

# Priority to Unemployed Immigrants? A Causal Machine Learning Evaluation of Training in Belgium

Bart Cockx<sup>§</sup>, Michael Lechner<sup>§§</sup>, Joost Bollens<sup>§§§</sup>

<sup>§</sup> Department of Economics, Ghent University

<sup>§§</sup> Swiss Institute for Empirical Economic Research (SEW), University of St. Gallen

<sup>§§§</sup> Vlaamse Dienst voor Arbeidsbemiddeling en Beroepsopleiding (VDAB)

This version: November 2022

*Date this version has been printed: 10 November 2022*

## Appendix for Online Publication

### **Addresses for correspondence:**

Bart Cockx, Ghent University, Department of Economics, Sint-Pietersplein 6, 9000 Gent, Belgium,  
[bart.cockx@ugent.be](mailto:bart.cockx@ugent.be), [sites.google.com/site/bartcockxsite](https://sites.google.com/site/bartcockxsite).

Michael Lechner, Swiss Institute for Empirical Economic Research (SEW), University of St. Gallen,  
Varnbuelstr. 14, 9000 St. Gallen Switzerland, [Michael.Lechner@unisg.ch](mailto:Michael.Lechner@unisg.ch), [www.michael-lechner.eu](http://www.michael-lechner.eu),  
[www.researchgate.net/profile/Michael\\_Lechner](https://www.researchgate.net/profile/Michael_Lechner).

Joost Bollens, VDAB, Keizerslaan 11, 1000 Brussels, Belgium, [joost.bollens@vdab.be](mailto:joost.bollens@vdab.be).

# Appendix A: Data

## A.1 Descriptive statistics by programme participation

*Table A.1: Means and standardized differences<sup>1</sup> for conditioning variables and outcomes*

	No ALMP participation (NOP) <sup>3</sup>	Short vocational training (SVT)	Long vocational training (LVT)	Orientation training (OT)
<i>A. Conditioning variables</i>				
(Pseudo-) Duration until the first treatment, in days (Daction)	92.66	124.39 (51.73)	122.81 (49.29)	119.17 (43.62)
Gender (female =1)(Woman)	.49	.31 (35.73)	.40 (16.48)	.47 (3.16)
Age in years (age)	35.09	34.01 (11.87)	34.00 (12.28)	33.63 (16.53)
Living in a city (city)	.36	.37 (2.43)	.32 (6.98)	.31 (11.02)
Knowledge of French (frans)	.63	.54 (17.87)	.72 (20.85)	.67 (9.07)
Knowledge of English (engels)	.70	.66 (7.86)	.82 (29.32)	.79 (20.30)
Knowledge of German (duits)	.25	.19 (13.20)	.34 (19.35)	.32 (16.31)
Knowledge of Italian (italiaans)	.03	.02 (6.85)	.02 (4.25)	.03 (3.92)
Knowledge of Spanish (spaans)	.08	.05 (12.83)	.07 (2.42)	.07 (4.02)
Proficiency in Dutch <sup>2</sup> (Lang_dutch)	2.44	2.51 (8.68)	2.72 (39.74)	2.64 (27.10)
Ever been in BIT before the current unemployment spell (cat_2)	.33	.36 (5.04)	.43 (20.62)	.40 (14.14)
Ever been in unemployment without U-benefit before the current unemployment spell (cat_3)	.24	.29 (9.64)	.24 (1.53)	.25 (2.39)
Ever been on welfare before the current unemployment spell (cat_5)	.07	.14 (22.27)	.04 (11.62)	.08 (4.27)
Having had an unemployment benefit sanction before the current unemployment spell (cat_14)	.06	.07 (6.63)	.04 (9.15)	.03 (12.71)
Ever have had a sickness benefit before the current unemployment spell (cat_76)	.12	.14 (4.52)	.10 (8.43)	.11 (4.86)
Ever been back to education before the current unemployment spell (cat_77)	.04	.04 (1.08)	.06 (10.59)	.06 (9.76)

Table A.1, continued

Ever been part-time working, part-time unemployed before the current unemployment spell (cat_80)	.19	.15 (9.68)	.14 (12.74)	.17 (3.20)
Ever been in BIT and part-time work before the current unemployment spell (cat_82)	.05	.03 (11.43)	.05 (2.37)	.06 (4.67)
Ever been in on-the-job-training before the current unemployment spell (cat_85)	.11	.16 (14.76)	.16 (16.18)	.17 (18.82)
Ever been in temporary agency work before the current unemployment spell (cat_89)	.21	.28 (16.80)	.27 (14.02)	.24 (7.47)
Ever been working full-time but looking for another job before the current unemployment spell (cat_90)	.29	.33 (8.76)	.42 (27.23)	.40 (23.20)
Ever been working part time + part time in education before the current unemployment spell (cat_91)	.01	.02 (8.68)	.01 (7.36)	.02 (2.10)
Ever been working part-time and part-time looking for a job before the current unemployment spell (cat_93)	.18	.14 (11.86)	.15 (6.84)	.18 (1.33)
Ever have had limited search obligations because of family or social reasons before the current unemployment spell (cat_96)	.02	.01 (.86)	.01 (7.73)	.02 (.29)
Ever have had limited search obligations because participation in training before the current unemployment spell (cat_97)	.03	.03 (.84)	.04 (4.67)	.04 (3.08)
Number of months in unemployment in the 10 years before the current spell (Unem_10jaar)	18.05	19.12 (5.44)	15.91 (11.14)	17.3 (3.80)
Number of months with sickness benefit in the 10 years before the current spell (ziek_10jaar)	1.05	.99 (1.09)	.72 (7.27)	1.03 (.32)
Number of months unknown position in the 10 years before the current spell (mystery_10jaar)	9.11	9.44 (1.53)	9.91 (3.55)	10.17 (4.76)
Number of months of work in the 10 years before the current spell (werk_10jaar)	48.69	54.18 (14.35)	59.48 (28.13)	54.89 (16.24)
Number of months in unemployment in the 5 years before the current spell (unem_5jaar)	10.52	11.24 (5.80)	8.91 (13.63)	9.81 (5.88)
Number of months with sickness benefit in the 5 years before the current spell (ziek_5jaar)	.66	.50 (5.11)	.48 (5.94)	.60 (1.79)

Number of months in unknown position in the 5 years before the current spell (mystery_5jaar)	2.89	3.32 (4.49)	3.00 (1.15)	3.39 (5.01)
--	------	----------------	----------------	----------------

Table A.1, continued

Number of months of work in the 5 years before the current spell (werk_5jaar)	31.81	35.30 (16.87)	38.64 (33.31)	36.3 (21.77)
Number of months in unemployment in the 2 years before the current spell (unem_2jaar)	3.89	3.77 (2.18)	2.97 (18.01)	3.18 (13.79)
Number of months with sickness benefit in the 2 years before the current spell (ziek_2jaar)	.22	.16 (4.39)	.19 (2.02)	.17 (3.24)
Number of months in unknown position in the 2 years before the current spell (mystery_2jaar)	.53	.64 (3.71)	.58 (1.50)	.67 (4.58)
Number of months of work in the 2 years before the current spell (werk_2jaar)	15.50	17.38 (22.57)	18.50 (36.82)	18.00 (30.32)
Having experience in the preferred occupation <sup>4</sup> (exp)	1.63	1.86 (21.21)	1.73 (8.73)	1.80 (14.90)
Duration of the previous job in months (duur_laatste_werk)	17.10	19.86 (20.66)	21.50 (32.67)	20.69 (26.87)
Number of professions the person is interested in (aant_beroepen)	2.77	3.04 (14.35)	2.98 (10.86)	3.01 (12.65)
Having participated in training before the current spell (vroeger_o)	.08	.12 (13.02)	.11 (10.77)	.11 (10.71)
Having participated in a training course Dutch Before the current spell (vroeger_n)	.01	.03 (10.78)	.01 (3.38)	.01 (.05)
Having participated in on-the-job-training before the current spell (vroeger_i)	.07	.10 (11.25)	.11 (13.29)	.12 (17.30)
Having participated in intensive counselling before the current spell (vroeger_t)	.06	.08 (5.04)	.05 (4.38)	.06 (.31)
Having participated in an orientation training before the current spell (vroeger_r)	.01	.01 (.09)	.03 (9.61)	.03 (10.54)
Having participated in another ALMP before the current spell (vroeger_a)	.04	.07 (13.53)	.05 (4.37)	.04 (3.04)
Drivers license car (rijbew_B)	.68	.63 (10.79)	.76 (16.39)	.71 (5.91)
Driver's license truck (rijbew_C)	.04	.05 (4.81)	.03 (4.98)	.03 (6.19)
Drivers license bus (rijbew_D)	.01	.01 (2.84)	.01 (.98)	.01 (2.46)
Educ. Attainment: High school (HS) drop-out (after 1 <sup>st</sup> level) (so1)	.14	.18 (10.29)	.04 (34.75)	.10 (13.63)

Table A.1, continued

Educ. Attainment: part-time HS degree, vocational track (dbso)	.04	.06 (9.51)	.03 (2.97)	.04 (1.73)
Educ. Attainment: HS drop-out (after 2 <sup>nd</sup> level HS), general or artistic tracks (akso2)	.02	.02 (1.72)	.02 (.84)	.03 (4.66)
Educ. Attainment: HS drop-out (after 2 <sup>nd</sup> level HS), vocational track (bso2)	.09	.13 (13.53)	.06 (9.47)	.08 (2.65)
Educ. Attainment: Special needs (for disabled) HS degree (buso)	.02	.04 (11.29)	.01 (5.18)	.02 (1.69)
Educ. Attainment: HS drop-out (after 2 <sup>nd</sup> level HS) technical track (tso2)	.04	.05 (4.11)	.04 (.02)	.05 (8.08)
Educ. Attainment: HS graduate, general or artistic track (akso3)	.08	.08 (1.39)	.13 (16.86)	.11 (10.42)
Educ. Attainment: HS graduate, vocational track (bso3)	.21	.26 (13.3)	.21 (1.41)	.22 (4.08)
Educ. Attainment: HS graduate, technical track (tso3)	.11	.11 (.27)	.22 (29.68)	.17 (17.82)
Educ. Attainment: Tertiary vocational education (hbo)	.02	.01 (8.48)	.02 (4.01)	.02 (.52)
Educ. Attainment: professional bachelor's degree (pba)	.13	.04 (31.38)	.14 (4.19)	.11 (4.89)
Educ. Attainment: academic bachelor's degree (aba)	.02	0.00 (11.01)	.01 (2.08)	.01 (1.02)
Educ. Attainment: academic master's degree (ma)	.10	.02 (32.94)	.06 (16.26)	.04 (24.03)
District of residence: Antwerpen (Antwerpen)	.23	.18 (11.78)	.17 (15.21)	.11 (31.53)
District of residence: Mechelen (Mechelen)	.05	.04 (5.04)	.04 (1.23)	.04 (2.37)
District of residence: Turnhout (Turnhout)	.06	.08 (6.27)	.09 (9.15)	.03 (13.70)
District of residence: Leuven (Leuven)	.06	.08 (4.83)	.07 (3.41)	.05 (4.95)
District of residence: Vilvoorde (Vilvoorde)	.09	.04 (16.56)	.04 (17.03)	.05 (14.31)
District of residence: Brugge (Brugge)	.04	.07 (13.63)	.05 (4.33)	.05 (4.78)
District of residence: Ieper or Diksmuide (Ieper)	.02	.03 (7.74)	.03 (6.16)	.03 (3.97)
District of residence: Kortrijk (Kortrijk)	.03	.04 (2.85)	.04 (3.76)	.05 (8.18)

Table A.1, continued

District of residence: Oostende (Oostende)	.02	.06 (16.33)	.04 (7.81)	.05 (13.80)
District of residence: Roeselare (Roeselare)	.02	.02 (2.97)	.03 (7.13)	.03 (7.55)
District of residence: Tielt or Veurne (Tielt)	.02	.02 (5.34)	.02 (3.06)	.02 (2.45)
District of residence: Aalst (Aalst)	.04	.04 (1.27)	.03 (9.14)	.04 (.88)
District of residence: Dendermonde (Dendermonde)	.03	.03 (3.77)	.03 (.54)	.03 (1.18)
District of residence: Eeklo (Eeklo)	.01	.01 (1.68)	.01 (4.94)	.01 (1.27)
District of residence: Gent (Gent)	.09	.06 (10.95)	.09 (.87)	.09 (1.06)
District of residence: Oudenaarde (Oudenaarde)	.02	.01 (9.38)	.00 (11.69)	.03 (6.14)
District of residence: Sint-Niklaas (Sintniklaas)	.04	.03 (3.83)	.04 (.50)	.02 (10.91)
District of residence: Hasselt (Hasselt)	.07	.09 (5.16)	.11 (14.26)	.11 (13.66)
District of residence: Maaseik (Maaseik)	.04	.05 (4.27)	.05 (7.98)	.08 (17.49)
District of residence: Tongeren (Tongeren)	.04	.03 (2.39)	.03 (1.59)	.09 (22.58)
Country of birth: Belgium (belg)	.65	.66 (2.09)	.77 (28.05)	.71 (13.01)
Country of birth: Western & Northern EU (eu_core)	.07	.05 (5.45)	.06 (1.42)	.06 (1.16)
Country of birth: Southern EU (eu_south)	.01	.01 (3.04)	.01 (6.33)	.02 (3.00)
Country of birth: Eastern EU (eu_east)	.06	.04 (10.12)	.03 (13.22)	.03 (12.13)
Country of birth: Turkey & Morocco (tm )	.05	.04 (8.27)	.02 (16.19)	.03 (9.71)
Country of birth: Rest of the world (row)	.16	.21 (11.68)	.10 (17.18)	.14 (4.68)
Calendar month start unemployment spell 1 (StartUI1)	.05	.04 (7.93)	.04 (7.04)	.06 (3.39)
Calendar month start unemployment spell 2 (StartUI2)	.08	.06 (5.41)	.09 (2.83)	.08 (1.55)
Calendar month start unemployment spell 3 (StartUI3)	.05	.06 (5.20)	.04 (.82)	.07 (9.22)

Table A.1, continued

Calendar month start unemployment spell 4 (StartUI4)	.05	.05 (.82)	.06 (4.67)	.06 (5.12)
Calendar month start unemployment spell 5 (StartUI5)	.05	.05 (.02)	.06 (4.79)	.04 (3.47)
Calendar month start unemployment spell 6 (StartUI6)	.04	.05 (4.70)	.04 (.93)	.04 (2.01)
Calendar month start unemployment spell 7 (StartUI7)	.05	.06 (4.15)	.06 (5.87)	.05 (1.17)
Calendar month start unemployment spell 8 (StartUI8)	.07	.06 (2.77)	.07 (.66)	.06 (2.38)
Calendar month start unemployment spell 9 (StartUI9)	.06	.05 (4.34)	.06 (1.1)	.05 (3.39)
Calendar month start unemployment spell 10 (StartUI10)	.07	.07 (1.17)	.07 (2.49)	.05 (7.78)
Calendar month start unemployment spell 11 (StartUI11)	.06	.07 (3.20)	.06 (1.67)	.06 (.42)
Calendar month start unemployment spell 12 (StartUI12)	.05	.05 (2.67)	.04 (7.03)	.05 (3.44)
Calendar month start unemployment spell 13 (StartUI13)	.05	.06 (4.77)	.04 (4.76)	.05 (2.23)
Calendar month start unemployment spell 14 (StartUI14)	.07	.07 (1.22)	.05 (7.33)	.05 (5.60)
Calendar month start unemployment spell 15 (StartUI15)	.05	.06 (5.45)	.04 (2.35)	.05 (.83)
Calendar month start unemployment spell 16) (StartUI16)	.04	.04 (.64)	.04 (.47)	.05 (3.98)
Calendar month start unemployment spell 17 (StartUI17)	.04	.03 (2.49)	.05 (6.41)	.04 (2.38)
Calendar month start unemployment spell 18 (StartUI18)	.04	.03 (2.16)	.05 (4.51)	.04 (1.79)
Calendar month start unemployment spell 19 (StartUI19)	.04	.04 (.09)	.05 (3.08)	.03 (4.86)
Preferred occupation 2 <sup>5</sup> (prof2)	.10	.04 (25.87)	.20 (29.63)	.20 (29.52)
Preferred occupation 19 <sup>5</sup> (prof19)	.05	.09 (16.09)	.05 (.39)	.05 (2.48)

Table A.1, continued

Preferred occupation 21 <sup>5</sup> (prof21)	.03	0.00 (20.17)	.02 (5.93)	.01 (14.05)
Preferred occupation 24 <sup>5</sup> (prof24)	.02	.01 (10.78)	.02 (5.95)	.02 (6.21)
Preferred occupation 25 <sup>5</sup> (prof25)	.05	.06 (3.85)	.05 (2.10)	.06 (5.01)
Preferred occupation 26 <sup>5</sup> (prof26)	.04	.05 (3.21)	.04 (3.93)	.04 (3.08)
Preferred occupation 27 <sup>5</sup> (prof27)	.02	.01 (3.29)	.01 (3.87)	.01 (9.65)
Preferred occupation 29 <sup>5</sup> (prof29)	.05	.07 (7.66)	.03 (13.38)	.03 (11.40)
Preferred occupation 30 <sup>5</sup> (prof30)	.01	.02 (3.46)	.01 (.25)	.01 (2.07)
Preferred occupation 31 <sup>5</sup> (prof31)	.07	.12 (18.21)	.06 (.68)	.06 (3.04)
Preferred occupation 32 <sup>5</sup> (prof32)	.04	.01 (15.57)	.06 (12.47)	.06 (11.31)
Preferred occupation 33 <sup>5</sup> (prof33)	.02	.01 (10.63)	.04 (10.59)	.04 (8.46)
Preferred occupation 34 <sup>5</sup> (prof34)	.02	.02 (1.41)	.01 (7.56)	.01 (3.33)
Preferred occupation 35 <sup>5</sup> (prof35)	.01	.02 (8.81)	.01 (2.30)	.02 (7.52)
Preferred occupation 36 <sup>5</sup> (prof36)	.01	.01 (4.18)	.01 (2.72)	.01 (2.55)
Preferred occupation 37 <sup>5</sup> (prof37)	.02	.03 (5.92)	.03 (5.33)	.03 (5.98)
Preferred occupation 38 <sup>5</sup> (prof38)	.01	.01 (.95)	.03 (10.76)	.02 (5.56)
Preferred occupation 39 <sup>5</sup> (prof39)	.01	.01 (.63)	.02 (7.66)	.01 (1.66)
Preferred occupation 40 <sup>5</sup> (prof40)	.05	.05 (2.33)	.03 (12.74)	.03 (8.74)
Preferred occupation 41 <sup>5</sup> (prof41)	.02	0.00 (16.53)	.02 (6.20)	.02 (.01)
Preferred occupation 42 <sup>5</sup> (prof42)	.18	.26 (19.08)	.18 (.44)	.18 (2.02)
Preferred occupation 43 <sup>5</sup> (prof43)	.07	.08 (4.95)	.03 (20.42)	.03 (16.08)
Preferred occupation 44 <sup>5</sup> (prof44)	.09	.02 (32.7)	.05 (15.81)	.03 (23.52)
The economic sector of previous job 0 <sup>6</sup> (sect0)	.30	.31 (.57)	.27 (7.40)	.29 (2.74)



Table A.1, continued

The economic sector of previous job 1 <sup>6</sup> (sect1)	.01	.02 (7.79)	.01 (1.12)	.01 (.65)
The economic sector of previous job 2 <sup>6</sup> (sect2)	.01	.01 (3.74)	.01 (.15)	.02 (.98)
The economic sector of previous job 3 <sup>6</sup> (sect3)	.03	.03 (1.09)	.02 (4.79)	.03 (1.52)
The economic sector of previous job 4 <sup>6</sup> (sect4)	.01	.02 (6.09)	.01 (2.33)	.01 (3.01)
The economic sector of previous job 5 <sup>6</sup> (sect5)	.03	.03 (1.31)	.02 (9.20)	.04 (4.84)
The economic sector of previous job 6 <sup>6</sup> (sect6)	.04	.01 (18.70)	.02 (12.32)	.01 (17.36)
The economic sector of previous job 7 <sup>6</sup> (sect7)	.01	.02 (3.28)	.01 (3.47)	.02 (1.41)
The economic sector of previous job 8 <sup>6</sup> (sect8)	.02	.02 (1.27)	.02 (2.02)	.02 (3.01)
The economic sector of previous job 9 <sup>6</sup> (sect9)	.01	.01 (2.03)	.01 (5.20)	.00 (9.56)
The economic sector of previous job 10 <sup>6</sup> (sect10)	.01	.01 (1.11)	.03 (8.62)	.02 (2.18)
The economic sector of previous job 11 <sup>6</sup> (sect11)	.02	.03 (7.94)	.02 (.79)	.02 (4.54)
The economic sector of previous job 12 <sup>6</sup> (sect12)	.01	.03 (11.83)	.02 (4.21)	.01 (2.80)
The economic sector of previous job 13 <sup>6</sup> (sect13)	.01	.01 (3.40)	.02 (1.62)	.01 (.16)
The economic sector of previous job 14 <sup>6</sup> (sect14)	.01	.01 (.51)	.01 (.55)	.02 (6.46)
The economic sector of previous job 15 <sup>6</sup> (sect15)	.02	.02 (1.01)	.02 (1.41)	.02 (1.13)
The economic sector of previous job 16 <sup>6</sup> (sect16)	.01	.01 (5.55)	.01 (2.44)	.01 (5.10)
The economic sector of previous job 17 <sup>6</sup> (sect17)	.01	0,00 (6.72)	0,00 (6.28)	0,00 (6.87)
The economic sector of previous job 18 <sup>6</sup> (sect18)	.02	.02 (1.33)	.03 (2.92)	.03 (5.99)
The economic sector of previous job 19 <sup>6</sup> (sect19)	.01	.01 (.92)	.02 (8.82)	.02 (3.83)

The economic sector of previous job 20 <sup>6</sup> (sect20)	.01	.02 (1.79)	.02 (6.13)	.01 (4.33)
The economic sector of previous job 21 <sup>6</sup> (sect21)	.03	.04 (4.90)	.04 (5.11)	.03 (1.08)
The economic sector of previous job 22 <sup>6</sup> (sect22)	.02	.02 (2.42)	.02 (1.91)	.03 (3.90)

Table A.1, continued

The economic sector of previous job 23 <sup>6</sup> (sect23)	.02	.03 (6.04)	.01 (3.87)	.01 (5.17)
The economic sector of previous job 24 <sup>6</sup> (sect24)	.02	.02 (5.74)	.03 (8.77)	.02 (3.64)
The economic sector of previous job 25 <sup>6</sup> (sect25)	.01	.01 (2.72)	.01 (6.91)	.01 (1.75)
The economic sector of previous job 26 <sup>6</sup> (sect26)	.01	.01 (1.82)	.01 (6.46)	.01 (4.73)
The economic sector of previous job 27 <sup>6</sup> (sect27)	.02	.01 (6.63)	.01 (9.14)	.01 (3.42)
The economic sector of previous job 28 <sup>6</sup> (sect28)	.01	.02 (6.81)	.02 (8.84)	.02 (5.50)
The economic sector of previous job 29 <sup>6</sup> (sect29)	.02	.04 (10.86)	.04 (10.77)	.03 (6.08)
The economic sector of previous job 30 <sup>6</sup> (sect30)	.01	.02 (4.86)	.01 (2.39)	.01 (1.08)
The economic sector of previous job 31 <sup>6</sup> (sect31)	.01	.02 (7.53)	.01 (.86)	.01 (.92)
The economic sector of previous job 32 <sup>6</sup> (sect32)	.02	.02 (1.29)	.01 (3.36)	.01 (7.06)
The economic sector of previous job 33 <sup>6</sup> (sect33)	.03	.01 (11.67)	.05 (7.34)	.04 (5.03)
The economic sector of previous job 34 <sup>6</sup> (sect34)	.04	.02 (12.25)	.05 (2.30)	.04 (.82)
The economic sector of previous job 35 <sup>6</sup> (sect35)	.06	.04 (9.90)	.06 (1.35)	.07 (5.45)

The economic sector of previous job 36 <sup>6</sup> (sect36)	.02	.02 (5.95)	.03 (2.73)	.02 (5.31)
Number of kids=0 (Nkids0)	.12	.17 (16.13)	.14 (7.96)	.17 (16.31)
Number of kids=1 (Nkids1)	.04	.04 (3.24)	.04 (3.22)	.04 (3.32)
Number of kids=2 (Nkids2)	.03	.03 (1.51)	.03 (2.39)	.03 (.91)
Number of kids > 3 (Nkids3)	.02	.01 (4.71)	.01 (7.42)	.01 (2.27)
Number of kids missing (Nkids4)	.80	.75 (10.19)	.79 (2.15)	.75 (11.15)
Household head: no (head0)	.45	.48 (5.31)	.49 (7.75)	.48 (5.00)
Household head: yes (head1)	.09	.08 (3.58)	.06 (11.06)	.07 (6.95)
Household head: unknown (head2)	.45	.44 (3.28)	.44 (1.79)	.45 (1.14)

Table A.1, continued

*B. Availability of labour market history data (LMHD)<sup>6</sup>*

Fraction of people in first unemployment spell (no LMHD)	0.14	0.07 (23)	0.05 (29)	0.06 (25)
Fraction of people with at least 2 years of LMHD	0.80	0.89 (24)	0.88 (24)	0.88 (23)
Fraction of people with at least 10 years of LMHD	0.41	0.44 (6)	0.46 (9)	0.46 (9)

*C. Outcomes*

Cumulative months in employment 9 months after programme start (Cwa9)	3.51	3.35 (4.83)	2.34 (36.91)	2.03 (46.35)
Cumulative months in unemployment 9 months after programme start (Cua9)	4.90	5.44 (16.16)	6.44 (49.18)	6.68 (56.21)
Cumulative months out of the labour force 9 months after programme start (Cia9)	.58	.21 (25.63)	.22 (25.93)	.29 (20.23)
Cumulative months in employment 30 months after programme start (Cwa30)	16.04	18.45 (24.06)	17.11 (10.83)	14.82 (11.92)
Cumulative months in unemployment 30 months after programme start (Cua30)	10.83	10.11 (8.42)	11.75 (10.66)	13.47 (29.89)
Cumulative months out of the labour force 30 months after programme start (Cia30)	3.13	1.44 (27.71)	1.14 (34.03)	1.71 (23.14)

Number of observations	59,964	1,305	1,220	1,115
------------------------	--------	-------	-------	-------

Notes: <sup>1</sup> The standardized difference is  $|\bar{x}^j - \bar{x}^{NOP}| / \sqrt{[\text{Var}(x^j) + \text{Var}(x^{NOP})]/2}$ , where  $\bar{x}^j$  and  $\text{Var}(x^j)$  are the sample mean and variance of the variable  $x^j$  for  $j \in \{SVT, LVT, OT\}$ .

<sup>2</sup> Proficiency in Dutch = 0 if no knowledge; = 1 if limited; = 2 if good; =3 if very good.

<sup>3</sup> For non-participants in ALMP (NOP) the date at which the ALMP starts is predicted (See Section 4.4).

<sup>4</sup> Experience: no experience (0), limited experience (1), good experience (2), a lot of experience (3)

<sup>5</sup> 2 = General clerk; 19 = Goods handlers; 21 = Managers of a department or service; 24 = Educators; 25 = Sales support staff; 26 = Sales representatives; 27 = Representatives; 29 = Hall staff; 30 = Food workers; 31 = Construction workers and technicians; 32 = Specialised administrative staff; 33 = Computer and ICT staff; 34 = Agricultural, horticultural and forestry workers and fishermen; 35 = Vehicle mechanics; 36 = Staff involved in tourism, leisure and sport; 37 = Metalworkers; 38 = Draughtsmen and designers; 39 = Transport and logistics personnel; 40 = Nurses and carers, Law enforcement and rescue workers, Medics, Paramedics and laboratory assistants; 41 = Private sector consultants, Bank and insurance experts, Business consultants; 42 = Craftsmen, Drivers, Apparel and leatherworkers, Miscellaneous production workers, Printers, Electricians and electricians; Woodworkers, Industrial technicians, Machinists and crane operators; Metal production workers, Operators chemistry and plastics, Precision technicians, Technical managers, Textile workers, Rail, water and air transport workers; 43 = Personal service providers, Cleaning and maintenance personnel; 44 = Architects and surveyors, Artists, artists and other cultural occupations, Controllers and inspectors, Instruction, training and education personnel, Media personnel, Teaching and management personnel in schools, Researchers and experts study service, Socio-cultural workers; <sup>6</sup> 0 = Missing; 1 = Construction of houses; 2 = Retail sale of clothing; 3 = Dining facility full service; 4 = Dining facility limited service; 5 = General cleaning of buildings; 6 = Ordinary general secondary education; 7 = Other social work; 8 = Retail sale of food; 9 = Institutions for the elderly and disabled; 10 = Electrical installation, plumbing and other construction installation; 11 = Finishing of buildings; 12 = Other specialised construction activities; 13 = Wholesale trade in consumer goods; 14 = Retail trade in consumer goods; 15 = Retail trade in other goods; 16 = Other social work activities without accommodation; 17 = Other personal services; 18 = Manufacture of food products; 19 = Manufacture of fabricated metal products, except machinery and equipment; 20 = Wholesale and retail trade, maintenance and repair of motor vehicles and motorcycles; 21 = Wholesale trade and commission trade, except of motor vehicles and motorcycles; 22 = Retail trade, except of motor vehicles and motorcycles; 23 = Land transport and transport via pipelines; 24 = Warehousing and support activities for transportation; 25 = Food and beverage service activities; 26 = Services to buildings and landscape gardening; 27 = Education; 28 = Industry 1; 29 = Industry 2; 30 = Industry 3, Electricity, gas, steam and air conditioning supply, Water supply; Waste and sewerage management and remediation services; 31 = Construction, wholesale and retail trade; Repair of motor vehicles and motorcycles; 32 = Transport and storage, Accommodation and food service activities; 33 = Information and communication, Financial and insurance activities, Real estate activities, Professional, scientific and technical activities; 34 = Administrative and support service activities 1; 35 = Administrative and support service activities 2, Public administration and defence; compulsory social security, Human health and social work activities; 36 = Arts, entertainment and recreation, Other services.

<sup>6</sup> This information is implicitly conditioned upon, as it can be obtained by (a combination of) values of conditioning variables: no LHHD corresponds to individuals for which  $duur\_laatste\_werk=0$ , individuals for which  $unem\_2jaar + ziek\_2jaar + mystery\_2jaar + werk\_2jaar = 24$  have at least 2 years of LMHD, and individuals for which  $unem\_10jaar + ziek\_10jaar + mystery\_10jaar + werk\_10jaar = 120$  have at least 10 years of LMHD.

## A.2 Analysis of selection between training programmes

*Table A.2: Means and standardized differences (SD)<sup>1</sup> of conditioning variables for which the maximum SD of LVT and OT relative to SVT is larger than 20%*

	Short vocational training (SVT)	Long vocational training (LVT)	Orientation training (OT)
<i>Conditioning variables</i>			
Gender (female =1)(Woman)	.31	.40 (18.99)	.47 (32.48)
Knowledge of French (frans)	.54	.72 (39.04)	.67 (27.04)
Knowledge of English (engels)	.66	.82 (37.30)	.79 (28.23)
Knowledge of German (duits)	.19	.34 (32.65)	.32 (29.59)
Proficiency in Dutch <sup>2</sup> (Lang_dutch)	2.51	2.72 (34.50)	2.64 (20.34)
Ever been on welfare before current unemployment spell (cat_5)	.14	.04 (33.35)	.08 (18.08)
Drivers license car (rijbew_B)	.63	.76 (27.27)	.71 (16.72)
Educ. Attainment: High school (HS) drop-out (after 1 <sup>st</sup> level) (so1)	.18	.04 (44.6)	.1 (23.86)
Educ. Attainment: HS drop-out (after 2 <sup>nd</sup> level HS) , vocational track (bso2)	.13	.06 (22.85)	.08 (16.15)
Educ. Attainment: HS graduate, technical track (tso3)	.11	.22 (29.94)	.17 (18.09)
Educ. Attainment: professional bachelor's degree (pba)	.04	.14 (35.36)	.11 (26.70)
District of residence: Tongeren (Tongeren)	.03	.03 (0.80)	.09 (24.80)
Country of birth: Belgium (belg)	.66	.77 (25.92)	.71 (10.91)
Country of birth: Rest of the world (row)	.21	.10 (28.8)	.14 (16.35)
Preferred occupation 2 <sup>3</sup> (prof2)	.04	.20 (54.05)	.20 (53.94)
Preferred occupation 29 <sup>3</sup> (prof29)	.07	.03 (20.74)	.03 (18.83)
Preferred occupation 31 <sup>3</sup> (prof31)	.12	.06 (18.88)	.06 (21.17)
Preferred occupation 32 <sup>3</sup> (prof32)	.01	.06 (26.99)	.06 (25.94)

Table A.2, continued

Preferred occupation 33 <sup>3</sup> (prof33)	.01	.04 (20.42)	.04 (18.47)
Preferred occupation 43 <sup>3</sup> (prof43)	.08	.03 (25.09)	.03 (20.86)
Number of observations	1,305	1,220	1,115

Notes: <sup>1</sup> The standardized difference is  $|\bar{x}^j - \bar{x}^{SVT}| / \sqrt{[Var(x^j) + Var(x^{SVT})]/2}$ , where  $\bar{x}^j$  and  $Var(x^j)$  are the sample mean and variance of the variable  $x^j$  for  $j \in \{LVT, OT\}$ .

<sup>2</sup> Proficiency in Dutch = 0 if no knowledge; = 1 if limited; = 2 if good; =3 if very good.

<sup>3</sup> 2 = General clerk; 29 = Hall staff; 31 = Construction workers and technicians; 32 = Specialised administrative staff; 33 = Computer and ICT staff; 43 = Personal service providers, Cleaning and maintenance personnel.

Table A.2 reports the means and SDs of conditioning variables for which the maximum SD of LVT and OT relative to SVT is larger than 20%. By considering only variables for which these SDs are larger than 20%, we can study in which direction there is an imbalance of participants in SVT relative to the other two programs. We first summarize for each set of variables the direction of the selection bias of SVT relative to LVT and then list the associated SD for each variable (we only retain variables for which the corresponding SD is larger than 20%) and between parentheses first the mean of participants in SVT versus (vs.) the mean of participants in LVT. Next, after a brief conclusion regarding what we can learn from this analysis, a brief comparison is made between the participants in LVT and OT.

- 1) With respect to education, we find that participants in SVT have a much lower level of educational attainment than those in LVT: (i) high school drop-out (level 1): SD=44.6 (.18 vs .04); (ii) HS drop-out – vocational track (level 2): SD=22.85 (.13 vs .06); (iii) HS graduate technical track: SD=29.94 (.11 vs .22); (iv) professional bachelor's: SD 35.36 (.04 vs .14).
- 2) Fewer Belgians and more individuals originating from the rest of the world are participating in SVT than in LVT: (i) Belgian: SD=25.92 (.66 vs .77); (ii) ROW: SD=28.8 (.21 vs .10).
- 3) The Dutch proficiency is lower for participants in SVT than in LVT: SD=34.5 (2.51 vs 2.72).

- 4) The proficiency in (locally useful) foreign languages is lower for participants in SVT than in LVT: (i) French: SD=39.04 (.54 vs .72); English: SD=37.3 (.66 vs .82); German: SD=32.65 (.19 vs .34).
- 5) A higher fraction of participants in SVT has ever been on welfare than those in LVT (cat\_5): SD=33.35 (.14 vs .04).
- 6) Fewer participants in SVT have a driver's licence for a car than those in LVT: SD=27.27 (.63 vs .76).
- 7) Some previous occupations for which the labour market is typically tight are less represented among participants in SVT than in LVT: (i) general clerk (prof2): SD=54.05 (.04 vs .20); Specialized administrative staff (prof32): SD=26.99 (.01 vs .06); Computer and ICT staff (prof33): SD=20.42 (.01 vs .04).
- 8) Other occupations, for which applicants typically exceed candidates, are more represented among participants in SVT than in LVT: (i) hall staff (prof29): SD=20.74 (.07 vs .03); (ii) Personal service providers, Cleaning and maintenance personnel (prof43): SD=25.09 (.08 vs .03).

We conclude that participants in SVT generally have characteristics that are less valuable in the labour market than participants in LVT. This means that they are *negatively* selected in terms of observables relative to participants in LVT.

We can also deduce from Table A.2 that participants in OT have a very similar (or maybe slightly worse) profile as those in LVT, which is, hence, systematically better than participants in SVT.

## Appendix B: Additional details of estimation

### B.1 Predictions of the pseudo durations for the NOP

For all participants in training log duration until the programme start is regressed on all the explanatory variables, including the interactions of all the explanatory variables with gender. This was done with a LASSO regression, using the R-package `glmnet` (Friedman et al., 2008). A ten-fold cross-validation approach is used to determine the penalty term. As the LASSO estimates are biased, the regression is re-run using the subset of variables selected by the LASSO, to obtain OLS estimates (Post-LASSO). Subsequently, the OLS results are used to predict the duration until the start in the NOP sample. Then we add to this prediction a draw from a Normal distribution with mean zero and the standard deviation being the one of the OLS regression.

The initial NOP sample has 112,128 observations. For 54,255 observations, the duration of the unemployment spell is smaller than the predicted duration until the start. For 6,732 observations, the predicted duration until the start exceeds 9 months. For 5,371 observations, both problems are present. As a result, there are 55,616 observations where either or both problems apply. These observations are dropped from the sample.

The LASSO regression is not reported while the Post-LASSO regression is reported in Table B.1.



*Table B.1: Post LASSO regression of the log duration until the programme start on the explanatory variables in a LASSO regression*

Intercept	4,96	(0,16)	***
Number of months in unemployment in the 10 years before current spell (Unem_10jaar)	0,00	(0,00)	***
Number of months unknown position in the 10 years before current spell (mystery_10jaar)	0,00	(0,00)	*
Number of months out of the labour force in the 10 years before current spell (olf_10jaar)	0,00	(0,00)	
Number of months of work in the 5 years before current spell (werk_5jaar)	0,00	(0,00)	*
Number of months in unemployment in the 2 years before current spell (unem_2jaar)	-0,01	(0,00)	*
Number of months unknown position in the 2 years before current spell (mystery_2jaar)	0,00	(0,00)	
Number of months out of the labour force in the 2 years before current spell (olf_2jaar)	0,00	(0,01)	
Drivers license car (rijbew_B)	-0,02	(0,03)	
Drivers license truck (rijbew_C)	-0,19	(0,06)	***
Educ. attainment: part time education, professional track (dbso)	0,03	(0,05)	
Educ. attainment; 3rd level secondary education, general track (aso3)	0,03	(0,03)	
Educ. attainment; 3rd level secondary education, artistic track (kso3)	0,08	(0,08)	
Educ. attainment: master (ma)	0,05	(0,07)	
Age in years (age)	-0,01	(0,03)	
Knowledge of english (engels)	-0,03	(0,02)	
Knowledge of german (duits)	0,02	(0,03)	
Proficiency in dutch very good (ned3)	0,06	(0,02)	**
District of residence: Antwerpen (Antwerpen)	-0,13	(0,08)	*
District of residence Mechelen(Mechelen)	0,08	(0,05)	
District of residence Turnhout (Turnhout)	0,07	(0,05)	
District of residence Leuven (Leuven)	0,13	(0,06)	**
District of residence Vilvoorde (Vilvoorde)	0,14	(0,05)	***
District of residence Brugge (Brugge)	-0,22	(0,07)	***
District of residence Ieper (Ieper)	-0,15	(0,07)	**
District of residence Kortrijk (Kortrijk)	0,06	(0,06)	
District of residence Dendermonde (Dendermonde)	-0,13	(0,08)	
District of residence Gent (Gent)	0,14	(0,05)	***
District of residence Oudenaarde (Oudenaarde)	-0,04	(0,11)	
District of residence Hasselt (Hasselt)	-0,03	(0,05)	
District of residence Maaseik (Maaseik)	0,09	(0,07)	
District of residence Tongeren (Tongeren)	-0,03	(0,06)	

Table B.1, continued

Household-head	-0,03	(0,01)	**
Number of occupations the person is interested in (aant_beroepen)	0,00	(0,01)	
Ever been in BIT before the current unemployment spell (cat_2)	-0,02	(0,03)	
Ever been in unemployment without UB before the current unemployment spell (cat_3)	-0,01	(0,02)	
Ever been in PWA before the current unemployment spell (cat_30)	0,15	(0,25)	
Ever been in a trajectory from sickness benefit to work before the current unemployment spell (cat_32)	-0,15	(0,10)	
Ever been in Arbeidszorg before the current unemployment spell (cat_33)	-1,28	(0,47)	***
Ever been in BIT and part time work before the current unemployment spell (cat_82)	-0,02	(0,04)	
Ever been in temporary agency work before the current unemployment spell (cat_89)	-0,11	(0,03)	***
Ever been working full-time but looking for another job before the current unemployment spell (cat_90)	-0,08	(0,02)	***
Ever been working part-time + part-time in education before the current unemployment spell (cat_91)	-0,14	(0,10)	
Ever been in the specific status given to some high skilled unemployed from outside the European Economic Area, before the current unemployment spell (cat_94)	0,10	(0,16)	
Ever have had limited search obligations because of family or social reasons before the current unemployment spell (cat_96)	0,06	(0,36)	
Ever have had limited search obligations because participation in training before the current unemployment spell (cat_97)	-0,04	(0,05)	
Unemployment rate in district of residence (Wlgr)	-2,42	(1,59)	
Educational attainment high (High)	0,02	(0,05)	
Age between >=22, < 25 (age_lt25)	-0,05	(0,03)	
Age between >=36, < 50 (age_lt50)	0,06	(0,05)	
Age between >=50 and <=55 (age_ge50)	0,13	(0,08)	
Country of birth: Belgium (belg)	-0,03	(0,03)	
Country of birth: Western & Northern EU (eu_core)	0,04	(0,06)	
Unemployment spell began in January (m1)	-0,12	(0,06)	**
Unemployment spell began in February (m2)	-0,22	(0,06)	***
Unemployment spell began in March (m3)	0,02	(0,03)	
Unemployment spell began in June (m6)	0,06	(0,05)	
Unemployment spell began in July (m7)	-0,05	(0,06)	
Unemployment spell began in August (m8)	-0,36	(0,06)	***
Unemployment spell began in September (m9)	-0,27	(0,05)	***
Unemployment spell began in October (m10)	-0,08	(0,05)	
Unemployment spell began in November (m11)	-0,16	(0,04)	***
Unemployment spell began in December (m12)	-0,16	(0,04)	***
Duration previous job in months (duur_laatste_werk)	0,00	(0,00)	**
Having participated in on-the-job-training before the current spell (vroeger_i)	-0,02	(0,03)	

Table B.1, continued

Having participated in a Dutch training course before the current spell (vroeger_n)	-0,02	(0,08)	
Having participated in a training before the current spell (vroeger_o)	-0,04	(0,04)	
Having participated in an orientation training before the current spell (vroeger_r)	-0,10	(0,07)	
Having participated in intensive counselling before the current spell (vroeger_t)	0,03	(0,04)	
Having limited experience in the preferred occupation (ervaring_beperkt)	-0,02	(0,03)	
Having no experience in the preferred occupation (ervaring_geen)	0,00	(0,04)	
Knowledge of another language (different from Dutch, English, German, Italian or Spanish) (andere_taal)	0,01	(0,04)	
District or residence Aalst, interaction with sex	-0,05	(0,07)	
District or residence Antwerpen, interaction with sex	0,13	(0,06)	**
District or residence Brugge, interaction with sex	0,22	(0,09)	**
District or residence Dendermonde, interaction with sex	0,17	(0,11)	
District or residence Eeklo, interaction with sex	-0,10	(0,11)	
District or residence Gent, interaction with sex	-0,08	(0,06)	
District or residence Leuven, interaction with sex	-0,20	(0,08)	***
District or residence Maaseik, interaction with sex	-0,13	(0,09)	
District or residence Oostende, interaction with sex	0,06	(0,10)	
District or residence Oudenaarde, interaction with sex	-0,27	(0,15)	*
District or residence Roeselare, interaction with sex	-0,10	(0,10)	
District or residence Sint-Niklaas, interaction with sex	0,08	(0,08)	
District or residence Turnhout, interaction with sex	0,07	(0,07)	
Educ. Attainment: academic bachelor, interaction with sex	-0,13	(0,15)	
Age between >=50 and <=55, interaction with sex	0,05	(0,08)	
Age between >=36, < 50, interaction with sex	-0,10	(0,04)	**
Knowledge of another language (different from Dutch, English, German, Italian or Spanish, interaction with sex	0,01	(0,05)	
Educ. Attainment; 2nd level secondary education, general track, interaction with sex	-0,17	(0,12)	
Educ. Attainment; 3rd level secondary education, professional track, interaction with sex	-0,03	(0,04)	
Having had an UB sanction before the current unemployment spell, interaction with sex	0,00	(0,07)	
Ever have had a sickness benefit before the current unemployment spell, interaction with sex	-0,09	(0,04)	**
Ever been back to education before the current unemployment spell, interaction with sex	-0,01	(0,06)	
Ever been part-time working, part-time unemployed before the current unemployment spell, interaction with sex	0,03	(0,03)	
Ever been in on-the-job-training before the current unemployment spell, interaction with sex	-0,03	(0,05)	

Table B.1, continued

Ever been in temporary agency work before the current unemployment spell, interaction with sex	0,10	(0,04)	**
Ever been working part-time + part-time in education before the current unemployment spell, interaction with sex	0,14	(0,15)	
Ever been in temporary unemployment before the current unemployment spell, interaction with sex	0,28	(0,24)	
Ever been working part-time and part-time looking for a job before the current unemployment spell, interaction with sex	-0,06	(0,03)	**
Ever have had limited search obligations because of family or social reasons before the current unemployment spell, interaction with sex	0,08	(0,37)	
Living in a city, interaction with sex	0,03	(0,03)	
Educ. attainment: part-time education, professional track, interaction with sex	0,05	(0,09)	
Knowledge of German, interaction with sex	0,09	(0,05)	**
No experience in preferred occupation, interaction with sex	0,13	(0,05)	**
Good experience in preferred occupation, interaction with sex	0,04	(0,03)	
Country of birth: Western & Northern EU, interaction with sex	0,06	(0,08)	
Country of birth: Southern EU, interaction with sex	0,18	(0,13)	
Knowledge of French, interaction with sex	0,07	(0,03)	**
Educ. attainment: higher professional education , interaction with sex	-0,10	(0,12)	
Knowledge of Italian , interaction with sex	-0,12	(0,09)	
Educational attainment, secondary, 3rd level, interaction with sex	-0,03	(0,06)	
Educational attainment is low , interaction with sex	0,02	(0,06)	
Unemployment spell began in January, interaction with sex	-0,29	(0,08)	***
Unemployment spell began in February, interaction with sex	0,19	(0,09)	**
Unemployment spell began in April, interaction with sex	0,01	(0,05)	
Unemployment spell began in June, interaction with sex	-0,15	(0,06)	**
Unemployment spell began in July, interaction with sex	-0,14	(0,07)	***
Unemployment spell began in August, interaction with sex	-0,31	(0,08)	***
Unemployment spell began in September, interaction with sex	0,07	(0,07)	
Unemployment spell began in October, interaction with sex	-0,07	(0,08)	
Unemployment spell began in November, interaction with sex	-0,14	(0,06)	**
Number of months in unknown position in the 5 years before the current spell, interaction with sex	0,00	(0,00)	**
No knowledge of Dutch, interaction with sex	0,19	(0,18)	
Number of months out of the labour market in the 10 years before the current spell, interaction with sex	0,00	(0,00)	
Number of months out of the labour market in the 5 years before the current spell, interaction with sex	0,00	(0,00)	
Educ. Attainment: professional bachelor, interaction with sex	0,03	(0,08)	
Driver license car, interaction with sex	-0,06	(0,04)	

Table B.1, continued

Driver license truck, interaction with sex	-0,11	(0,22)	
Driver license bus, interaction with sex	-0,16	(0,20)	
Knowledge of Spanish, interaction with sex	0,04	(0,05)	
Country of birth: Turkey & Morocco, interaction with sex	0,12	(0,07)	*
Number of months in unemployment in the 10 years before the current spell, interaction with sex	0,00	(0,00)	
Number of months in unemployment in the 2 years before the current spell, interaction with sex	0,01	(0,00)	
Number of months of work in the 10 years before the current spell, interaction with sex	0,00	(0,00)	
Number of months with sickness benefit in the 5 years before the current spell, interaction with sex	0,00	(0,00)	
Having participated in a Dutch training course before the current spell, interaction with sex	-0,22	(0,13)	*
Having participated in a training before the current spell, interaction with sex	-0,10	(0,06)	*

Note: Standard errors are in brackets. \*, \*\*, \*\*\* indicate the precision of the estimate by showing whether the p-value of a two-sided significance test is below 10%, 5%, and 1% respectively.

## B.2 Tuning parameters of the MCF

The estimations have been performed with the Python package MCF, version 0.2.0, which can be freely downloaded from PyPI via *pip install*. Here, we will give some key tuning parameters used.

*Variables for leaf splitting:* The number of randomly selected variables are used in each leaf splitting was drawn from a Poisson distribution (for each split) with means 9 and 62. For the main employment outcome, a mean of 62 led to the smallest value of the objective function (which is to be minimized) in the out-of-bag samples.

The *minimum leaf size* was set to 5.

The number of *subsampling replications* was set to 1000 and *43% of the observations used to build the forest are contained in each subsample*. 43%-subsampling was also used when populating the leaves with outcomes.

### B.3 Tuning parameters of the Post-LASSO for the IATEs

The penalty term of the LASSO was determined by a grid of 100 different values, starting with the model without covariates to the model with all covariates. On this grid, the optimal value was determined by ten-fold cross-validation using the mean squared prediction error of the Post-LASSO.

### B.4: Algorithm used for computing policy trees

We modify Algorithm 2 of Zhou, Athey, and Wager (2019) in three ways. First, instead of using one global approximation level for computing possible splitting rules ( $A$  in their notation), we use a finer grid at higher levels of the tree (1<sup>st</sup> level, i.e., at the bottom of the tree:  $A$ , 2<sup>nd</sup> level:  $A/2$ , 3<sup>rd</sup> level,  $A/4$ , 4<sup>th</sup> level, i.e., at the top of the tree  $A/8$ ). Second, we (implicitly) enforce constraints to maximum individual programme shares as well as overall programme shares. Third, we allow explicitly for unordered categorical variables.

The algorithm we implemented can be written in three parts. The augmented baseline algorithm of Zhou, Athey, and Wager (2019) is denoted by Algorithm 1. Algorithm 2 is the module enforcing constraints, while Algorithm 3 deals with categorical variables. While in the presentation of Algorithm 1, we closely follow Zhou, Athey, and Wager (2019), we also keep the description at a stylised level and abstract from many implementational details.

A few remarks are in order about the elements that differ from the original proposal that essentially all address some approximations that reduce the computational costs of finding ‘good’ trees that are close to the optimal ones. For Algorithm 1 reducing  $A$  at higher levels leads to additional precision at reasonable additional computational costs. This precision gain is more relevant when fewer splitting points are available as in the higher levels.

Second, the constraints are enforced by finding treatment-specific cost values (in terms of outcome). These costs are found in (fast) iterative procedures such that an unconstrained

Black-box treatment will closely obey the restrictions if such costs are subtracted from policy scores. Here, it was only necessary to have such costs for SVT. Once these costs are found, they are subtracted from the treatment-specific policy scores. These modified scores are subsequently used to build the policy trees.

Finally, to deal with categorical variables, we allow for all possible splits. While this is computationally intensive, it should give much better results than for example breaking-up such variables into dummy variables as is common in the econometric literature.

## Appendix C: Additional results

### C.1 Average treatment effect on the treated

Table C.1 shows how the effects differ across the populations of programme participants for all average population effects. Numbers in bold always relate to the effect of the programme on its own population of participants (ATET). If caseworkers maximise effects, then one should expect the bold number to be the largest entry in each row.



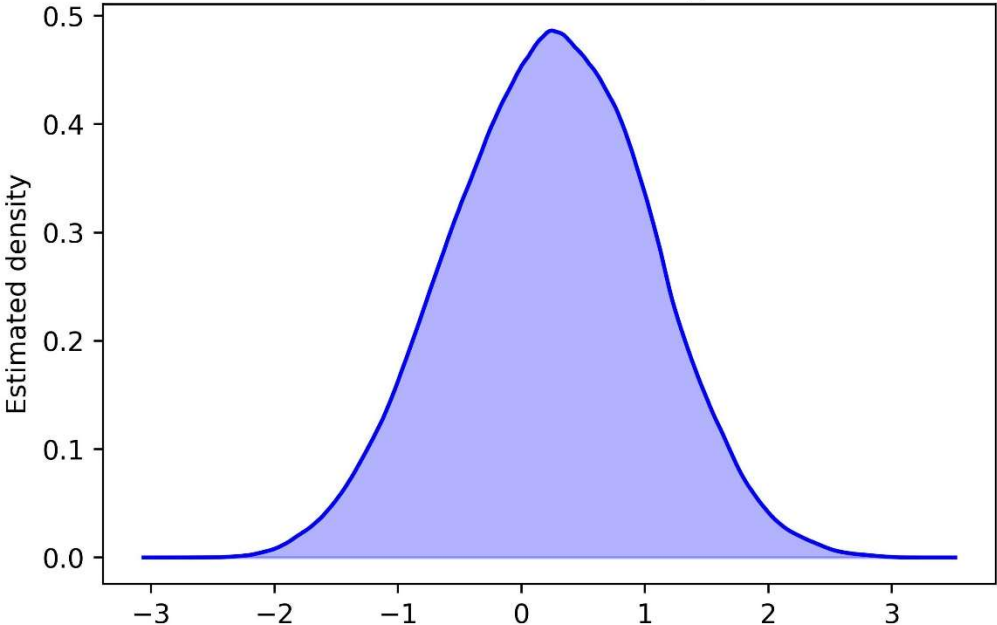
*Table C.1: Comparison of the effects of the different programmes on cumulative months in employment, unemployment and out of the labour force for different populations by participation status*

	No ALMP participation (NOP)	Short vocational training (SVT)	Long vocational training (LVT)	Orientation training (OT)
<i>Cumulative months in employment 9 months after programme start</i>				
SVT – NOP	-0.1 (0.1)	<b>-0.1</b> (0.1)	-0.1 (0.1)	-0.1 (0.1)
LVT – NOP	-1.3 (0.1) ***	-1.2 (0.1) ***	<b>-1.2</b> (0.1) ***	-1.2 (0.1) ***
OT – NOP	-1.5 (0.1) ***	-1.5 (0.1) ***	-1.5 (0.1) ***	<b>-1.5</b> (0.1) ***
LVT – SVT	-1.2 (0.2) ***	-1.1 (0.2) ***	<b>-1.1</b> (0.2) ***	-1.1 (0.2) ***
OT – SVT	-1.5 (0.2) ***	-1.4 (0.2) ***	-1.3 (0.2) ***	<b>-1.4</b> (0.2) ***
OT – LVT	-0.3 (0.2)	-0.3 (0.2)	-0.3 (0.2) *	<b>-0.3</b> (0.2)
<i>Cumulative months in employment between month 22 and month 30 after programme start</i>				
SVT – NOP	1.3 (0.1) ***	<b>1.3</b> (0.1) ***	1.2 (0.1) ***	1.3 (0.1) ***
LVT – NOP	0.9 (0.2) ***	0.9 (0.2) ***	<b>0.9</b> (0.1) ***	0.9 (0.1) ***
OT – NOP	0.7 (0.2) ***	0.6 (0.2) ***	0.6 (0.2) ***	<b>0.6</b> (0.1) ***
LVT – SVT	-0.3 (0.2) *	-0.4 (0.2) *	<b>-0.3</b> (0.2)	-0.3 (0.2)
OT – SVT	-0.6 (0.2) ***	-0.6 (0.2) ***	-0.6 (0.2) ***	<b>-0.7</b> (0.2) ***
OT – LVT	-0.3 (0.2)	-0.3 (0.2)	-0.4 (0.2) *	<b>-0.3</b> (0.2)
<i>Cumulative months in employment 30 months after programme start</i>				
SVT – NOP	2.5 (0.4) ***	<b>2.6</b> (0.4) ***	2.4 (0.4) ***	2.4 (0.4) ***
LVT – NOP	0.2 (0.4)	0.2 (0.4)	<b>0.3</b> (0.4)	0.3 (0.4)
OT – NOP	-1.4 (0.4) ***	-1.3 (0.4) ***	-1.5 (0.4) ***	<b>-1.4</b> (0.4) ***
LVT – SVT	-2.3 (0.5) ***	-2.4 (0.5) ***	<b>-2.1</b> (0.5) ***	-2.2 (0.5) ***
OT – SVT	-3.9 (0.6) ***	-3.9 (0.6) ***	-3.9 (0.6) ***	<b>-3.8</b> (0.6) ***
OT – LVT	-1.6 (0.6) ***	-1.5 (0.6) ***	-1.8 (0.6) ***	<b>-1.6</b> (0.6) ***
<i>Cumulative months in unemployment 30 months after programme start</i>				
SVT – NOP	-1.2 (0.3) ***	<b>-1.2</b> (0.3) ***	-1.0 (0.3) ***	-1.1 (0.3) ***
LVT – NOP	1.4 (0.4) ***	1.4 (0.4) ***	<b>1.3</b> (0.4) ***	1.4 (0.4) ***
OT – NOP	2.9 (0.4) ***	2.8 (0.4) ***	2.9 (0.4) ***	<b>2.8</b> (0.4) ***
LVT – SVT	2.6 (0.5) ***	2.6 (0.5) ***	<b>2.4</b> (0.5) ***	2.5 (0.5) ***
OT – SVT	4.0 (0.5) ***	3.9 (0.5) ***	3.9 (0.5) ***	<b>3.9</b> (0.5) ***
OT – LVT	1.5 (0.5) ***	1.3 (0.5) ***	1.5 (0.5) ***	<b>1.4</b> (0.5) ***
<i>Cumulative months out-of-the-labour force 30 months after programme start</i>				
SVT – NOP	-1.5 (0.2) ***	<b>-1.4</b> (0.2) ***	-1.4 (0.2) ***	-1.4 (0.2) ***
LVT – NOP	-1.9 (0.2) ***	-1.9 (0.2) ***	<b>-1.9</b> (0.2) ***	-1.9 (0.2) ***
OT – NOP	-1.3 (0.2) ***	-1.3 (0.2) ***	-1.2 (0.2) ***	<b>-1.3</b> (0.2) ***
LVT – SVT	-0.5 (0.3) *	-0.4 (0.3)	<b>-0.4</b> (0.3)	-0.5 (0.3) *
OT – SVT	-0.2 (0.3)	0.1 (0.3)	0.2 (0.3)	<b>0.2</b> (0.3)
OT – LVT	0.6 (0.3)	0.6 (0.3) *	0.7 (0.3) **	<b>0.6</b> (0.3) **

Note: Outcomes measured in months. All effects are population averages for the respective programme participants given in the column. Standard errors are in brackets. \*, \*\*, \*\*\* the precision of the estimate by indicating whether the p-value of a two-sided significance test is below 10%, 5%, and 1% respectively.

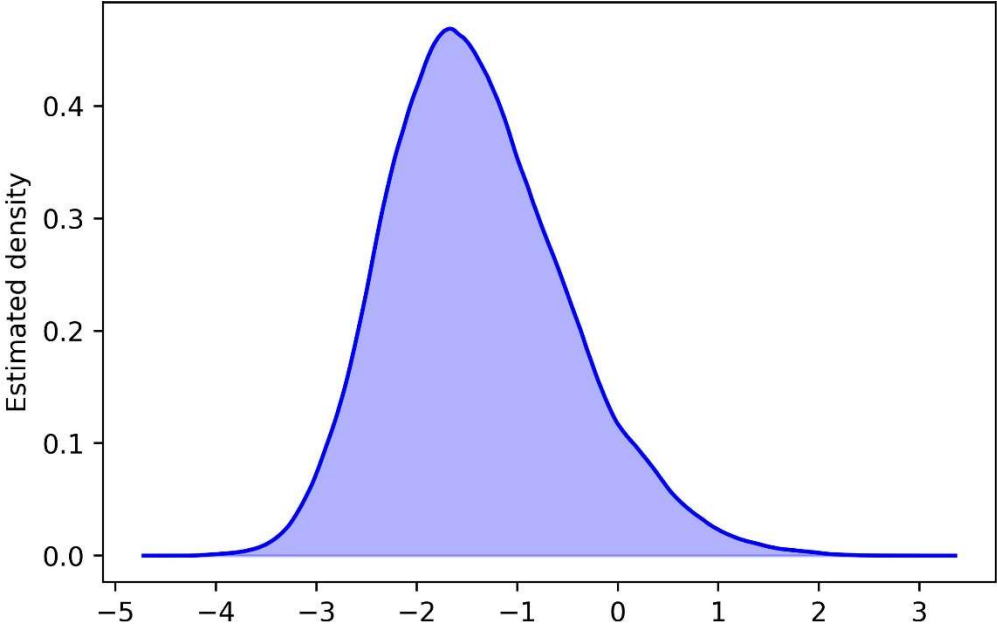
## C.2 Distribution of estimated IATEs

Figure C.1: Distribution of estimated IATE of LVT vs. NOP



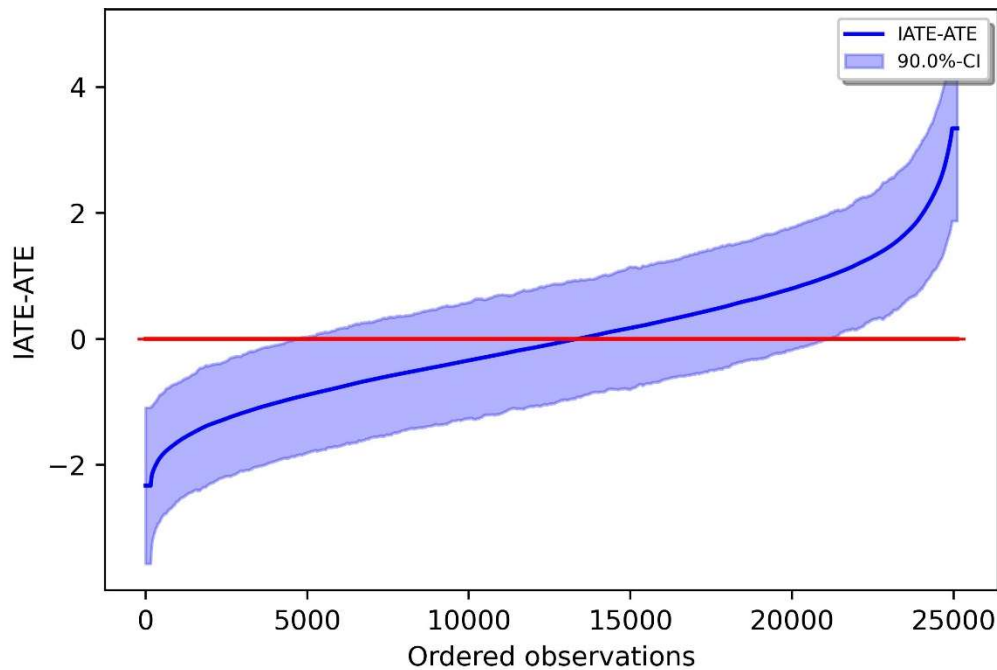
Note: The horizontal axis measures the average gain of participating in LVT relative to NOP in the cumulative number of months that one is employed as measured 30 months after the (hypothetical) programme start. Kernel smooth with Epanechnikov Kernel and Silverman (normality) bandwidth.

Figure C.2: Distribution of estimated IATE of OT vs. NOP



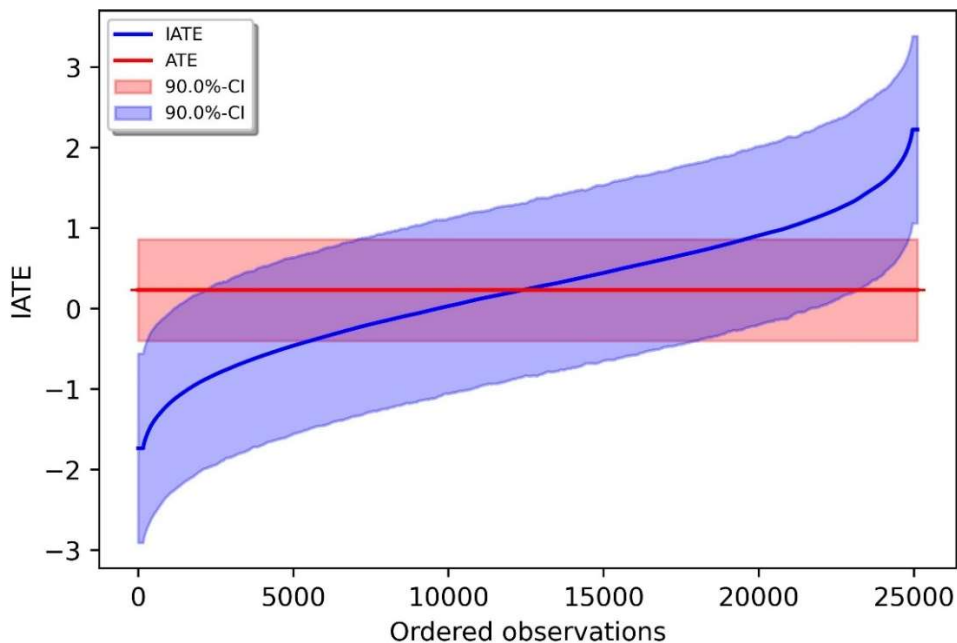
Note: The horizontal axis measures the average gain of participating in OT relative to NOP in the cumulative number of months that one is employed as measured 30 months after the (hypothetical) programme start. Kernel smooth with Epanechnikov Kernel and Silverman (normality) bandwidth.

Figure C.3: Overall heterogeneity: sorted effects *minus* ATE of SVT relative to NOP – Employment 30 months after the programme start



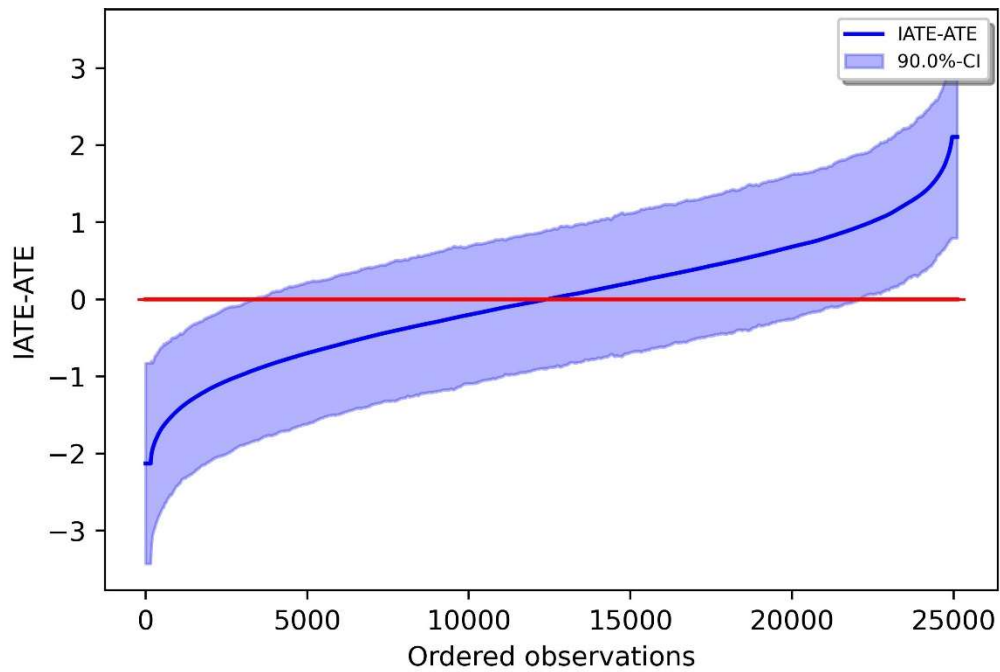
Note: (IATEs-ATE) are sorted according to their size. 90%-confidence interval based on estimated standard errors and normal distribution. The vertical axis measures, in deviation from the ATE, the average gain of participating in SVT relative to NOP in the cumulative number of months that one is employed as measured 30 months after the (hypothetical) programme start. The horizontal axis shows the rank of the ordered observations. Standard errors are smoothed by Nadaraya-Watson regression (Epanechnikov kernel with Silverman bandwidth).

Figure C.4: Overall heterogeneity: sorted effects of LVT relative to NOP – Employment 30 months after the programme start



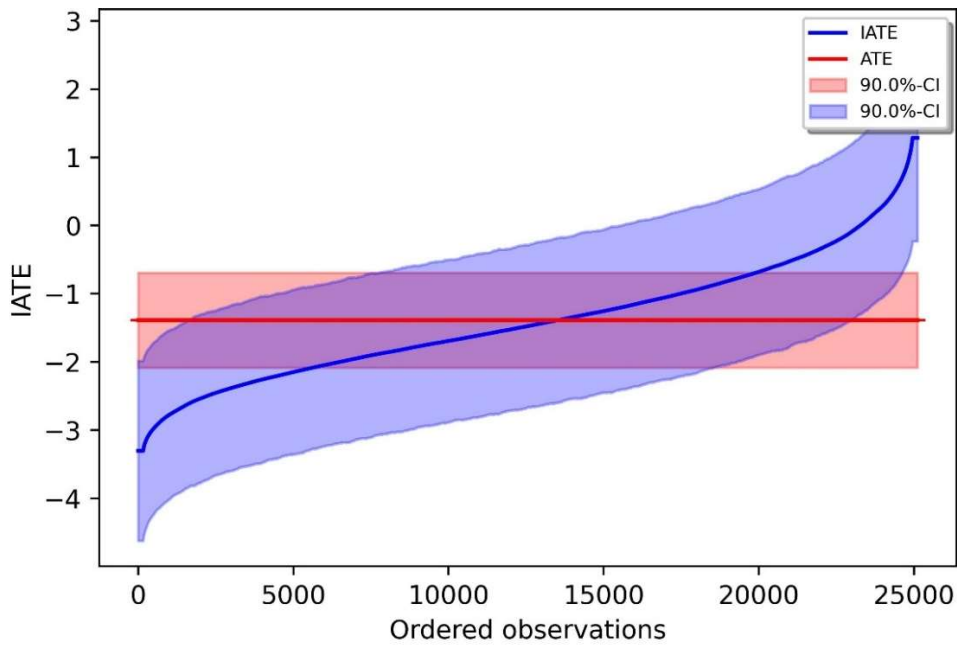
Note: IATEs are sorted according to their size. 90%-confidence interval based on estimated standard errors and normal distribution. The vertical axis measures the average gain of participating in LVT relative to NOP in the cumulative number of months that one is employed as measured 30 months after the (hypothetical) programme start. The horizontal axis shows the rank of the ordered observations. Standard errors are smoothed by Nadaraya-Watson regression (Epanechnikov kernel with Silverman bandwidth).

Figure C.5: Overall heterogeneity: sorted effects **minus** ATE of LVT relative to NOP –  
Employment 30 months after the programme start



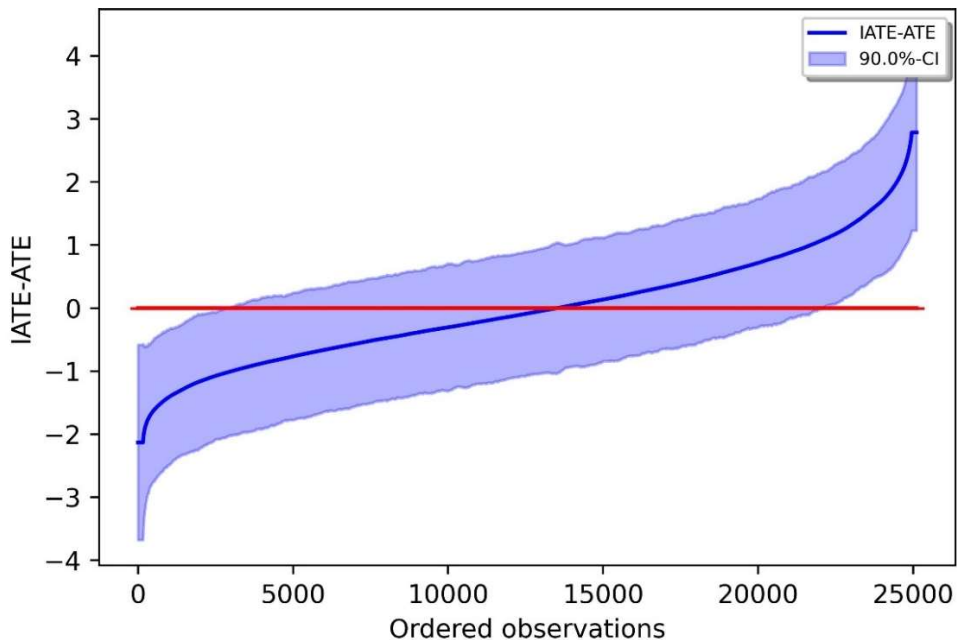
Note: (IATEs-ATE) are sorted according to their size. 90%-confidence interval based on estimated standard errors and normal distribution. The vertical axis measures, in deviation from the ATE, the average gain of participating in LVT relative to NOP in the cumulative number of months that one is employed as measured 30 months after the (hypothetical) programme start. The horizontal axis shows the rank of the ordered observations. Standard errors are smoothed by Nadaraya-Watson regression (Epanechnikov kernel with Silverman bandwidth).

Figure C.6: Overall heterogeneity: sorted effects of OT relative to NOP – Employment 30 months after the programme start



Note: IATEs are sorted according to their size. 90%-confidence interval based on estimated standard errors and normal distribution. The vertical axis measures the average gain of participating in OT relative to NOP in the cumulative number of months that one is employed as measured 30 months after the (hypothetical) programme start. The horizontal axis shows the rank of the ordered observations. Standard errors are smoothed by Nadaraya-Watson regression (Epanechnikov kernel with Silverman bandwidth).

Figure C.7: Overall heterogeneity: sorted effects *minus* ATE of OT relative to NOP – Employment 30 months after the programme start



Note: (IATEs-ATE) are sorted according to their size. 90%-confidence interval based on estimated standard errors and normal distribution. The vertical axis measures, in deviation from the ATE, the average gain of participating in OT relative to NOP in the cumulative number of months that one is employed as measured 30 months after the (hypothetical) programme start. The horizontal axis shows the rank of the ordered observations. Standard errors are smoothed by Nadaraya-Watson regression (Epanechnikov kernel with Silverman bandwidth).

## C.4 Matching results

In Table C.3 we report the estimates of the ATE on the cumulative number of months in employment 30 months after the programme start, i.e. our benchmark, based on the conventional propensity score radius matching method with bias adjustment as suggested by Lechner, Miquel and Wunsch (2011). We only report the programme effects relative to No ALMP participation (NOP). The same set of matching variables is chosen as in the MCF, but no polynomials or interactions are considered. The categorical variables are broken up into dummy variables and some of these dummies are merged in the case of small groups. The comparisons are based on slightly different common supports. The standard errors are analytical ones. These are known to be slightly conservative.

The estimated ATE are not very different from the ones obtained by MCF reported in Table 5.1 in the main text. The point estimate for the ATE of OT is nearly identical: -1.3 versus -1.4 by MCF. For LVT the point estimate obtained by the propensity score method is slightly smaller than the one obtained by MCF, but well within one standard error of the MCF estimate: 0.8 versus 1.0. Finally, the difference between the point estimates of the ATE of SVT is somewhat larger than for the other programmes, but the difference remains well within two standard errors of the MCF estimate: 2.6 versus 3.4 for the MCF estimate.

*Table C.3: Estimates of ATE using the bias-corrected propensity score estimator. Reference is No ALMP participation (NOP).*

	Short vocational training (SVT)	Long vocational training (LVT)	Orientation training (OT)
ATE (relative to NOP)	2.61 (0.51)***	0.83 (0.49)*	-1.34 (0.48)**

Note: Outcomes measured in months. All effects are population averages (ATE) relative to No ALMP Participation (NOP). Analytical standard errors are in brackets. \*, \*\*, \*\*\* indicate the precision of the estimate by showing whether the p-value of a two-sided significance test is below 10%, 5%, and 1% respectively.

## Appendix D: Common support analysis<sup>1</sup>

The assumption of common support plays a similar role in the causal forests as in matching type approaches. In our implementation, we deal with potential common support problems in two ways. First, when creating the trees for the MCF, splitting stops when one treatment is no longer available in a potential new leaf. This ‘stop splitting’ rule is very similar to extrapolation in parametric models without support (or ‘bad’ matches in matching estimation). This is technically necessary as the algorithm is based on computing within-leaf treatment effects. Therefore, usually, leaf sizes will increase when treatment shares become more asymmetric (as in the case in this application). As this procedure will not detect areas of lack of common support, we also compute treatment probabilities for the different treatments using random forests. Based on these probabilities, we used 5 different rules to deal with a potential lack of overlap:

- (i) No adjustment;
- (ii) Find for each programme comparison the minimum/maximum treatment probability within each of the four subsamples defined by treatment status; observations smaller than the largest of these four minima and larger than the smallest of these four maxima are considered to be off support and are therefore deleted;
- (iii-v) To guard against outliers, instead of considering the largest/smallest minimum/maximum, consider instead the smallest/largest 99.5<sup>th</sup>/0.5<sup>th</sup>, 99<sup>th</sup>/1<sup>st</sup>, 97.5<sup>th</sup>/2.5<sup>th</sup> quantile.

The data used to determine the cut-off values are those used to build the trees (as opposed to the ones used to populate the leaves with outcomes). These cut-offs are then applied to the second sample. This way prediction and inference will be less affected by the common support

---

<sup>1</sup> We note that the estimates of the preferred model reported in this section (in particular the ATEs for the third rule reported in Table D.1) differ slightly from the ones that we report in the main text. The reason is that we could not replicate the code, because there was a problem with fixing the seed of the random number generator.

step. This procedure also leads to a somewhat higher share of observations marked as out-off support in the sample used to determine the cut-offs than in the second sample.

*Table D.1: Sensitivity of the estimates of ATE relative to non-participation (NOP) for different rules of enforcing common support*

Common support rule	ATE relative to non-participation			Observations deleted in %
	Short vocational training (SVT)	Long vocational training (LVT)	Orientation training (OT)	
No restrictions	2.61 (0.37)	0.35 (0.36)	-1.33 (0.40)	0%
Min/max & max/min	2.46 (0.37)	0.17 (0.36)	-1.44 (0.39)	4%
<b>Min/q99.5 &amp; max/q0.5</b>	<b>2.63 (0.38)</b>	<b>0.52 (0.39)</b>	<b>-1.49 (0.42)</b>	<b>19%</b>
Min/q99 & max/q1	2.93 (0.40)	0.80 (0.41)	-1.64 (0.45)	31%
Min/q97.5 & max/q2.5	2.48 (0.44)	1.00 (0.44)	-2.20 (0.50)	43%

Note: Outcomes measured in months. All effects are population averages (ATE) relative to No ALMP Participation (NOP). Standard errors are in brackets. \*, \*\*, \*\*\* indicate the precision of the estimate by showing whether the p-value of a two-sided significance test is below 10%, 5%, and 1% respectively.

In Table D.1 we show the sensitivity of the estimated ATEs of accumulated employment after 30 months relative to nonparticipation for different common support rules. The final row contains the number of observations deleted due to the particular common support rule.

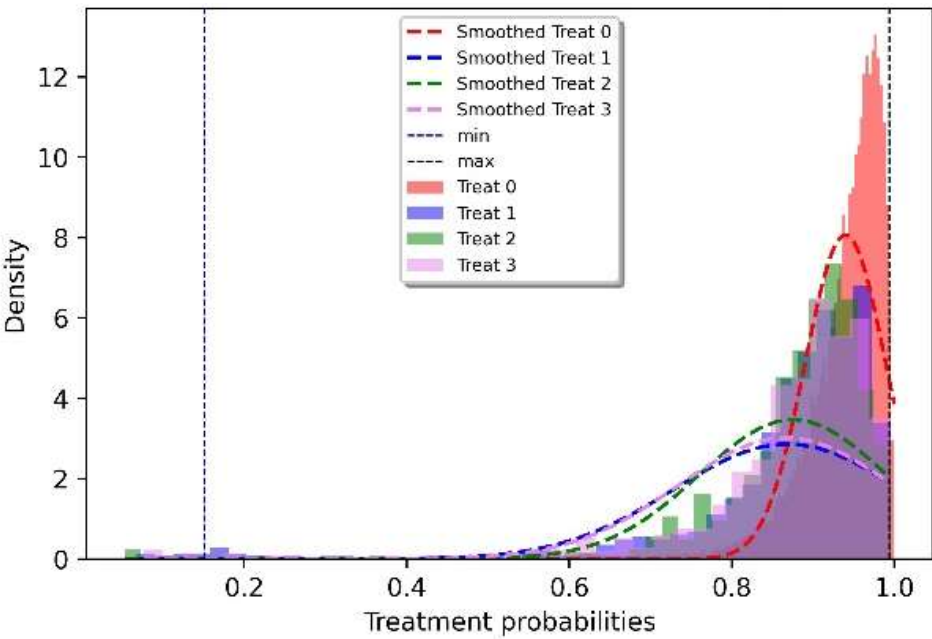
The first observation is that the number of observations to be considered off support increases up to almost half of the sample. The point estimates vary between one and two magnitudes of their estimated standard errors. Not surprisingly, this variation is most excessive when larger chunks of the data are removed. However, the qualitative interpretation of the estimated effects (i.e., their relative ordering between programmes) does not change much, especially when the first three rules are compared. Not surprisingly, as more than 30% of the observations are discarded in the last two rules, the estimates become more sensitive, especially the ATE of LVT and OT which, respectively, increases and decreases substantially. Interestingly, the standard errors do not increase by much despite the large decrease in the sample size, indicating that the remaining data are a more homogeneous population. In other words, reduced variability within the remaining data offsets the additional variability of the ATEs due to the reduced sample size.



For considering which rule to use to determine which observations are on and off support, it is reassuring that the results are not overly sensitive. Nevertheless, we want to keep the data labelled as ‘off-support’ to a minimum as otherwise the population to which our results apply may change substantially from the initial population of interest. We only want to remove observations with non-overlapping (extreme) propensity scores for which we expect inconsistent estimation. Investigating propensity scores helps in this respect.

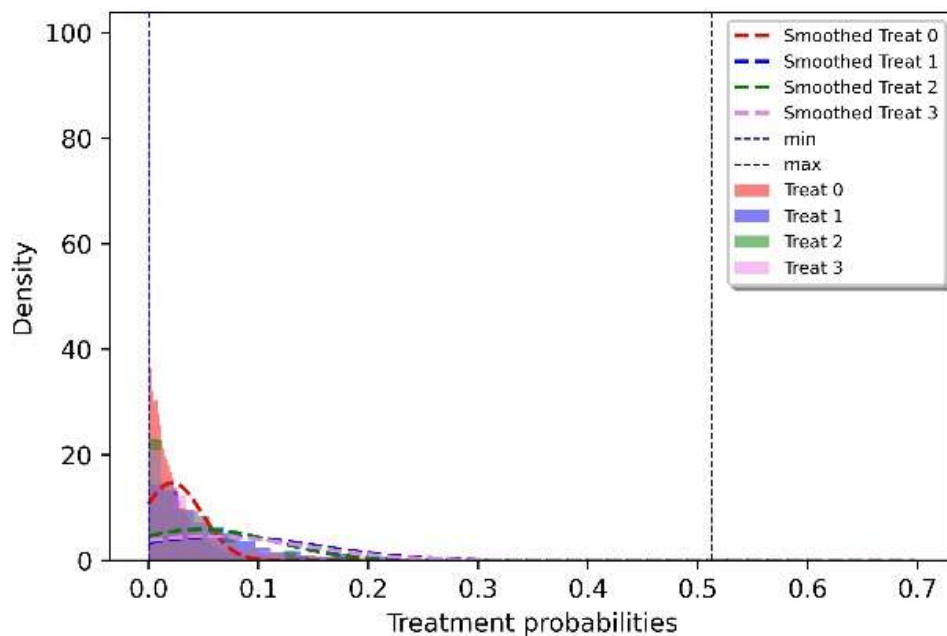
*Figure D.1: Distribution of treatment probabilities in subsamples by treatment state (min-/max-rule)*

*Figure D.1.1: Probability of being in NOP*



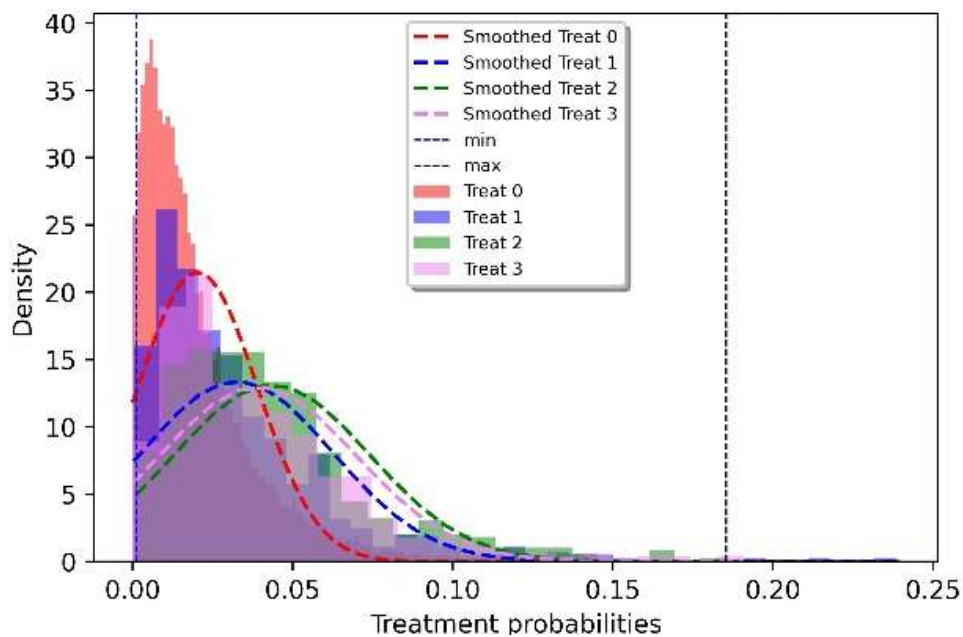
Note: The distribution shown is for the subsample used to populate the leaves with outcomes.

Figure D.1.2: Probability of being in SVT



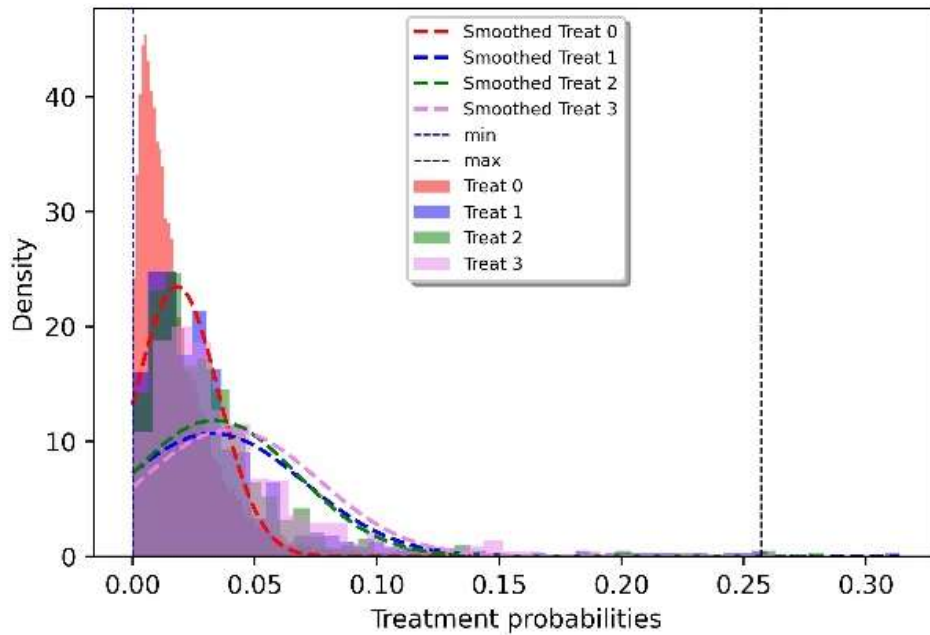
Note: The distribution shown is for the subsample used to populate the leaves with outcomes.

Figure D.1.3: Probability of being in LVT



Note: The distribution shown is for the subsample used to populate the leaves with outcomes.

Figure D.1.4: Probability of being in OT

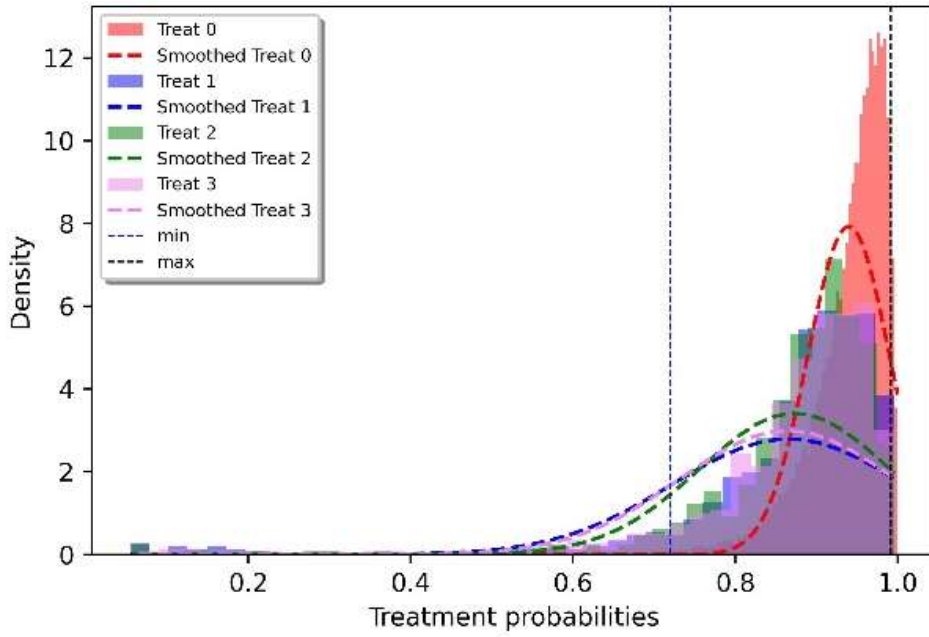


Note: The distribution shown is for the subsample used to populate the leaves with outcomes.

In Figures D.1 and D.2 we show graphically the histograms of such propensity scores, i.e., the probabilities of being in the four different treatment states. Each subplot refers to the distribution of one probability in the four subpopulations. In addition, the vertical lines refer to the cut-offs used. Both figures have the same propensity scores, but Figure D.1 shows the cut-offs of the min-/max-rule while Figure D.2 is based on the more conservative 99.5% rule which is less prone to outliers.

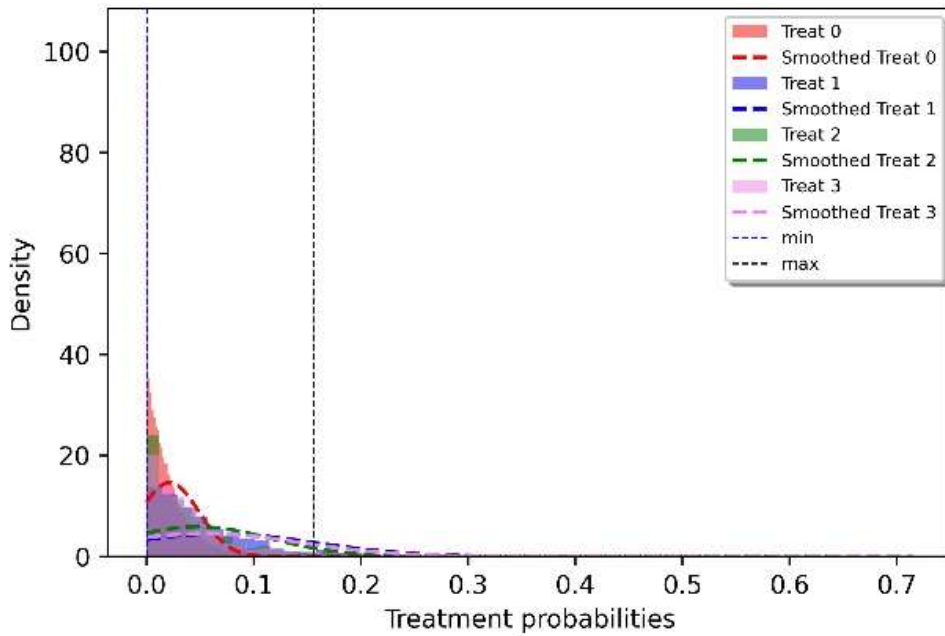
Figure D.2: Distribution of treatment probabilities in subsamples by treatment state  
(99.5% rule)

Figure D.2.1: Probability of being in NOP



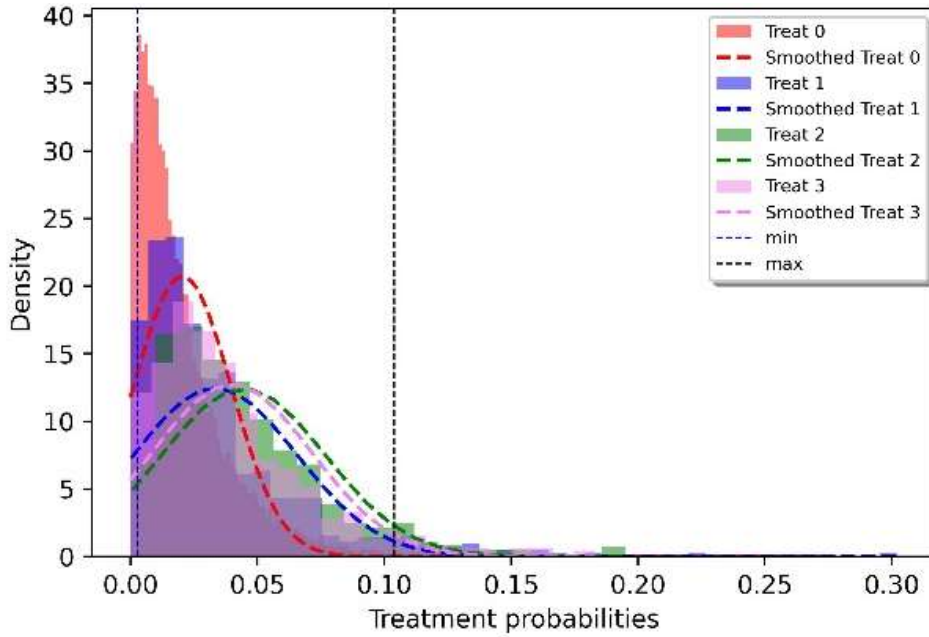
Note: The distribution shown is for the subsample used to populate the leaves with outcomes.

Figure D.2.2: Probability of being in SVT



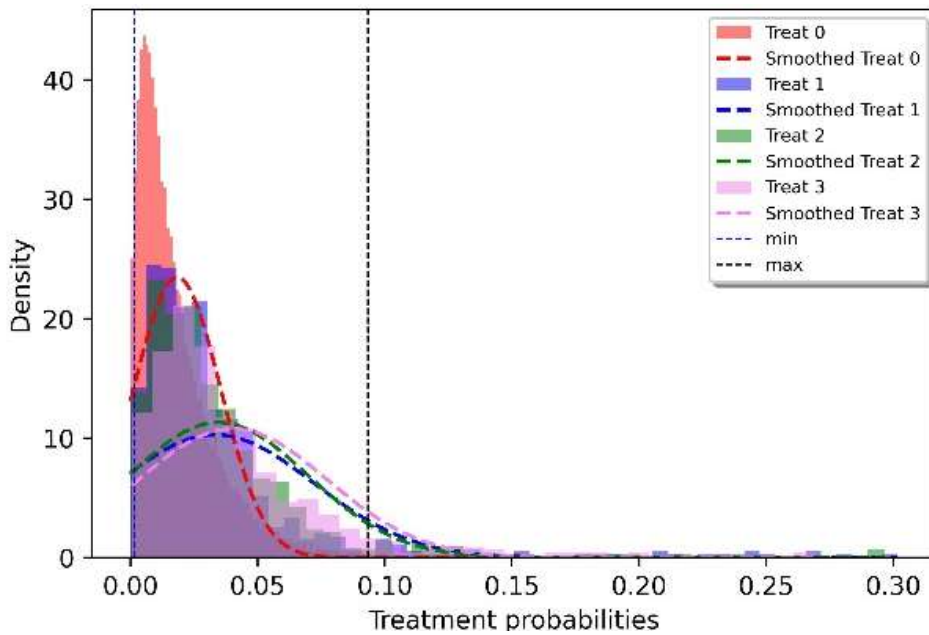
Note: The distribution shown is for the subsample used to populate the leaves with outcomes.

Figure D.2.3: Probability of being in LVT



Note: The distribution shown is for the subsample used to populate the leaves with outcomes.

Figure D.2.4: Probability of being in OT



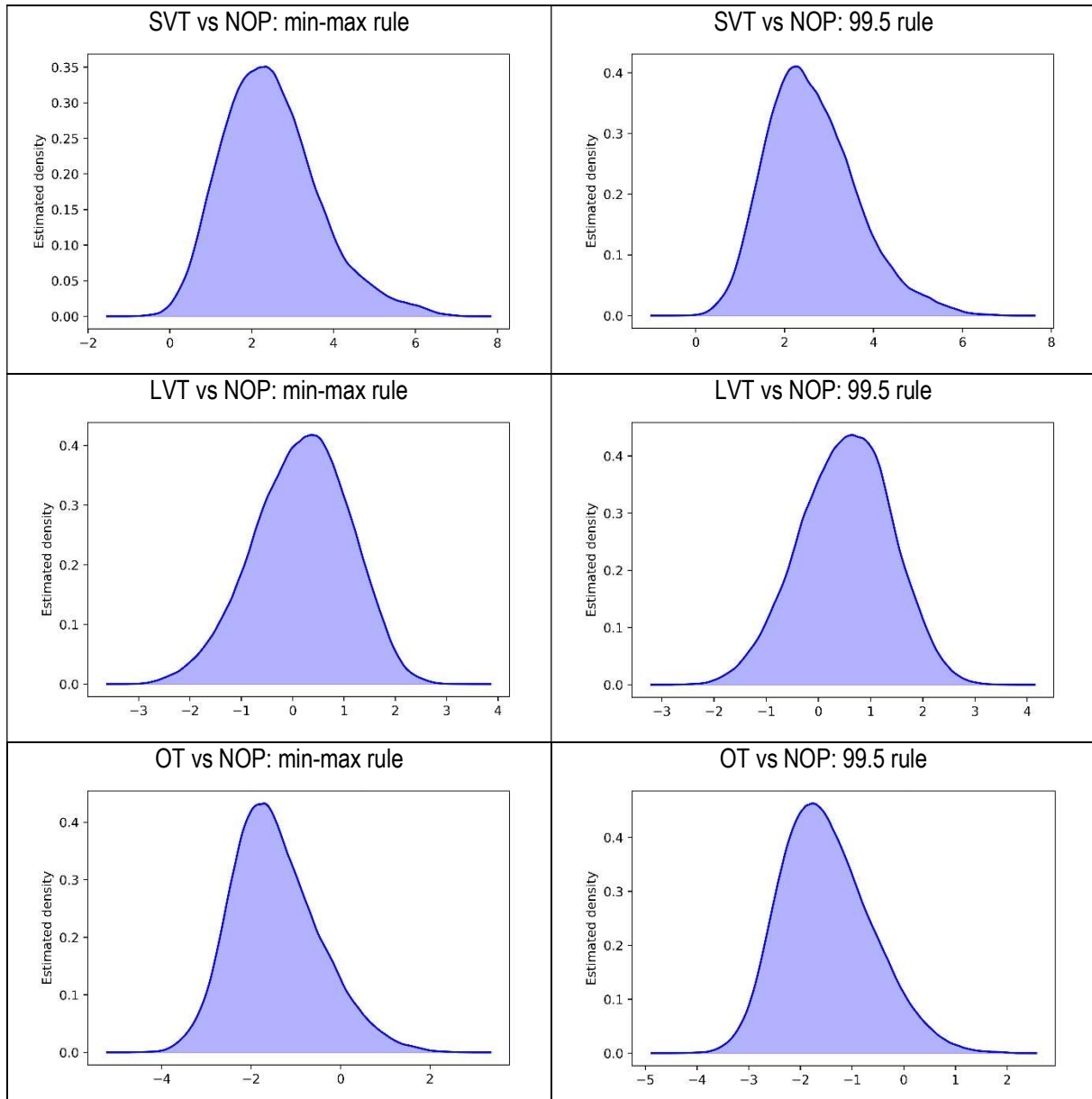
Note: The distribution shown is for the subsample used to populate the leaves with outcomes.

From visible inspection, it comes clear that the min-max rule (Figure D.1) includes observations that appear to be extreme outliers in terms of their treatment probabilities. By contrast, the more prudent 99.5% rule which is shown in Figure D.2, appears to successfully remove extreme observations, without reducing the data much more than required. At the

thresholds in Figure D.2 only the densities in the NOP subsample seems visually very low. However, this is not a concern, because this subsample is much larger than the other subsamples. Thus, it does not appear necessary to use a more stringent rule that would lead to a larger reduction of the data used for estimation.

Next, Figure D.3 shows how using the min/max rule or 99.5% rule changes the distribution of the estimated IATE (by changing the estimation as well as the distribution of the covariate values for which the IATEs are computed). The main difference between the distributions is that the distribution emerging from the 99.5% rule is that it seems to be slightly more compressed. This is in line with the aforementioned observation that by dropping the observations off support the variability is reduced. While this difference is difficult to spot in Figure D.3, this difference (not shown) is more pronounced if the 99.5% rule is compared to the no-adjustment case. This means that the imposition of common support matters not so much for the estimation of the ATEs, but more for the variability of the IATEs.

Figure D.3: Distribution of IATEs: Min-max rule vs. 99.5% rule



Finally, Table D.2 shows how the distribution of some key confounders differ in the subsamples kept and removed by applying the 99.5%-rule, which is used in the main body of the paper. Here, again, the results for the random subsample that is used for populating the leaves with outcome variables are shown. The results for the other sample used to build the forest are very similar and are omitted for the sake of brevity.

The table shows that the sizes of all programme groups are reduced by similar magnitudes (13% to 16%). Generally, the means of the variables do not differ much between the

subsamples. The most striking difference is the removal of a more than proportional fraction of individuals with the lowest level of education. Besides, the difference in the composition of the on and off support samples seems the largest for the NOP sample. However, given that the ATEs hardly differ whether off-support observations are dropped, suggests that these differences do not systematically bias the estimated treatment effects in a particular direction.

*Table D.2: Means of some key variables in the sample on and off support (99.5% rule)*

Variable	Non-participation (NOP)			Short vocational training (SVT)			Long vocational training (LVT)			Orientation training (OT)		
	All	On	Off	All	On	Off	All	On	Off	All	On	Off
Women (share in %)	48.5	47.1	55.5	30.9	31.3	31.0	41.0	40.4	44.0	46.3	44.6	55.9
Age (in years)	35.2	35.6	34.1	34.1	34.4	31.6	34.1	34.7	30.8	33.5	33.7	32.4
Proficiency in Dutch (0-3) <sup>1</sup>	2.4	2.5	2.1	2.5	2.5	2.5	2.7	2.7	2.7	2.6	2.7	2.5
Months unemployed in the last 10 years	17.9	19.3	11.0	18.7	18.6	19.2	15.8	16.3	13.0	16.7	16.5	17.9
Months unemployed in last 2 years	3.9	4.0	3.1	3.5	3.4	4.3	2.8	2.9	2.6	3.1	3.0	3.8
Lowest education level (share in %)	14.2	11.7	26.7	19.6	17.3	34.4	5.1	4.3	9.8	10.2	9.6	13.1
Number of observations	28155	23543	4612	663	576	87	603	512	91	561	477	84
Share in %	100	84	16	100	87	13	100	85	15	100	85	15

Notes: *On / off* refers to the subsamples that are kept/deleted for estimation. *All* refers to the full sample. All statistics refer to the evaluation sample only (i.e., the sample from which the outcome variables are taken to populate the leaves).

<sup>1</sup> Proficiency in Dutch = 0 if no knowledge; = 1 if limited; = 2 if good; =3 if very good. <sup>2</sup> For non-participants in ALMP (NOP) the date at which the ALMP starts (is predicted) (See Section 4.4).

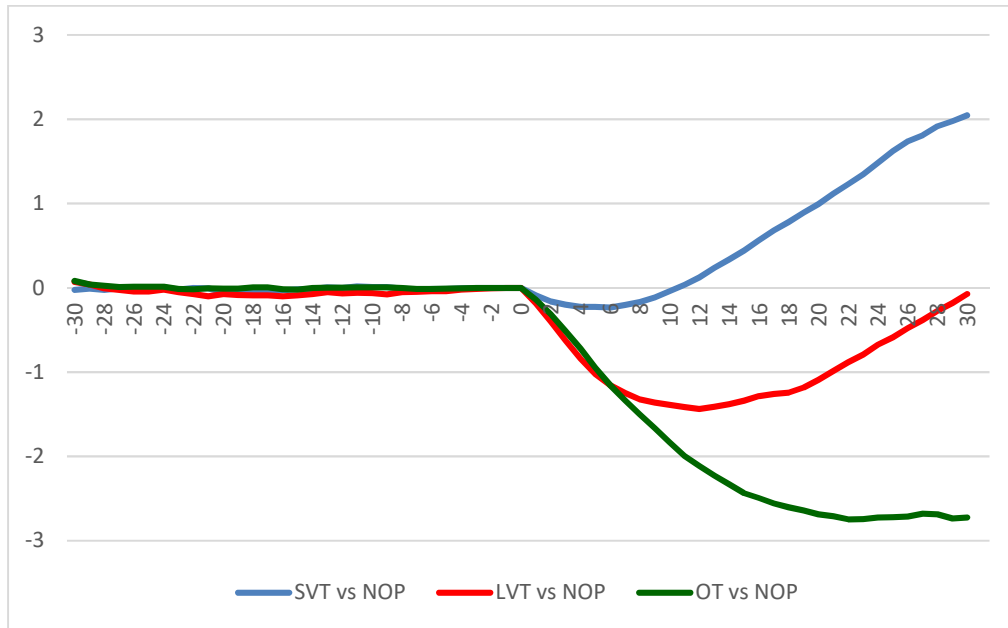


## Appendix E: The Second placebo test

For this placebo analysis, we retain all individuals for whom at least 30 months of labour market history before the start of the unemployment spell is available. We retain hereby 46'455 individuals (77% of the full sample). We then estimate the ATEs during the 30 pre-entry months in the same way (using the same conditioning information) as we do in the main analysis for the 30 months since the programme start. Furthermore, we also re-estimate the cumulative ATEs during the 30 months since the programme start on this restricted sample to check to what extent this subsample is selective.

The findings of this analysis are reported in Figure E.1. During the 30 months before entry into unemployment, all ATEs are very close to zero and never statistically different from zero. The ATEs after entry into unemployment should be compared to those reported 9 and 30 months after the programme start in Table 5.1. The ATEs are very similar after 9 months. After 30 months the point estimates are a bit more different, in particular for OT, for which the average cumulative number of months in employment is reduced by nearly 3 months instead of 1.4 months in the main analysis. For SVT the ATE is also slightly smaller, plus 2 instead of plus 2.5 months, a difference of about one standard error. Finally, for LVT the ATEs are both very close to zero. Overall, the particular sample selection might induce only some doubt that this placebo test is not representative of the ATE of OT.

Figure E.1: Cumulative monthly employment effects



Note: On the horizontal axis the negative numbers count the number of months before the month of entry in unemployment which is denoted by zero; the positive numbers denote the number of months that have elapsed since programme entry, as in the main analysis. The vertical axis denotes the average difference in the cumulative number of months in employment between programme participation (SVT, LVT or OT) and the counterfactual of no programme participation (NOP), i.e. the ATEs of programme participation relative to NOP. The ATEs are estimated on the sample of 46'455 individuals for whom at least 30 months of labour market history is available.