(192.611)	(278.181)	(545.866)
F-stat	135	91.5	43.8	13.6
N.obs	9623	7491	5141	2593

*, **, *** statistically significant at the 10%, 5% and 1% level, respectively.
Note. This table shows the OLS, reduced form, and IV results using gross monthly income registered in June 2016 as the outcome. We use two different specifications for age: linear (top panel) and quadratic (bottom panel). The columns show the results for different age windows: 21–29, 22–28, 23–27 and 24–26. The control variables are the same as reported in the first stage regressions. Standard errors in parenthesis.

is useful to know as much as possible about what is the sub-population whose behaviour is changed by the instrument (the so-called compliant population).

In the following section, we will tackle these issues by first showing that our FRDD design is well-defined, and second, by more closely analysing the characteristics of the compliant sub-population.

6.1 Testing for manipulation

The validity of the FRDD rests on the local randomisation of the treatment status, that is, the fact that individuals are not able to precisely control the running variable. In our case, the running variable is age measured at registration at the SEA, using the exact date of birth.

Hence, the underlying identification assumption is that individuals just below and just above the threshold of 25 years old are comparable, except for their exposure to the priority rule, which only applies to those below the cut-off. Such an assumption may be violated if individuals can anticipate participation in the intervention, that is, when many people below age 25 suddenly start registering at the SEA offices (after January 2014) specifically because they want to participate in the YG's VT programmes. In this scenario, we would observe a peak in SEA registration by persons 25 years of age. Furthermore, anticipation would bias our results if the anticipatory behaviour is more pronounced in certain selected groups, i.e. the most motivated NEETs.

To make this more explicit, assume that the priority rule was well known amongst the target population, i.e. the pool of NEETs aged 15–29. Assume that the most motivated NEETs below 25 years of age decide to register at the SEA in 2014 in order to participate in the training programme. Moreover, assume that the registration rate among the NEETs above 25 remains unaffected, since they are aware of having little chance of participating in the training programme. In this setting, by comparing registered unemployed persons below the age of 25 with those above 25 years of age, we would be comparing individuals who are also different in terms of unobserved characteristics (e.g. motivation). This would violate the identifying assumption underlying the RDD, i.e. that individuals just below and just above the threshold are similar in all respects except for the treatment assignment.

To check for the presence of manipulation, we provide three pieces of evidence. First we show that the distribution of the running variable is continuous around the cut-off. From Figure 2, we clearly see that this condition is fulfilled. As a second check, we run a formal test as in McCrary (2008). The logic behind the testing procedure is that if individuals had precise control over the assignment process, we would expect the density of the running variable to exhibit a jump at the cut-off age, being higher on the left-hand side of the cut-off (if the treatment is assigned to individuals with smaller values with respect to the cut-off and everybody is willing to receive the treatment). Conversely, if the individuals have imprecise control over assignment, the density of the running variable should be continuous around the cut-off. Hence, the McCrary test allows for the testing of the continuity of the density of the running variable around the threshold. As can be seen from Figure 3, there is no discontinuity in age at the cut-off.

In addition, to be sure that YG did not change the incentives for registering at the SEA, we plot the age distribution at SEA registration in 2013, before the start of the YG (see Figure 4). Figure 4 suggests that the age at SEA registration for 2013 is very similar to that observed during 2014, shown in Figure 2. Note

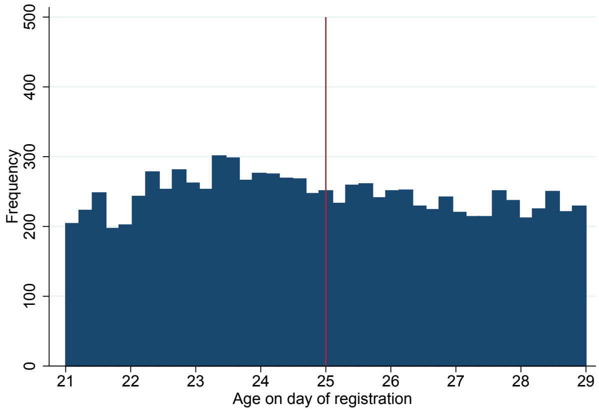


Figure 2: Distribution of age at SEA registration in 2014.

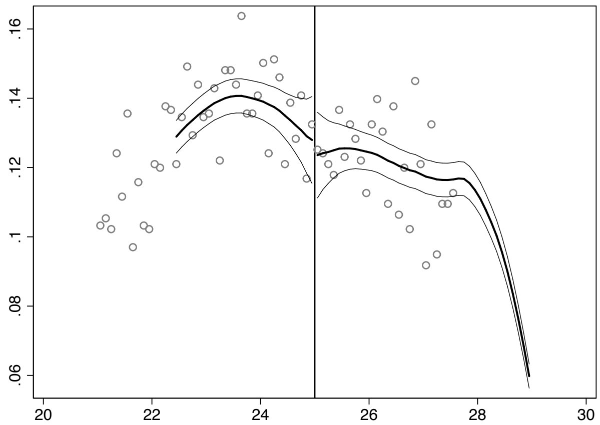


Figure 3: McCrary test

that in Figure 4, the vertical red line at 25 years of age indicates the "placebo" priority rule for participating in the training programme, even though it was not yet put in place. Therefore, in 2013, there is no reason to expect a peak at 25 in the age distribution at registration. By contrast, the absence of the peak in the corresponding graphs in 2014 (see Figure 2), when the YG was in place, supports our identification strategy. In Figure 5, we also show the number of individuals aged 15–29 registered at the SEA, by date of entry for 2013 and 2014. The vertical line on January 1st 2014 defines the introduction of the YG. According to the figure, registration at the SEA is not uniformly distributed over time, but follows a trend. The number of registrations increases in the second semester of 2013, reaches a peak in January 2014, and then decreases in the first semester of 2014. This does not invalidate our analysis.

Lastly, we provide further informal tests that there is no discontinuity in other individual characteristics before the treatment takes place. In the plots in Figure 6, we show the distribution of characteristics such as level of education, average income calculated over five distinctive time points before participation in the training programme (between January 2012 and December 2013), and area of residence (dummy for residing in the capital city, Riga). The figure also explains why we discard very young individuals (i.e. those aged 15–21) from the FRDD analysis, as those individuals are very likely to differ from those around the cut-off age of 25 in terms of their observable and unobservable characteristics (e.g. school drop-outs). Figure 6 shows that there is no discontinuity at the cut-off age in the distribution of each of these observable variables. This additional evidence supports the empirical strategy used in the analysis.

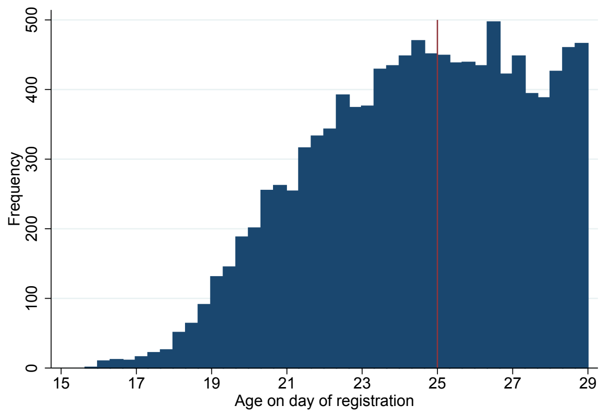


Figure 4: Distribution of age at SEA registration in 2013.

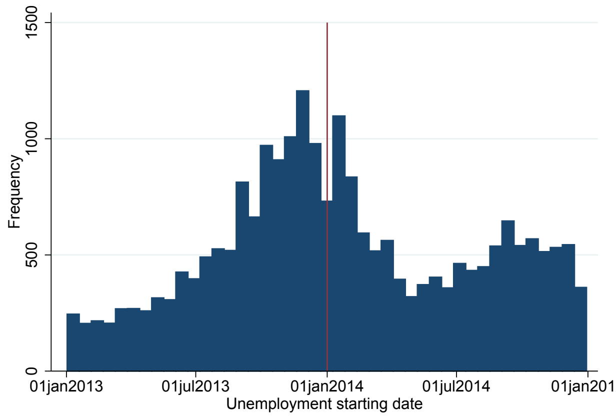


Figure 5: Distribution of date of registration at SEA in 2013-2014.

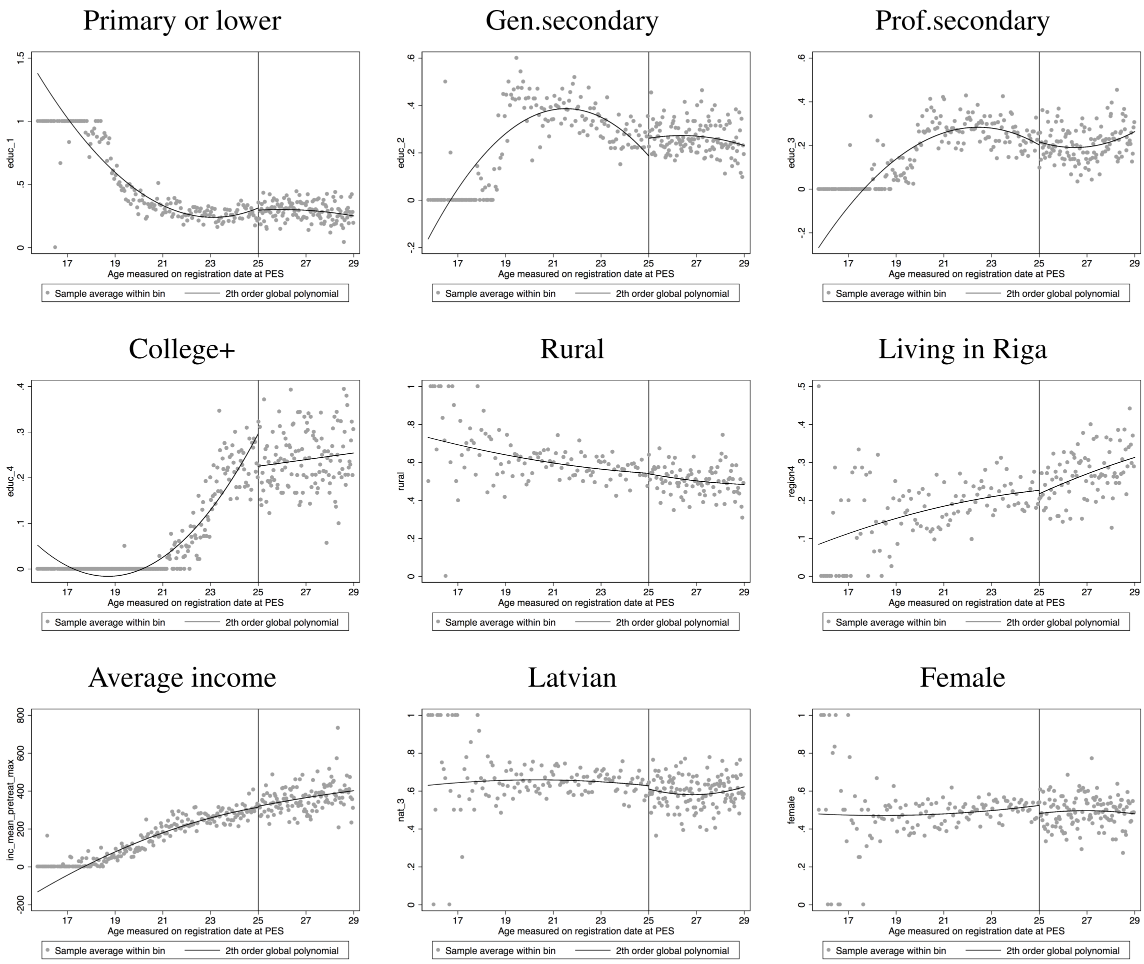


Figure 6: Testing for continuity.

Note: This figure shows that there is no discontinuity at the cut-off point in the distribution of the covariates for which we control for, namely, level of education (primary or lower, general secondary, professional secondary, college or more), living in a rural area, living in the capital city, average income, being of Latvian nationality, being female).

7 Heterogeneous effects

7.1 Characterization of the compliant sub-population

Each instrument helps identify a unique causal parameter. Using a FRDD setting, at best, we can identify the "local effect" (Local Average Treatment Effect, i.e. LATE) for the group of compliers, that is, the sub-population whose treatment status (participation in the YG or non-participation) is determined by the instrument (the priority rule).

The compliant sub-population associated with the priority rule instrument is composed of youths who, in absence of the rule, would not have participated in the training programme.

Although the compliers cannot be listed from observed data, we can learn something about their characteristics by exploiting Bayes theorem when both the endogenous variable the instrument are binary.[30] This can be done based on pre-treatment characteristics (X_i), which follow a Bernoulli distribution (dummy variables), in order to answer questions such as: Are compliers of the VT programme more likely to be female or reside in rural areas, compared to the full sample? For this exercise, we use the following binary indicators, which are observed before the individuals participate in the training programme: an indicator for being female, an indicator for having Latvian citizenship, an indicator for living in a rural area and one for living in the capital city, four indicators for level of education (i.e. primary education or lower, general secondary, professional secondary and college), and finally, an indicator for not observing any positive income before participation in the training programme.

Table 6 summarises our results. We report the unconditional mean of the pre-treatment dummy variables (X) calculated over the whole sample (column 3), the conditional mean for the complier sub-population (column 4), and in column 5, we report the relative likelihood that a complier has $X = 1$ (that is, the ratio between column 4 and column 3).

With respect to the whole sample, compliers are 14% more likely to be female, 1% more likely to have Latvian citizenship,[31] 13% more likely to live in rural areas, and 52% less likely to be living in Riga. Concerning educational achievement, we find that the compliers are 21% more likely to have primary education or lower and 13% more likely to have a general secondary education. Conversely, they are less likely to have a professional secondary or college education. Finally, we find that compliers are 20% more likely to have had no income in the period of 2012–2013.[32] All in all, this characterisation confirms that before enrolling in the programme, the compliers, on average, seem to have poorer skills and lower employability compared to the whole sample, which is consistent with the aim of the programme.

[30] See Angrist (2004) and Angrist and Pischke (2009) for the methodology.

[31] Our data show that about 62% of the individuals in our sample have Latvian citizenship. The remainder is composed of individuals of Russian nationality (17%), as well as citizens of neighbouring countries, and 15% is unspecified.

[32] Past income is measured in the months of January 2012, June 2012, December 2012, June 2013, and December 2015.

Table 6: Characterization of the compliant population.

Exogenous variable (X)	N	$Pr(X=1)$	$Pr(X=1 \mid \text{compliers})$	$\frac{Pr(X=1 \mid \text{compliers})}{Pr(X=1)}$
			$Z = age < 25$	
Female	11603	0.492	0.561	1.140
Latvian citizenship	11603	0.627	0.634	1.011
Rural area	11603	0.553	0.626	1.133
Riga	11603	0.224	0.119	0.529
Primary or lower	11603	0.320	0.388	1.213
General secondary	11603	0.297	0.338	1.139
Professional secondary	11603	0.225	0.150	0.668
College	11603	0.157	0.096	0.608
No income before T	11603	0.344	0.414	1.206

Note: Using Bayes theorem and exploiting the monotonicity assumption, one gets:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

$$\frac{P(x_{1i} = 1|D_{1i} > D_{0i})}{P(x_{1i} = 1)} = \frac{P(D_{1i} > D_{0i}|x_{1i} = 1)}{P(D_{1i} > D_{0i})} = \frac{E[D_i|Z_i = 1, x_{1i} = 1] - E[D_i|Z_i = 0.x_{1i} = 1]}{[E[D_i|Z_i = 1] - E[D_i|Z_i = 0]}$$

Angrist and Pischke (2009), pages 261–263.

7.2 Data limitations and robustness checks

In this section, we discuss some of the limitations of our study, mostly related to data availability and their content.

Since we use information from tax records to check the employment status of both participants and non-participants after the end of the training (2015-2017), we are only able to observe formal employment, disregarding other work arrangements such as self-employment or informal employment. Furthermore, as in other studies, it is not possible to determine whether individuals continue their education after the training experience, if they migrate, or whether they commute to work in neighbouring countries (e.g. Estonia, Lithuania, etc). We discuss each of these limitations in turn.

First, disregarding self-employment may bias the impact of the programme downwards if the programme were to boost entrepreneurship, contributing to the creation of new jobs. Unfortunately, tax registers are not integrated in Latvia, so we are only able to employ monthly income for a job spell from formal dependent employment. More precisely, the Latvian State Revenue Service (SRS) collects tax declarations only for self-employed persons earning an income above a certain threshold (i.e. above 213 €, according to Eurofound). To check for the presence of any self-employment prevalence among our VT participants and non-participants, we exploit information on the industry in which they worked before the implementation of YG, namely, before January 2014. We rely on the NACE code, using a one-digit classification. We then compare the distribution of participants and non-participants in our data across different sectors, with the distribution of the self-employed across sectors obtained using data on Latvia from the European Union Labour Force Survey (EU-LFS).[33] Since the EU-LFS provides information on individuals' ages using 5-year bands (0–4, 5–9, 10–14, 15–19, and so on), we compare the results for the age-group 15–29 (the policy's target group).

Based on the EU-LFS data for the period of 2012–2014, the three sectors with the highest share of self-employed individuals are "Agriculture" (19.17%), "Other service activities" (28.76%) and "Professional activities" (13.50%). In contrast, the majority of individuals in our data (based on Latvian tax records) worked in sectors such as "Water supply" (28, 34%), "Information" (16,76%) and "Arts" (9.66%) during the same reference period. Therefore, the VT programme seems to target unemployed individuals without any job experience or with experience as employees. More generally, Latvia is the only Baltic country that has registered a decline in self-employment over time, accounting for less than 10% of total employment (Hazans, 2005).

A second point is related to the presence of informal employment and in this regard, two issues are worth discussing. First, with the data at hand, it is not possible to measure informal employment. From the point of view of the programme providers, informal employment is not an outcome of interest, therefore not observing it does not negatively bias the impact of the programme, as the main objective of the programme is to boost formal employment. Second, official labour statistics from Latvia suggest that the phenomenon of undeclared labour has decreased over the years. Hazans (2011) provides an overview of informal employment in 30 European countries, using data from the European Social Survey (ESS) for the period of 2008–2009, and finds that informal employment decreases from South to West to East to North. Overall, he shows a declining trend of informal employment for most countries. In the Baltic countries, Estonia is the country with the lowest informality rate, followed by Lithuania and Latvia, where only about 8% of the labour force is working informally. However, he finds that informality is higher for self-employed occupations than in dependent employment. Given that our intervention mainly targets unemployed persons with no working experience or with past experience as dependent workers, we expect a small bias even in this case, if any.

Third, given that we look at young people, another possible outcome after completing the programme may be to enroll in a university or vocational programme. This would be a positive outcome of the programme, despite not being among the main aims, which are to provide sufficient qualifications for unemployed youths to find a job.

Fourth, another possibility is that after completing the programme, unemployed youths may seek em-

[33] The EU-LFS is a large household sample survey providing quarterly and yearly data on the labour participation of people aged 15 and over, as well as on persons outside the labour force. The dataset includes all persons aged 15 years and over living in private households.

Table 7: Percentage of workers aged 15–29 in various sectors for the period 2012-2014: EU-LFS for Latvia (self-employed) vs. Latvian tax records (all workers).

	EU-LFS (self-employed)	Latvian tax data (all workers)
Agriculture	19.17	1.86
Mining and quarrying	0.00	1.32
Manufacturing	3.07	1.05
Electricity	0.00	2.29
Water supply	0.00	28.34
Construction	6.54	1.48
Wholesale	3.67	7.87
Transportation	2.50	0.79
Accommodation	1.66	3.73
Information	9.94	16.76
Financial activities	0.00	2.16
Real estate activities	2.56	5.32
Professional activities	13.50	5.32
Administrative activities	6.22	4.08
Public administration	0.00	1.63
Education	2.48	0.30
Human health	0.72	3.24
Arts	1.67	9.66
Other service activities	28.76	2.80

Note: Author's calculations. We show the percentage of self-employed workers aged 15–29 out of the total employed across different sectors. The first column uses data from the EU-LFS. The second column, uses Latvian tax data. In both cases, we have pooled the data for the years 2012–2014 and use the NACE job sector classification (first digit).

ployment in neighbouring countries. If they commute, we would observe them as employed in another country according to Latvian tax data. This would negatively bias the impact of the programme. In Eastern countries, migration and commuting are quite important. To get a sense of the importance of this issue in our analysis, we run the analysis excluding from our sample individuals who reside in municipalities that share the border with: Estonia, Belarus, Russia and Lithuania. The results are not different from the baseline ones where we consider all municipalities,[34] so we may exclude this possibility.

8 Conclusion

In this paper, we provide new evidence from the implementation of the Youth Guarantee (YG) initiative in Latvia. In particular, we focus on the evaluation of a vocational (VT) training programme targeted at

[34] The estimated coefficients are of the same magnitude, but statistically not different from zero. Results are available upon request.

unemployed youths aged 15–29 who are not NEETs. We exploit a fuzzy regression discontinuity design (FRDD) thanks to a specific eligibility criterion adopted by the Latvian government, which gave a higher priority for participation in this programme to young people under the age of 25. In this setting, the priority rule provides presumably exogenous variation for youth participation in the VT programme and allows us to estimate its causal effect of on labour market outcomes (employment status and income) later in life, and net out confounding effects due to self-selection and motivation. For this, we rely on rich administrative data from the State Employment Agency (SEA), which provides information on the population of registered unemployed individuals at a given date (including both participants and non-participants in the VT programme), and match it with data from the State Revenue Service (SRS), which provides information on individuals' income at specific dates before and after the programme. Our results show that the priority rule strongly predicts youth participation in the programme. The results from our first stage are strong and statistically significant, whereas we find a positive but non-significant effect of the training programme on employment status and income. Since we analyse labour market outcomes between June 2016 and June 2017, after the VT programme's completion, we can reject the possibility that the lack of an effect is explained by "lock-in effects", i.e. a less intense job search by individuals participating in the programme. Our results are broadly consistent with the existing literature on ALMPs targeted at youths, which often find little to no effect when these measures are not coupled with demand-driven interventions (e.g. tax rebates or firm-provided training).

Several explanations could be proposed for the lack of a positive effect of YG's VT programme on dependent employment in the formal sector and on declared labour incomes. First, since all registered persons at the SEA were provided with job search assistance, it may be the case that these services were just as effective for both VT participants and non-participants, and the VT programme did not have any additional positive effect. A second explanation is that although the VT provision was implemented according to an estimate of the sectors and occupations in high demand, it might not have taken into account the geographical distribution of labour demand and supply. Alternatively, it may have been imprecise, or demand might have been met by immigrant workers. Last but not least, our estimates may also suggest that the very low-skilled individuals targeted by the programme may be more likely to find work in the informal sector, which would make it very difficult to estimate the effect of the programme on individuals' overall employability and on declared and undeclared incomes. Unfortunately, owing to data limitations, we are not able to further explore these potential explanations.

References

Alfonsi, L., O. Bandiera, V. Bassi, R. Burgess, I. Rasul, M. Sulaiman, and A. Vitali (2017, December). Tackling Youth Unemployment: Evidence from a Labour Market Experiment in Uganda. Technical report, Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD, LSE.

Angrist, J. D. (2004). Treatment effect heterogeneity in theory and practice. *The Economic Journal 114*(494), C52–C83.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton: Princeton University Press.

Attanasio, O., A. Guaran, C. Medina, and C. Meghir (2017, April). Vocational training for disadvantaged youth in Colombia: A long–term follow–up. *American Economic Journal: Applied Economics 9*(2), 131–43.

Attanasio, O., A. Kugler, and C. Meghir (2011). Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial. *American Economic Journal: Applied Economics 3*(3), 188–220.

Biewen, M., B. Fitzenberger, A. Osikominu, and M. Paul (2014). The effectiveness of public-sponsored training revisited: The importance of data and methodological choices. *Journal of Labor Economics 32*(4), 837–897.

Blundell, R., M. Costa Dias, C. Meghir, and J. Van Reenen (2004). Evaluating the employment impact of a mandatory job search program. *Journal of European Economic Association 2*(4), 569–606.

Cabasés Piqué, M., V. M. Pardell, and T. Strecker (2016). The EU youth guarantee? A critical analysis of its implementation in Spain. *Journal of Youth Studies 19*(5), 684–704.

Caliendo, M. and R. Schmidl (2016, Jan). Youth unemployment and active labor market policies in Europe. *IZA Journal of Labor Policy 5*(1).

Card, D., P. Ibarrarán, F. Regalia, D. Rosas-Shady, and Y. Soares (2011). The labor market impacts of youth training in the dominican republic. *Journal of Labor Economics 29*(2), 267–300.

Card, D., J. Kluve, and A. Weber (2010). Active labour market policy evaluations: a meta-analysis. *Economic Journal 120*(548), F452–F477.

Card, D., J. Kluve, and A. Weber (2015). What works? A meta analysis of recent active labor market program evaluations. Working Paper 21431, National Bureau of Economic Research.

Carling, K. and L. Larsson (2005). Does early intervention help the unemployed youth? *Labour Economics 12*(3), 301–319.

Cave, G., H. Bos, F. Doolittle, and C. Toussaint (1993). JOBSTART. Final report on a program for school dropouts.

Couch, K. A. (1992). New Evidence on the Long–Term Effects of Employment Training Programs. *Journal of Labor Economics 10*(4), 380–388.

Doerr, A., B. Fitzenberger, T. Kruppe, M. Paul, and A. Strittmatter (2017). Employment and earnings effects of awarding training vouchers in germany. *Industrial and Labour Relations Review 70*(3), 767–812.

Escudero, V. and E. L. Mourelo (2015). The youth guarantee programme in Europe: Features, implementation and challenges. Technical report, International Labour Organisation Office, Research Department, Working Paper 4.

Eurostat (2017). Employment and unemployment: European union labour force survey.

Fitzenberger, B. and R. Völter (2007). Long–run effects of training programs for the unemployed in East Germany. *Labour Economics 14*(4), 730–755.

Flores-Lagunes, A., A. Gonzales, and T. Neumann (2010). Learning but not earning? The impact of JOB CORPS training on Hispanic youth. *Economic Inquiry 48*(3), 651–667.

Grubb, M. (1999). Lessons from education and training for youth: five precepts.

Hämäläinen, K., U. Hämäläinen, and T. J. (2014). The labour market impacts of a youth guarantee: Lessons for europe? Technical report, Government Institute for Economic Research Working Papers, No. 60.

Hazans, M. (2005). Working too hard? working and employment conditions in new eu member states. convergence or diversity?

Hazans, M. (2011). Informal Workers Across Europe: Evidence from 30 Countries. Technical report, IZA Institute of Labour, Discussion Papers Series No. 5871.

Heckman, J. J., R. J. Lalonde, and J. A. Smith (1999). The economics and econometrics of active labor market programs. *Handbook of Labor Economics 3*, 1865 – 2097.

Ibarrarán, P., J. Kluve, L. Ripani, and D. Rosas Shady (2015). Experimental Evidence on the Long-Term Impacts of a Youth Training Program. IZA Discussion Papers 9136, Institute for the Study of Labor (IZA).

Ibarrarán, P., L. Ripani, B. Taboada, J. M. Villa, and B. Garcia (2014). Life skills, employability and training for disadvantaged youth: Evidence from a randomized evaluation design. *IZA Journal of Labor & Development 3*(1), 10.

Kluve, J. (2010). The effectiveness of European active labor market programs. *Labour Economics 17*, 904–918.

Lechner, M., R. Miquel, and C. Wunsch (2011). Long–run effects of public sector sponsored training in West Germany. *Journal of the European Economic Association 9*(4), 742–784.

Martin, H., L. Michael, and S. Anthony (2018). Direct and indirect effects of training vouchers for the unemployed. *Journal of the Royal Statistical Society: Series A 181*(2), 441–463.

Martin, J. P. and D. Grubb (2001). *What works and for whom: a review of OECD countries' experiences with active labour market policies*. Uppsala: Working Paper Series 2001:14, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Mascherini, M. (2012). Youth Guarantee: Experiences from Finland and Sweden. Technical report, Eurofund, Dublin.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Pastore, F. (2015). The European youth guarantee: Labor market context, conditions and opportunities in Italy. *IZA Journal of European Labor Studies 4*(1), 1–22.

Schochet, P. Z., J. Burghardt, and S. McConnell (2008). Does job corps work? Impact findings from the national job corps study. *The American Economic Review 98*(5), 1864–1886.

Strittmatter, A. (2016). What effect do vocational training vouchers have on the unemployed? *IZA World of Labor* (316).

Appendix A: The VT programme

Figure 7 summarises the procedure that a young candidate needs to follow to receive job-search assistance and eventually register in any ALMP programme.

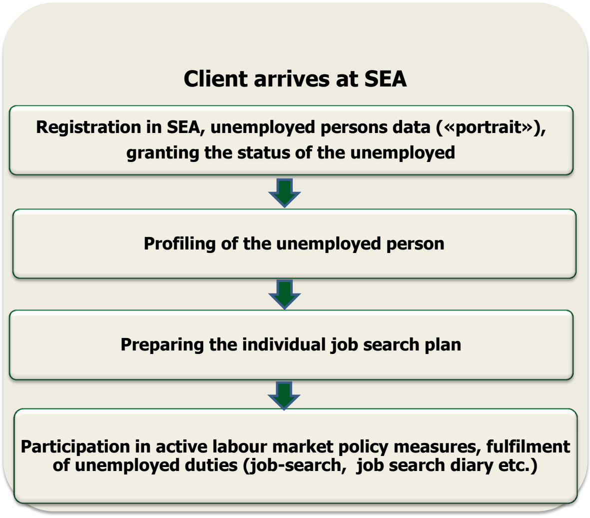


Figure 7: Procedure to receive job-search assistance and participate in ALMP programmes.

The VT programme itself is implemented using a voucher type of procedure which consists in the following steps:

1. *Registration*: after registering at the SEA, unemployed individuals receive job-search assistance. In order to promote efficient and targeted provision of the measures offered by the SEA to the unemployed, the SEA carries out the profiling of unemployed, which includes the determination of the most suitable available active employment measures for the unemployed and the preferable sequence for receiving the measures. Such profiling takes place in a meeting between the unemployed and the SEA officer. During the meeting the unemployed applies for participation in the training measures. The SEA officer checks that candidates satisfy the eligibility requirement for participation before registering the application in the SEA database.

2. *Selection of a programme*: the unemployed may choose a suitable programme from the list of training programmes (approximately 75 VT programmes).

3. *Voucher receipt*: the SEA officer makes a phone call and invites the unemployed to receive a training voucher. The voucher consists of two parts: one is for the training provider and the other should be returned to the SEA officer. It contains information on the maximum amount of expenses covered by the SEA.

4. *Choice of the training provider*: the unemployed selects a training provider within the first 10 working days after the SEA's job search assistance. The choice is made from the list of procured training providers published on the SEA's website.[35] The training provider has to determine the suitability of a person for participating in a training programme.

Appendix B: Additional figures

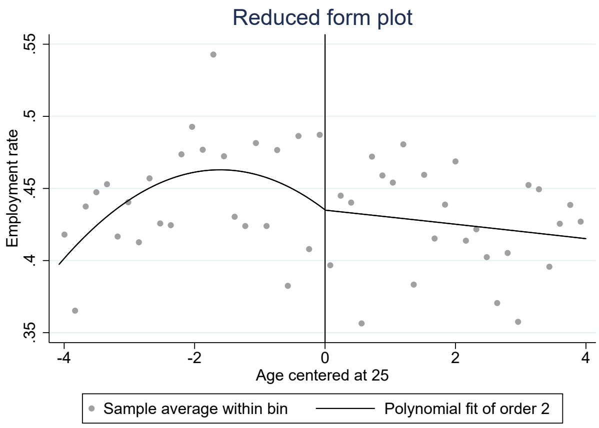


Figure B.1: Probability of being employed in December 2016, by age of registration at SEA.

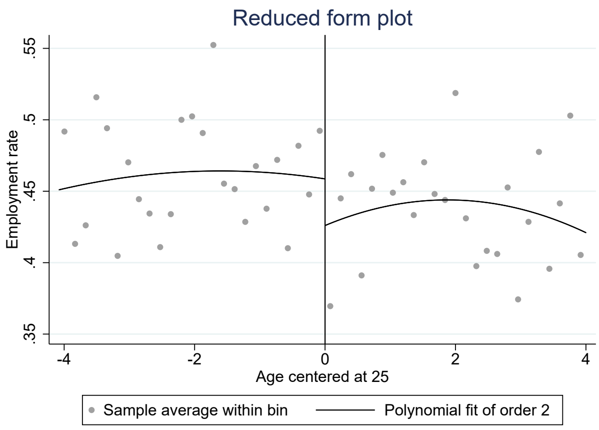


Figure B.2: Probability of being employed in June 2017, by age of registration at SEA.

[35] However, other training providers may also be selected, if they are ready to make an agreement with the unemployed and follow the next procurement procedure.

Appendix C: Additional tables

We extend our analysis by looking at labour market outcomes in December 2016 and June 2017 (Table C.1 and C.3), respectively as from 1 to 3 years after the completion of the programme. In line with the results in Table 4, we also find positive effects of participating in the VT on employment in December 2016 and June 2017, but the results are not statistically different from 0. Participating in the VT predicts an increase the probability of being employed 1 to 1,5 years after by 3.5 pp (Table C.1). Similarly to Table 4 the IV estimates are not statistically significant, pointing to a zero effect.

Table C.1: Results for the probability of being employed in December 2016.

Window width	21–29	22–28	23–27	24–26
		Linear		
		OLS		
Treated	0.018	-0.002	0.015	0.057
	(0.021)	(0.023)	(0.028)	(0.041)
		RF		
Age< 25	0.016	-0.003	0.005	0.040
	(0.019)	(0.022)	(0.027)	(0.038)
		IV		
Treated	0.141	-0.027	0.060	0.601
	(0.174)	(0.210)	(0.305)	(0.588)
F-stat	135	91.8	43.1	13.1
N.obs	9623	7491	5141	2593
		Quadratic		
		OLS		
Treated	0.018	-0.002	0.014	0.055
	(0.021)	(0.023)	(0.028)	(0.041)
		RF		
Age< 25	0.016	-0.003	0.007	0.043
	(0.019)	(0.022)	(0.027)	(0.038)
		IV		
Treated	0.140	-0.029	0.078	0.623
	(0.174)	(0.211)	(0.302)	(0.578)
F-stat	135	91.5	43.8	13.6
N.obs	9623	7491	5141	2593

*, **, *** statistically significant at the 10%, 5% and 1% levels.
Note. This table shows the OLS, reduced form and IV results, using as outcome an indicator for being employed in December 2016. We use two different specifications for age: linear (top panel) and quadratic (bottom panel). The columns show the results for different age windows: 21–29, 22–28, 23–27 and 24–26. The control variables are the same reported in the first stage regressions. Standard errors in parenthesis.

Table C.2: Results for monthly income in December 2016.

Window width	21–29	22–28	23–27	24–26
		Linear		
		OLS		
Treated	-30.018	-56.412 **	-43.760	-8.044
	(20.946)	(23.265)	(27.762)	(40.890)
		RF		
Age<25	8.955	-8.143	5.177	57.312
	(19.740)	(22.269)	(26.969)	(38.411)
		IV		
Treated	80.593	-77.205	58.324	856.672
	(177.649)	(210.684)	(303.362)	(619.197)
F-stat	135	91.8	43.1	13.1
N.obs	9623	7491	5141	2593
		Quadratic		
		OLS		
Treated	-30.027	-56.390 **	-44.342	-10.186
	(20.947)	(23.267)	(27.769)	(40.906)
		RF		
Age<25	8.783	-8.072	5.986	59.732
	(19.746)	(22.277)	(26.982)	(38.433)
		IV		
Treated	79.022	-76.626	66.892	873.888
	(177.636)	(211.011)	(301.107)	(608.307)
F-stat	135	91.5	43.8	13.6
N.obs	9623	7491	5141	2593

*, **, *** statistically significant at the 10%, 5% and 1% levels.

Note. This table shows the OLS, reduced form and IV results, using as outcome gross monthly income registered in December 2016. We use two different specifications for age: linear (top panel) and quadratic (bottom panel). The columns show the results for different age windows: 21–29, 22–28, 23–27 and 24–26. The control variables are the same reported in the first stage regressions. Standard errors in parenthesis.

Table C.3: Results for the probability of being employed in June 2017.

Window width	21–29	22–28	23–27	24–26
		Linear		
		OLS		
Treated	0.021	0.018	0.020	0.016
	(0.021)	(0.023)	(0.028)	(0.041)
		RF		
Age< 25	0.013	0.012	0.032	0.065 *
	(0.019)	(0.022)	(0.027)	(0.038)
		IV		
Treated	0.113	0.110	0.362	0.970
	(0.174)	(0.210)	(0.310)	(0.625)
F-stat	135	91.8	43.1	13.1
N.obs	9623	7491	5141	2593
		Quadratic		
		OLS		
Treated	0.021	0.018	0.019	0.014
	(0.021)	(0.023)	(0.028)	(0.041)
		RF		
Age< 25	0.013	0.012	0.035	0.066 *
	(0.019)	(0.022)	(0.027)	(0.038)
		IV		
Treated	0.117	0.112	0.386	0.970
	(0.174)	(0.211)	(0.308)	(0.612)
F-stat	135	91.5	43.8	13.6
N.obs	9623	7491	5141	2593

*, **, *** statistically significant at the 10%, 5% and 1% levels.

Note. This table shows the OLS, reduced form and IV results, using as outcome an indicator for being employed in June 2017. We use two different specifications for age: linear (top panel) and quadratic (bottom panel). The columns show the results for different age windows: 21–29, 22–28, 23–27 and 24–26. The control variables are the same reported in the first stage regressions. Standard errors in parenthesis.

Table C.4: Results for monthly income in June 2017.

Window width	21–29	22–28	23–27	24–26
		Linear		
		OLS		
Treated	-21.795	-37.039	-27.035	-19.765
	(21.071)	(23.643)	(28.803)	(41.850)
		RF		
Age< 25	6.613	6.876	11.318	63.362
	(19.857)	(22.626)	(27.975)	(39.312)
		IV		
Treated	59.511	65.191	127.512	947.100
	(178.588)	(214.366)	(315.200)	(642.768)
F-stat	135	91.8	43.1	13.1
N.obs	9623	7491	5141	2593
		Quadratic		
		OLS		
Treated	-21.798	-37.115	-27.909	-20.968
	(21.072)	(23.645)	(28.806)	(41.878)
		RF		
Age< 25	6.560	6.657	12.562	64.752 *
	(19.863)	(22.634)	(27.986)	(39.346)
		IV		
Treated	59.026	63.192	140.382	947.333
	(178.581)	(214.687)	(312.909)	(629.671)
F-stat	135	91.5	43.8	13.6
N.obs	9623	7491	5141	2593

*, **, *** statistically significant at the 10%, 5% and 1% levels.

Note. This table shows the OLS, reduced form and IV results, using as outcome the gross monthly income registered n June 2017. We use two different specifications for age: linear (top panel) and quadratic (bottom panel). The columns show the results for different age windows: 21–29, 22–28, 23–27 and 24–26. The control variables are the same reported in the first stage regressions. Standard errors in parenthesis.