Online Appendices

Heterogeneous Employment Effects of Job Search Programmes: A Machine Learning Approach

Michael C. Knaus[†], Michael Lechner[‡], and Anthony Strittmatter[*]

Date this version has been printed: 17 February 2020

[†] Michael C. Knaus assistant professor of econometrics at the University St. Gallen and is also affiliated with IZA, Bonn. Email: michael.knaus@unisg.ch

[‡] Michael Lechner is fully professor of econometrics at the University St. Gallen and is also affiliated with CEPR, London, CESIfo, Munich, IAB, Nuremberg, and IZA, Bonn. Email: michael.lechner@unisg.ch

[*] Anthony Strittmatter is assistant professor of econometrics at the University St. Gallen. Email: anthony.strittmatter@unisg.ch

Appendix A: Sample selection

Table A.1 documents the sample selection steps. Additional observation are trimmed to ensure common support as shown in Figure D.1.

Table A.1: Sample selection steps.

Selection criteria		Remaining sample size
Population: all new jobseekers during the year 2003		238,902
Exclude Geneva and five other employment offices	-19,464	219,438
Exclude jobseekers not (yet) assigned to a caseworker	-4,289	215,149
Exclude foreigners with work permit shorter than one year	-5,399	209,750
Exclude jobseekers without unemployment benefit claim	-18,434	191,316
Exclude jobseekers who applied for or claim disability insurance	-3,163	188,153
Restrict to prime-age population (24 to 55 years old)	-51,649	136,504
Exclude unemployed whose caseworker did not respond to the questionnaire	-31,469	105,035
Exclude unemployed whose caseworkers did not respond to the cooperativeness question	-4,915	100,120
Exclude participants in other ALMP than JSP	-8,787	91,333
Exclude individuals employed at (pseudo) treatment date	-6,135	85,198

Note: Only the last two sample selection steps differ from Huber, Lechner, Mellace (2017).

Appendix B: Descriptive statistics

The following table shows unconditional means and standard deviations by participation status as well as standardised differences to illustrate selection into participation.

Table B.1: Descriptive statistics of confounding variables by JSP participation status.

	Participants		Non-Participants		Std. Diff.
	Mean	S.D.	Mean	S.D.	
	(1)	(2)	(3)	(4)	(5)
Characteristics of unemployed persons					
Female	0.45	-	0.44	-	0.56
× French speaking REA	0.04	-	0.11	-	19.51
× Italian speaking REA	0.01	-	0.04	-	11.85
Age (in 10 years)	3.73	0.88	3.66	0.86	5.60
Unskilled	0.22	-	0.23	-	1.80
× French speaking REA	0.03	-	0.05	-	8.36
× Italian speaking REA	0.01	-	0.03	-	8.62
Semi-skilled qualification	0.15	-	0.16	-	2.45
× French speaking REA	0.02	-	0.05	-	12.10
× Italian speaking REA	0.002	-	0.01	-	5.16
Skilled qualification without degree	0.03	-	0.05	-	4.72
× French speaking REA	0.003	-	0.02	-	11.22
× Italian speaking REA	0.002	-	0.01	-	4.11
Employability rating low	0.12	-	0.14	-	3.98
× French speaking REA	0.01	-	0.02	-	9.87
× Italian speaking REA	0.004	-	0.01	-	4.94
Employability rating medium	0.77	-	0.74	-	5.80
× French speaking REA	0.07	-	0.19	-	26.32
× Italian speaking REA	0.02	-	0.05	-	11.57
# of unemp. spells in last 2 years	0.41	0.98	0.64	1.27	13.86
× French speaking REA	0.05	0.36	0.19	0.76	16.84
× Italian speaking REA	0.02	0.22	0.07	0.46	10.16
# of emp. spells in last 5 years	0.10	0.13	0.13	0.15	14.70
Fraction of months emp. in last 2 years	0.83	0.22	0.79	0.25	12.57
× French speaking REA	0.06	0.22	0.19	0.35	30.04
× Italian speaking REA	0.02	0.13	0.06	0.22	15.77
Past income (in 10,000 CHF)	4.58	2.02	4.16	2.05	14.50
Prev. job in primary sector	0.06	-	0.10	-	10.44
Prev. job in secondary sector	0.16	-	0.13	-	6.04
Prev. job in tertiary sector	0.63	-	0.58	-	7.07
Prev. job self-employed	0.004	-	0.01	-	3.01
Prev. job manager	0.08	-	0.07	-	1.85
Prev. job skilled worker	0.63	-	0.60	-	4.70
Prev. job unskilled worker	0.26	-	0.29	-	5.01

Table continues on next page >

Table B.1 continued.

	Participants		Non-Participants		Std. Diff.
	Mean	S.D.	Mean	S.D.	
	(1)	(2)	(3)	(4)	(5)
Characteristics of unemployed persons					
Native language not German, French, or Italian	0.29	-	0.32	-	5.40
× French speaking REA	0.02	-	0.08	-	18.01
× Italian speaking REA	0.01	-	0.02	-	9.80
Married	0.47	-	0.49	-	2.35
Foreigner with temporary residence permit	0.11	-	0.14	-	6.96
Foreigner with permanent residence permit	0.23	-	0.25	-	3.12
Foreigner with mother tongue similar to canton's language	0.12	-	0.11	-	2.40
Lives in big city	0.17	-	0.17	-	0.05
Lives in medium sized city	0.16	-	0.13	-	4.83
Start of JSP participation in the second unemp. quarter	0.45	-	0.46	-	0.38
Caseworker characteristics					
Female	0.45	-	0.41	-	6.94
× French speaking REA	0.02	-	0.09	-	22.33
× Italian speaking REA	0.01	-	0.02	-	6.15
Age (in 10 years)	4.43	1.16	4.44	1.16	0.77
× French speaking REA	0.37	1.29	1.14	2.04	31.79
× Italian speaking REA	0.11	0.70	0.34	1.19	16.43
Tenure (in years)	5.54	3.23	5.86	3.31	6.84
× French speaking REA	0.47	1.78	1.59	3.07	31.36
× Italian speaking REA	0.21	1.39	0.60	2.29	14.58
Own unemp. experience	0.63	-	0.63	-	0.54
× French speaking REA	0.05	-	0.17	-	26.33
× Italian speaking REA	0.02	-	0.05	-	11.73
Education above vocational training	0.45	-	0.43	-	2.36
× French speaking REA	0.04	-	0.10	-	17.68
× Italian speaking REA	0.01	-	0.03	-	9.46
Education tertiary track	0.21	-	0.24	-	4.68
× French speaking REA	0.02	-	0.09	-	21.92
× Italian speaking REA	0.004	-	0.02	-	8.25
Vocational training degree	0.26	-	0.23	-	5.63
× French speaking REA	0.002	-	0.01	-	9.64
× Italian speaking REA	0.01	-	0.04	-	11.28
Indicator for missing caseworker characteristics	0.04	-	0.04	-	0.13

Table continues on next page >

Table B.1 continued.

	Participants		Non-Participants		Std. Diff.
	Mean	S.D.	Mean	S.D.	
	(1)	(2)	(3)	(4)	(5)
Allocation of unemployed persons to caseworkers					
By industry	0.66	-	0.53	-	17.73
× French speaking REA	0.05	-	0.10	-	12.88
× Italian speaking REA	0.01	-	0.04	-	11.32
By occupation	0.58	-	0.56	-	3.08
× French speaking REA	0.06	-	0.17	-	25.14
× Italian speaking REA	0.01	-	0.05	-	14.27
By age	0.04	-	0.03	-	2.58
By employability	0.07	-	0.07	-	0.12
By region	0.09	-	0.12	-	7.55
Other allocation type	0.07	-	0.07	-	1.37
Local labour market characteristics					
French speaking REA	0.08	-	0.25	-	33.30
Italian speaking REA	0.03	-	0.08	-	16.81
Cantonal unemployment rate (in %)	3.64	0.77	3.75	0.86	9.23
× French speaking REA	0.32	1.10	1.05	1.86	33.93
× Italian speaking REA	0.11	0.69	0.34	1.16	16.61
Cantonal GDP per capita (in 10,000 CHF)	5.13	0.92	4.92	0.93	15.75
# of caseworker	989		1,282		
# of observations	12,998		72,200		

Note: We report unconditional means for all variables, standard deviations (S.D.) for all non-binary variables, and standardised differences between participants and non-participants. Rosenbaum and Rubin (1983) consider a standardised difference of more than 20 as being 'large'. We report the descriptive statistics of the outcome variables in Table 1 of the main text. REA is the abbreviation for regional employment agency. For many variables we include interactions with a dummy for French (× French speaking REA) and Italian (× Italian speaking REA) speaking regional employment agencies. To account for categorical variables, we omit the dummies some qualification degree, employability rating high, lives in small city, and German speaking regional employment agencies.

Appendix C: Theorem 1 and Equation (3)

C.1 Proof of Theorem 1

The following proof is based on the seminal contributions of Rosenbaum and Rubin (1983). To proof the identification of CATEs, we use the definition (see Section 4.1 in the main text),

$$\gamma(z) = E\left[Y_i^1 - Y_i^0 \middle| Z_i = z\right] = E\left[Y_i^1 \middle| Z_i = z\right] - \left[Y_i^0 \middle| Z_i = z\right].$$

Then we apply the law of iterative expectations

$$\gamma(z) = E_{X|Z=z}\left[E\left[Y_i^1 \middle| X_i = x, Z_i = z\right] \middle| Z_i = z\right] - E_{X|Z=z}\left[E\left[Y_i^0 \middle| X_i = x, Z_i = z\right] \middle| Z_i = z\right].$$

When we condition on the confounders X_i, the potential outcomes are independent of the treatment indicator D_i (Assumption 1),

$$\begin{aligned}\gamma(z) = {} & E_{X|Z=z}\left[E\left[Y_i^1 \middle| D_i = 1, X_i = x, Z_i = z\right] \middle| Z_i = z\right] \\ & - E_{X|Z=z}\left[E\left[Y_i^0 \middle| D_i = 0, X_i = x, Z_i = z\right] \middle| Z_i = z\right].\end{aligned}$$

Conditional on the treatment status, the potential outcomes equal the observed outcome,

$$\begin{aligned}\gamma(z) = {} & E_{X|Z=z}[E[Y_i | D_i = 1, X_i = x, Z_i = z] | Z_i = z] \\ & - E_{X|Z=z}[E[Y_i | D_i = 0, X_i = x, Z_i = z] | Z_i = z].\end{aligned}$$

C.2 Consistency of Estimator in Equation (3)

We consider the structural working model (see equation (2))

$$Y_i = Z_i \beta_t + \frac{T_i Z_i}{2} \delta + u_i,$$

which includes a constant, such that $E[u_i] = 0$. The IPW weights are

$$\begin{aligned} w(D_i, X_i, Z_i) = \widetilde{W}_i &= \frac{\dfrac{D_i - p(X_i, Z_i)}{p(X_i, Z_i)\left(1 - p(X_i, Z_i)\right)}}{D_i \sum_{i=1}^N \dfrac{D_i}{p(X_i, Z_i)} + (1 - D_i) \sum_{i=1}^N \dfrac{1 - D_i}{1 - p(X_i, Z_i)}} \\ &= \left(\frac{D_i}{\bar{q}_1 p(X_i, Z_i)} - \frac{1 - D_i}{\bar{q}_0 \left(1 - p(X_i, Z_i)\right)}\right), \end{aligned}$$

with $\bar{q}_1 = \sum_{i=1}^N D_i / p(X_i, Z_i)$ and $\bar{q}_0 = \sum_{i=1}^N (1 - D_i) / \left(1 - p(X_i, Z_i)\right)$.

The objective function is:

$$\underset{\hat{\delta}}{\operatorname{argmin}} \sum_{i=1}^{N} \left[\widetilde{W}_i T_i \left(Y_i - \frac{T_i Z_i}{2} \hat{\delta} \right)^2 \right].$$

The first order condition is:

$$-2 \sum_{i=1}^{N} \left[\frac{Z_i' \widetilde{W}_i}{2} \left(Y_i - \frac{T_i Z_i}{2} \hat{\delta} \right) \right] = 0.$$

We derive the closed form solution for $\hat{\delta}$:

$$\hat{\delta} = 2 \left(\sum_{i=1}^{N} Z_i' \widetilde{W}_i T_i Z_i \right)^{-1} \sum_{i=1}^{N} Z_i' \widetilde{W}_i Y_i.$$

We plug-in the structural working model:

$$\hat{\delta} = 2 \left(\sum_{i=1}^{N} Z_i' \widetilde{W}_i T_i Z_i \right)^{-1} \sum_{i=1}^{N} Z_i' \widetilde{W}_i \left(Z_i \beta_t + \frac{T_i Z_i}{2} \delta + u_i \right),$$

$$= 2\beta_t \left(\frac{1}{N} \sum_{i=1}^{N} Z_i' \widetilde{W}_i T_i Z_i \right)^{-1} \frac{1}{N} \sum_{i=1}^{N} Z_i' \widetilde{W}_i Z_i + \delta + 2 \left(\frac{1}{N} \sum_{i=1}^{N} Z_i' \widetilde{W}_i T_i Z_i \right)^{-1} \frac{1}{N} \sum_{i=1}^{N} Z_i' \widetilde{W}_i u_i.$$

We take the probability limit

$$plim(\hat{\delta}) = 2\beta_t E\left[Z_i' \widetilde{W}_i T_i Z_i \right]^{-1} E\left[Z_i' \widetilde{W}_i Z_i \right] + \delta$$
$$+ 2E\left[Z_i' \widetilde{W}_i T_i Z_i \right]^{-1} plim \left(\frac{1}{N} \sum_{i=1}^{N} Z_i' \widetilde{W}_i u_i \right),$$

where $plim\left(\left(\frac{1}{N} \sum_{i=1}^{N} Z_i' \widetilde{W}_i T_i Z_i \right)^{-1} \right) = E\left[Z_i' \widetilde{W}_i T_i Z_i \right]^{-1}$ under the assumption that $\left(\frac{1}{N} \sum_{i=1}^{N} Z_i' \widetilde{W}_i T_i Z_i \right)$ is non-singular w.p.a.1 and $plim\left(\frac{1}{N} \sum_{i=1}^{N} Z_i' \widetilde{W}_i Z_i \right) = E\left[Z_i' \widetilde{W}_i Z_i \right]$. Under the CIA, $plim\left(\frac{1}{N} \sum_{i=1}^{N} Z_i' \widetilde{W}_i u_i \right) = E\left[Z_i' \widetilde{W}_i u_i \right] = 0$ should hold. Applying Slutsky's theorem gives

$$plim(\hat{\delta}) = \delta + 2\beta_t E\left[Z_i' \widetilde{W}_i T_i Z_i \right]^{-1} E\left[Z_i' \widetilde{W}_i Z_i \right].$$

Accordingly, we have to show that $E\left[Z_i' \widetilde{W}_i Z_i \right] = 0$ to achieve $plim(\hat{\delta}) = \delta$.

$$E\left[Z_i' \widetilde{W}_i Z_i \right] = E\left[Z_i' \left(\frac{D_i}{\bar{q}_1 p(X_i, Z_i)} - \frac{1 - D_i}{\bar{q}_0 \left(1 - p(X_i, Z_i) \right)} \right) Z_i \right],$$

$$= E_{X,Z} \left[E\left[Z_i' \left(\frac{D_i}{\bar{q}_1 p(X_i, Z_i)} - \frac{1 - D_i}{\bar{q}_0 \left(1 - p(X_i, Z_i) \right)} \right) Z_i \middle| X_i = x, Z_i = z \right] \right],$$

$$= E_{X,Z}\left[E\left[Z_i'\frac{D_i}{\bar{q}_1 p(X_i,Z_i)}Z_i\middle|X_i=x,Z_i=z\right]\right.$$

$$\left.-E\left[Z_i'\frac{1-D_i}{\bar{q}_0\big(1-p(X_i,Z_i)\big)}Z_i\middle|X_i=x,Z_i=z\right]\right],$$

$$= E_{X,Z}\left[\frac{1}{\bar{q}_1 p(X_i,Z_i)}E[Z_i'D_iZ_i|X_i=x,Z_i=z]\right.$$

$$\left.-\frac{1}{\bar{q}_0\big(1-p(X_i,Z_i)\big)}E[Z_i'(1-D_i)Z_i|X_i=x,Z_i=z]\right]$$

$$= E\left[\frac{\bar{q}_0-\bar{q}_1}{\bar{q}_1\bar{q}_0}Z_i'Z_i\right].$$

Now we show that $E[\bar{q}_1]=E[\bar{q}_0]$ to finish the proof:

$$E[\bar{q}_1]-E[\bar{q}_0]=E\left[\sum_{i=1}^{N}\frac{D_i}{p(X_i,Z_i)}-\sum_{i=1}^{N}\frac{1-D_i}{1-p(X_i,Z_i)}\right],$$

$$=\sum_{i=1}^{N}\left(E\left[\frac{D_i}{p(X_i,Z_i)}\right]-E\left[\frac{1-D_i}{1-p(X_i,Z_i)}\right]\right)$$

$$=\sum_{i=1}^{N}E_{X,Z}\left[E\left[\frac{D_i-p(X_i,Z_i)}{p(X_i,Z_i)\big(1-p(X_i,Z_i)\big)}|X_i=x,Z_i=z\right]\right],$$

$$=\sum_{i=1}^{N}E_{X,Z}\left[\frac{E[D_i|X_i=x,Z_i=z]-p(X_i,Z_i)}{p(X_i,Z_i)\big(1-p(X_i,Z_i)\big)}\right]=0,$$

with $E[D_i|X_i=x,Z_i=z]=p(X_i,Z_i)$.

q.e.d.

Appendix D: Propensity score and matching quality

Table D.1 reports the average marginal effects of the propensity score estimation to illustrate selection into participation. Most of the significant coefficients confirm the observation that unemployed with higher skills and labour market success are more likely to participate in the program.

Table D.2 shows that inverse probability weighting successfully balances the covariates as indicated by a maximum standardised difference of 2.44 and a mean standardised difference of 0.7.

Table D.1: Average marginal effects in the propensity score estimation.

	Av. Marg. Eff.	S.E.
	(1)	
Characteristics of unemployed persons		
Female	0.01	(0.004)
× French speaking REA	0.01	(0.01)
× Italian speaking REA	-0.03	(0.02)
Age (in 10 years)	-0.01	(0.01)
Age^2/10,000	0.21	(0.17)
Unskilled	0.01*	(0.01)
× French speaking REA	0.10***	(0.02)
× Italian speaking REA	0.05**	(0.02)
Semi-skilled qualification	0.002	(0.01)
× French speaking REA	0.06***	(0.01)
× Italian speaking REA	0.03	(0.03)
Skilled qualification without degree	0.01*	(0.01)
× French speaking REA	-0.03	(0.03)
× Italian speaking REA	0.02	(0.03)
Employability rating low	-0.04***	(0.01)
× French speaking REA	0.10***	(0.03)
× Italian speaking REA	0.13***	(0.03)
Employability rating medium	-0.02	(0.01)
× French speaking REA	0.09***	(0.02)
× Italian speaking REA	0.07***	(0.02)
# of unemp. spells in last 2 years	-0.01***	(0.002)
× French speaking REA	0.003	(0.004)
× Italian speaking REA	0.004	(0.01)
Number of emp. spells in last 5 years	-0.08***	(0.01)
Fraction of months emp. in last 2 years	0.03***	(0.01)
× French speaking REA	-0.03	(0.02)
× Italian speaking REA	-0.05*	(0.02)
Past income (in 10,000 CHF)	0.09***	(0.01)
Prev. job in primary sector	-0.04***	(0.01)
Prev. job in secondary sector	0.04***	(0.01)
Prev. job in tertiary sector	0.01**	(0.01)
Prev. job self-employed	-0.09***	(0.02)
Prev. job manager	-0.05***	(0.01)
Prev. job skilled worker	-0.02**	(0.01)
Prev. job unskilled worker	-0.02**	(0.01)
Native language not German, French, or Italian	-0.01**	(0.01)
× French speaking REA	-0.03**	(0.01)
× Italian speaking REA	-0.01	(0.02)
Married	0.002	(0.003)
Foreigner with temporary residence permit	-0.02***	(0.01)
Foreigner with permanent residence permit	0.002	(0.004)
Foreigner with mother tongue similar to canton's language	0.03***	(0.004)
Lives in big city	-0.01	(0.01)
Lives in medium sized city	0.02***	(0.01)
Start JSP participation in second unemp. quarter	0.02***	(0.004)

Table continues on next page >

Table D.1 continued.

	Av. Marg. Eff. (1)	S.E.
Caseworker characteristics		
Female	0.02**	(0.01)
× French speaking REA	-0.05*	(0.03)
× Italian speaking REA	0.05*	(0.02)
Age (in 10 years)	0.003	(0.004)
× French speaking REA	0.001	(0.001)
× Italian speaking REA	0.001	(0.001)
Tenure (in years)	0.002	(0.001)
× French speaking REA	-0.01	(0.01)
× Italian speaking REA	0.003	(0.004)
Own unemp. experience	0.01	(0.01)
× French speaking REA	-0.03	(0.03)
× Italian speaking REA	0.05	(0.03)
Education above vocational training	0.0001	(0.01)
× French speaking REA	0.01	(0.03)
× Italian speaking REA	0.01	(0.03)
Education tertiary track	0.002	(0.01)
× French speaking REA	-0.02	(0.04)
× Italian speaking REA	0.01	(0.04)
Vocational training degree	0.02*	(0.01)
× French speaking REA	-0.09	(0.07)
× Italian speaking REA	0.02	(0.03)
Indicator for missing caseworker characteristics	0.002	(0.02)
Allocation of unemployed to caseworkers		
By industry	0.02***	(0.01)
× French speaking REA	0.05*	(0.03)
× Italian speaking REA	-0.04*	(0.02)
By occupation	0.02**	(0.01)
× French speaking REA	0.04*	(0.03)
× Italian speaking REA	-0.06**	(0.03)
By age	0.01	(0.01)
By employability	-0.02	(0.01)
By region	-0.04**	(0.01)
Other allocation type	-0.03**	(0.01)
Local labour market characteristics		
French speaking REA	-0.06	(0.09)
Italian speaking REA	-0.19	(0.11)
Cantonal unemployment rate (in %)	0.03***	(0.01)
× French speaking REA	-0.07***	(0.01)
× Italian speaking REA	-0.03*	(0.02)
Cantonal GDP per capita (in 10,000 CHF)	-0.03***	(0.01)
# of caseworker	1,282	
# of observations	85,198	

Note: The estimation is based on a Probit model with the outcome JSP participation. The Probit model includes a constant term. We obtain standard errors (S.E.) from a clustered bootstrap at the caseworker level with 4,999 replications. *, **, *** mean statistically different from zero at the 10%, 5%, 1% level, respectively. REA is the abbreviation for regional employment agency. For many variables we include interactions with a dummy for French (× French speaking REA) and Italian (× Italian speaking REA) speaking regional employment agencies. To account for categorical variables, we omit the dummies some qualification degree, employability rating high, lives in small city, and German speaking regional employment agencies.

Table D.2: Balance of confounders after IPW.

	Participants	Non-Participants	Std. Diff.
	Mean	Mean	
	(1)	(2)	(3)
Characteristics of unemployed persons			
Female	0.44	0.44	0.21
× French speaking REA	0.08	0.08	0.01
× Italian speaking REA	0.03	0.03	0.66
Age (in 10 years)	3.65	3.67	1.57
$Age^2/10{,}000$	0.14	0.14	1.59
Unskilled	0.24	0.24	0.31
× French speaking REA	0.05	0.05	0.56
× Italian speaking REA	0.02	0.02	0.44
Semi-skilled qualification	0.16	0.16	0.27
× French speaking REA	0.03	0.04	1.43
× Italian speaking REA	0.01	0.01	0.45
Skilled qualification without degree	0.04	0.04	1.42
× French speaking REA	0.01	0.01	1.30
× Italian speaking REA	0.01	0.01	1.46
Employability rating low	0.15	0.14	1.00
× French speaking REA	0.01	0.01	0.06
× Italian speaking REA	0.01	0.01	0.03
Employability rating medium	0.75	0.75	0.02
× French speaking REA	0.14	0.15	0.70
× Italian speaking REA	0.04	0.04	0.20
# of unemp. spells in last 2 years	0.59	0.57	1.01
× French speaking REA	0.12	0.11	0.36
× Italian speaking REA	0.05	0.05	0.68
# of emp. spells in last 5 years	0.12	0.12	1.19
Fraction of months emp. in last 2 years	0.80	0.80	0.48
× French speaking REA	0.13	0.13	0.83
× Italian speaking REA	0.05	0.05	1.23
Past income (in 10,000 CHF)	4.21	4.24	0.85
Prev. job in primary sector	0.08	0.08	0.45
Prev. job in secondary sector	0.14	0.14	0.17
Prev. job in tertiary sector	0.59	0.60	0.72
Prev. job self-employed	0.01	0.01	0.40
Prev. job manager	0.07	0.07	0.10
Prev. job skilled worker	0.59	0.60	1.19
Prev. job unskilled worker	0.30	0.30	0.89

Table continues on next page >

Table D.2 continued.

	Participants	Non-Participants	Std. Diff.
	Mean	Mean	
	(1)	(2)	(3)
Characteristics of unemployed persons			
Native language not German, French, or Italian	0.32	0.32	0.23
× French speaking REA	0.05	0.05	0.50
× Italian speaking REA	0.02	0.02	0.30
Married	0.48	0.48	0.36
Foreigner with temporary residence permit	0.13	0.13	0.06
Foreigner with permanent residence permit	0.25	0.25	0.61
Foreigner with mother tongue similar to canton's language	0.11	0.11	0.10
Lives in big city	0.16	0.17	0.98
Lives in medium sized city	0.15	0.14	1.09
Start of JSP participation in the second unemp. quarter	0.43	0.45	1.72
Caseworker characteristics			
Female	0.41	0.41	0.35
× French speaking REA	0.05	0.05	0.15
× Italian speaking REA	0.02	0.02	0.15
Age (in 10 years)	4.43	4.43	0.24
× French speaking REA	0.79	0.80	0.50
× Italian speaking REA	0.27	0.29	0.92
Tenure (in years)	5.77	5.75	0.29
× French speaking REA	1.09	1.09	0.22
× Italian speaking REA	0.48	0.52	1.26
Own unemp. experience	0.63	0.63	0.36
× French speaking REA	0.11	0.11	0.46
× Italian speaking REA	0.04	0.04	0.71
Education above vocational training	0.44	0.45	0.95
× French speaking REA	0.08	0.08	0.50
× Italian speaking REA	0.02	0.02	1.73
Education tertiary track	0.23	0.22	1.41
× French speaking REA	0.06	0.06	2.44
× Italian speaking REA	0.01	0.01	0.18
Vocational training degree	0.23	0.24	0.40
× French speaking REA	0.01	0.01	0.47
× Italian speaking REA	0.03	0.04	1.36
Indicator for missing caseworker characteristics	0.04	0.04	0.24

Table continues on next page >

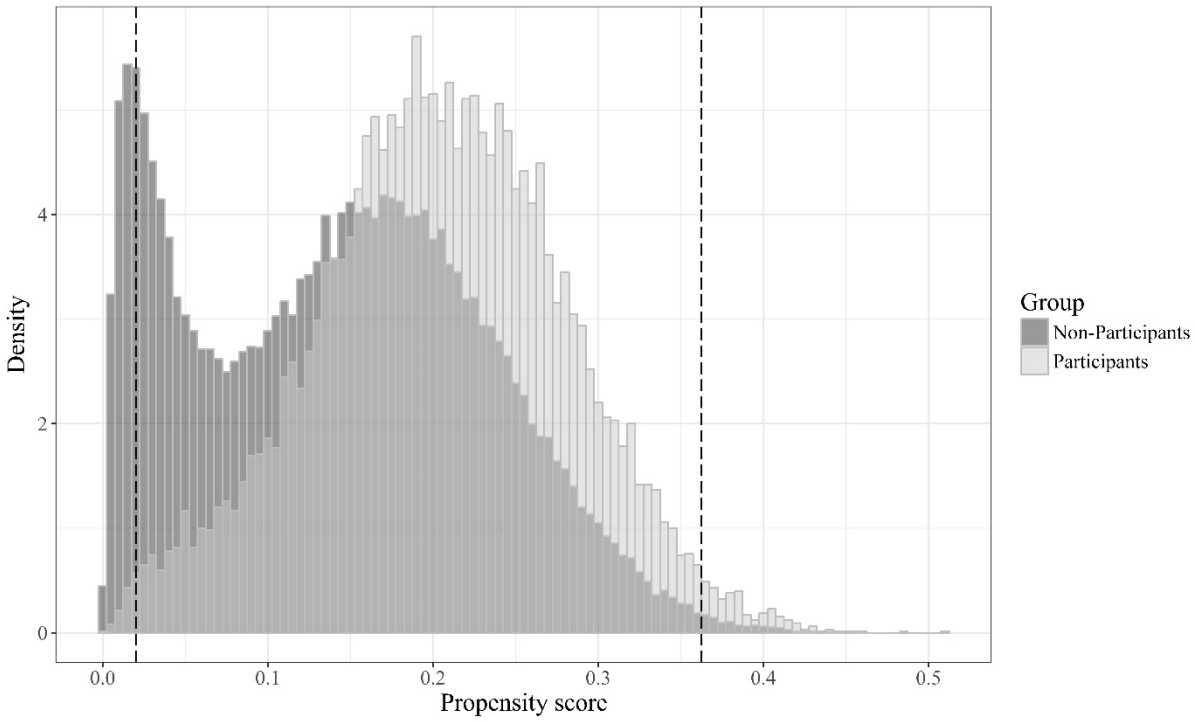
Table D.2 continued.

	Participants	Non-Participants	Std. Diff.
	Mean	Mean	
		(2)	(3)
Allocation of unemployed to caseworkers			
By industry	0.58	0.58	0.56
× French speaking REA	0.09	0.09	0.54
× Italian speaking REA	0.03	0.03	1.16
By occupation	0.56	0.56	0.01
× French speaking REA	0.13	0.13	0.13
× Italian speaking REA	0.04	0.04	1.19
By age	0.03	0.03	0.53
By employability	0.08	0.07	1.23
By region	0.11	0.11	1.31
Other allocation type	0.08	0.07	1.89
Local labour market characteristics			
French speaking REA	0.17	0.18	0.41
Italian speaking REA	0.06	0.07	1.16
Cantonal unemployment rate (in %)	3.68	3.68	0.29
× French speaking REA	0.69	0.71	1.02
× Italian speaking REA	0.27	0.29	1.10
Cantonal GDP per capita (in 10,000 CHF)	4.98	4.98	0.18
# of trimmed observations	582	6,118	
# of observations after trimming	12,712	65,786	

Note: We report IPW re-weighted means for all variables and standardised differences between participants and non-participants. Rosenbaum and Rubin (1983) consider a standardised difference of more than 20 as being 'large'. REA is the abbreviation for regional employment agency. For many variables we include interactions with a dummy for French (× French speaking REA) and Italian (× Italian speaking REA) speaking regional employment agencies. To account for categorical variables, we omit the dummies some qualification degree, employability rating high, lives in small city, and German speaking regional employment agencies.

Figure D.1 plots the distribution of the propensity score by participation status. We enforce common support by trimming observations below the 0.5 quantile of the participants and above the 99.5 quantile of the non-participants. In total, we trim 6,700 observations (582 participants, 6,118 non-participants). This procedure shows good final sample performance in the study Lechner and Strittmatter (2017).

Figure D.1: Histogram of the propensity score by participation status.



Note: The histogram has a binwidth of 0.005. The dashed lines show the lower and upper threshold of trimming.

Appendix E: Additional results

This Appendix complements the results shown in section 5.3 and 5.4 of the main text by providing results for additional outcome variables.

Table E.1: Post-LASSO coefficients of selected variables for the employment outcome 31 months after start participation.

	Months employed during first 31 months after the start of participation	
	Coef.	S.E.
	(1)	
Constant	-1.37***	(0.32)
Female × CW education above vocational training	0.07	(0.54)
Unskilled × CW education above vocational training	-1.24	(1.56)
Unskilled × prev. job unskilled	4.66***	(1.58)
# of unemp. spells in last 2 years × unemp. person and CW have primary education	0.29	(0.20)
Fraction of months emp. in last 2 years × past income 57 - 75k	0.32	(0.63)
GDP per capita × prev. job self-employed	4.56	(5.60)
CW education: above vocational training × past income 25 - 50k	1.09*	(0.66)
CW education: tertiary track × past income 25 - 50k	0.21	(0.84)
Degree in vocational training for caseworkers × past income 50 - 75k	-0.70	(0.81)
Married × past income 50 - 75k	-0.37	(0.70)
Foreigner with permanent residence permit × past income 50 - 75k	0.32	(0.76)
Medium city × prev. job unskilled	-0.48	(0.94)
Single household × no emp. spell last 2 years	-0.24	(0.56)
Past income 0 - 25k × # emp. spells past 5 years	3.60	(2.47)
# emp. spells past 5 years × unemp. person and CW have primary education	-1.41	(1.87)
# emp. spells past 5 years × unemp. person and CW have same gender, age, and education	-6.34	(5.93)
No emp. spell last 2 years × skilled worker	-0.56	(0.53)
Prev. job in primary sector × unskilled	-0.20	(1.30)
Unskilled × unemp. person and CW have primary education	0.09	(0.56)
Regional emp. agency No. 44	-0.98	(1.27)
# of selected variables	20 of 1,268	

Note: We apply one-step efficiency augmentation. We partition the data randomly into selection and estimation samples. We choose the penalty term based on Post-LASSO RMSE, which we optimise with 10-fold cross-validation. We obtain standard errors (S.E.) from a clustered bootstrap at the caseworker level with 4,999 replications. *, **, *** mean statistically different from zero at the 10%, 5%, 1% level, respectively. CW is the abbreviation for caseworker. 25 - 50k is the abbreviation for 25,000-50,000 CHF. 50 – 75k is the abbreviation for 50,000-75,000 CHF. We omit the results for the outcome cumulated employment between months 25-31 after start of JSP participation, because we do not identify any heterogeneity variables for this outcome in the random sample spilt we consider.

Table E.2: Differences between AITEs on cumulated employment during the first 6 months after the start JSP participation by characteristics of unemployed persons.

	Difference	S.E.
	(1)	(2)
Unskilled	0.26***	(0.03)
Some degree	-0.25***	(0.02)
# of unemployment spells last 2 years	0.24***	(0.03)
Employability rating low	0.23***	(0.03)
Skilled qualification w/o degree	0.21***	(0.05)
Past income	-0.19***	(0.02)
Foreigner with temporary permit	0.12***	(0.03)
Foreigner with permanent permit	0.12***	(0.02)
Employability rating high	-0.11***	(0.01)
Employability rating medium	-0.10***	(0.02)
Fraction employed last year	-0.09***	(0.01)
Semi-skilled qualification	0.05**	(0.02)
Mother tongue of canton	0.04***	(0.01)
Age	0.03***	(0.01)
Female	0.02	(0.02)

Note: This table reports the differences between AITE by low and high values of the respective characteristic of unemployed persons (see also Figure 3). A low value is zero when the variable is binary or below the median when the variable is non-binary. A high value is one when the variable is binary or not below the median when the variable is non-binary. The AITEs are based on 30 random sample splits. For each partition, we choose the penalty term based on Post-LASSO RMSE, which we optimise with 10-fold cross-validation. We apply one-step efficiency augmentation. We report standard errors based on 1,000 bootstrap replications. *, **, *** mean statistically different from zero at the 10%, 5%, 1% level, respectively.

Table E.3: Differences between AITEs on cumulated employment during the first 6 months after start JSP participation by caseworker characteristics.

	Difference	S.E.
	(1)	(2)
CW own unemployment experience	0.07***	(0.02)
CW & UE tertiary education	0.06***	(0.02)
CW special training	-0.03*	(0.01)
CW & UE upper secondary education	0.02*	(0.01)
CW & UE primary education	0.02**	(0.01)
CW & UE age difference	-0.02	(0.01)
CW education tertiary track	-0.01	(0.02)
CW age	-0.01	(0.01)
Female	-0.01	(0.01)
CW & UE secondary education	0.01	(0.01)
CW & UE same gender	0.01	(0.01)
CW cooperative	0.01	(0.01)
CW & UE same gender, age, and education	0.01	(0.02)
CW & UE same age ± 5 years	-0.01	(0.01)
CW education above vocational training	0.00	(0.02)
CW tenure	0.00	(0.01)

Note: This table reports the differences between AITE by low and high values of the respective characteristic of unemployed persons (see also Figure 4). A low value is zero when the variable is binary or below the median when the variable is non-binary. A high value is one when the variable is binary or not below the median when the variable is non-binary. The AITEs are based on 30 random sample splits. For each partition, we choose the penalty term based on Post-LASSO RMSE, which we optimise with 10-fold cross-validation. We apply one-step efficiency augmentation. We report standard errors based on 1,000 bootstrap replications. *, **, *** mean statistically different from zero at the 10%, 5%, 1% level, respectively.

Table E.4: Average number of times a specific variable is selected over thirty sample splits.

Past income	8.0
Qualification	5.9
Past unemployment	5.8
Previous job unskilled worker	4.4
Employability	3.9
Single household	2.8
CW own unemp. Experience	2.6
Size of city	2.4
Foreigner with B permit	2.3
RAV dummies	2.2
Age caseworker - age unemployed	2.0
Previous job in primary sector	1.9
Foreigner with C permit	1.8
Previous industry	1.7
Same gender	1.6
Previous job in tertiary sector	1.6
Female	1.6
Previous occupation	1.5
CW Education: above vocational training	1.4
Same age ± 5 years	1.2
CW female	1.1
CW vocational training degree	1.1
CW tertiary track	1.1
Mother tongue other than German, French, Italian	1.0
Both primary education	1.0
Both upper secondary education	1.0
Both secondary education	1.0
Both tertiary education	0.9
Previous job skilled worker	0.9
Never employed in last two years	0.8
CW cooperative	0.8
Number of employment spells in last 5 years	0.7
Previous job in secondary sector	0.6
Foreigner with mother tongue in canton's language	0.5
Previous job manager	0.5
CW tenure	0.4
Married	0.4
Number of people in household	0.2
Previous job self-employed	0.2
Same gender, age, and education	0.2
Fraction of months emp. in last 2 years	0.2
CW age	0.1
UE age	0.1

Note: Number of selected variables in the 30 sample splits of MCM with one-step efficiency augmentation and Post-LASSO.

Table E.5: K-means clustering. K=3

Cluster	1	2	3
Number of observations in cluster	56328	21133	1037
AITE for employment during first 6 months	-0.89	-0.52	-0.42
AITE for employment during first 12 months	-1.24	-0.74	-0.48
AITE for employment during first 31 months	-1.38	-0.65	2.02
AITE for employment during months 25-31	-0.06	-0.02	0.28
Female	0.42	0.50	0.49
Past income	0.47	0.31	0.30
Fraction of months emp. in last 2 years	0.83	0.74	0.73
# of unemp. spells in last 2 years	0.27	1.36	0.94
Unskilled	0.14	0.51	0.06
Qualification: semiskilled	0.15	0.19	0.01
Qualification: skilled without degree	0.01	0.06	0.91
Some qualification degree	0.69	0.24	0.02
Foreigner with mother tongue in canton's language	0.11	0.12	0.13
Employability rating low	0.11	0.23	0.21
Employability rating medium	0.77	0.71	0.74
Employability rating high	0.12	0.06	0.05
Age in 10 year	3.67	3.67	3.69
Foreigner with B permit	0.08	0.26	0.41
Foreigner with C permit	0.22	0.32	0.35
CW age	44.29	44.46	45.21
Female	0.42	0.38	0.36
CW tenure	5.75	5.78	5.58
Own unemp. experience	0.59	0.72	0.66
Education: above vocational training	0.39	0.59	0.47
Education: tertiary track	0.24	0.16	0.20
Vocational training degree	0.24	0.22	0.24
CW cooperative	0.49	0.48	0.44
Same gender	0.58	0.59	0.59
Age caseworker - age unemployed	8.55	8.68	9.08
Same age ± 5 years	0.23	0.23	0.22
Both primary education	0.72	0.69	0.66
Both secondary education	0.32	0.26	0.28
Both upper secondary education	0.44	0.32	0.36
Both tertiary education	0.53	0.63	0.58
Same gender. age. and education	0.04	0.03	0.03

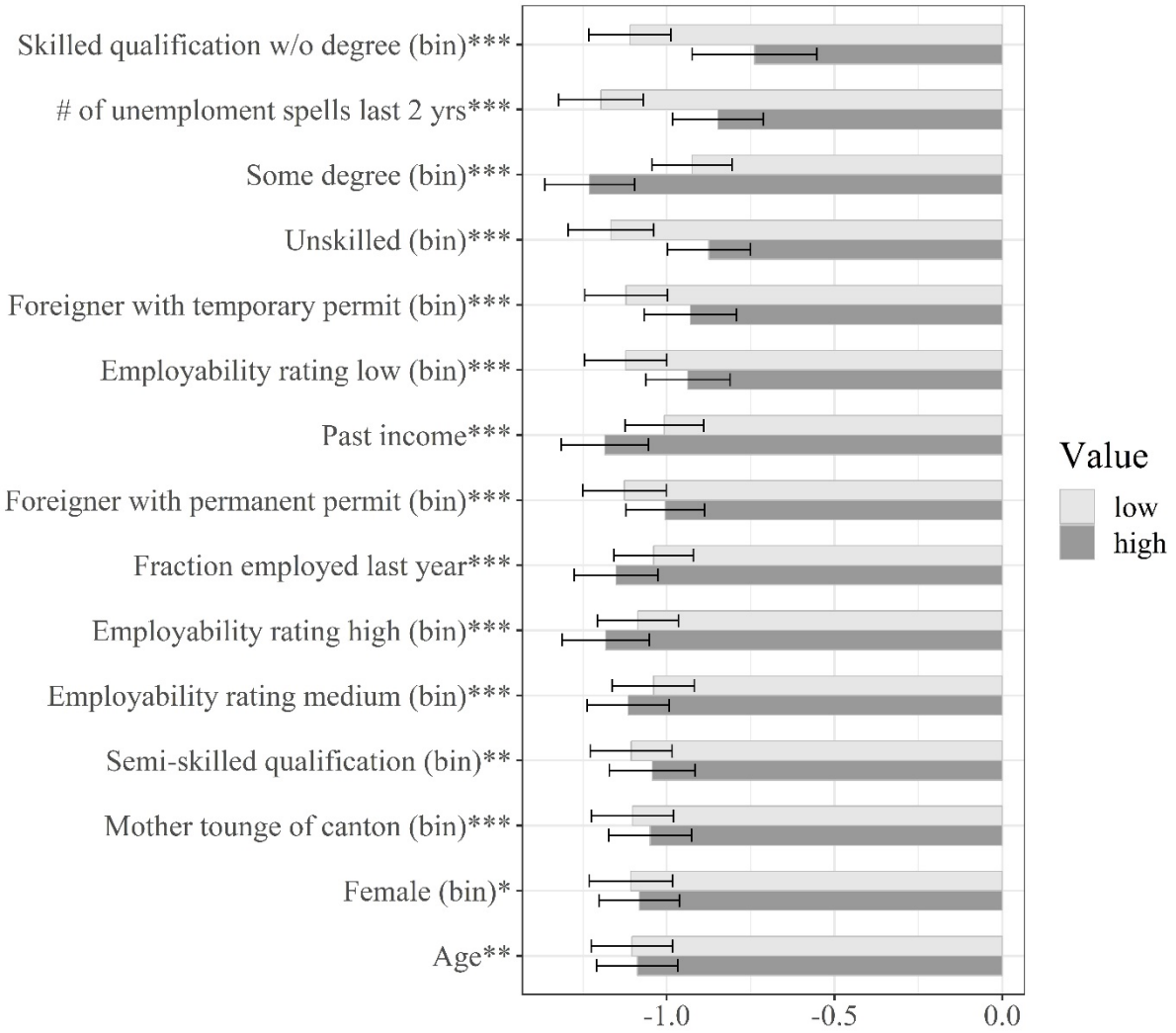
Note: Joint k-means clustering of AITEs of the four outcomes. Lower panel shows means of covariates.

Table E.6: K-means clustering. K=5

Cluster	1	2	3	4	5
Number of observations in cluster	44327	3878	19031	10295	967
AITE for employment during first 6 months	-0.92	-0.89	-0.63	-0.45	-0.42
AITE for employment during first 12 months	-1.27	-1.35	-0.92	-0.62	-0.48
AITE for employment during first 31 months	-1.32	-2.43	-1.01	-0.37	2.11
AITE for employment during months 25-31	-0.05	-0.16	-0.04	0.00	0.29
Female	0.44	0.13	0.49	0.51	0.49
Past income	0.49	0.46	0.36	0.26	0.31
Fraction of months emp. in last 2 years	0.83	0.84	0.78	0.71	0.75
# of unemp. spells in last 2 years	0.14	0.60	0.98	1.61	0.89
Unskilled	0.09	0.23	0.43	0.56	0.02
Qualification: semiskilled	0.13	0.21	0.21	0.17	0.01
Qualification: skilled without degree	0.01	0.02	0.05	0.05	0.97
Some qualification degree	0.77	0.53	0.31	0.21	0.01
Foreigner with mother tongue in canton's language	0.10	0.11	0.13	0.12	0.13
Employability rating low	0.08	0.15	0.22	0.24	0.20
Employability rating medium	0.79	0.76	0.71	0.70	0.74
Employability rating high	0.13	0.10	0.07	0.05	0.05
Age in 10 year	3.65	3.82	3.69	3.65	3.69
Foreigner with B permit	0.07	0.08	0.17	0.32	0.38
Foreigner with C permit	0.18	0.31	0.34	0.30	0.37
CW age	44.30	44.40	44.27	44.57	45.30
Female	0.43	0.30	0.40	0.37	0.37
CW tenure	5.71	5.82	5.85	5.75	5.61
Own unemp. Experience	0.59	0.55	0.65	0.78	0.64
Education: above vocational training	0.42	0.29	0.43	0.66	0.45
Education: tertiary track	0.23	0.29	0.21	0.14	0.20
Vocational training degree	0.24	0.26	0.23	0.21	0.24
CW cooperative	0.48	0.48	0.50	0.48	0.43
Same gender	0.58	0.68	0.58	0.59	0.59
Age caseworker - age unemployed	8.72	7.52	8.29	8.95	9.15
Same age ± 5 years	0.23	0.24	0.23	0.23	0.22
Both primary education	0.71	0.79	0.71	0.68	0.67
Both secondary education	0.31	0.40	0.31	0.24	0.28
Both upper secondary education	0.40	0.63	0.44	0.27	0.38
Both tertiary education	0.52	0.60	0.60	0.64	0.58
Same gender, age, and education	0.03	0.06	0.04	0.02	0.03

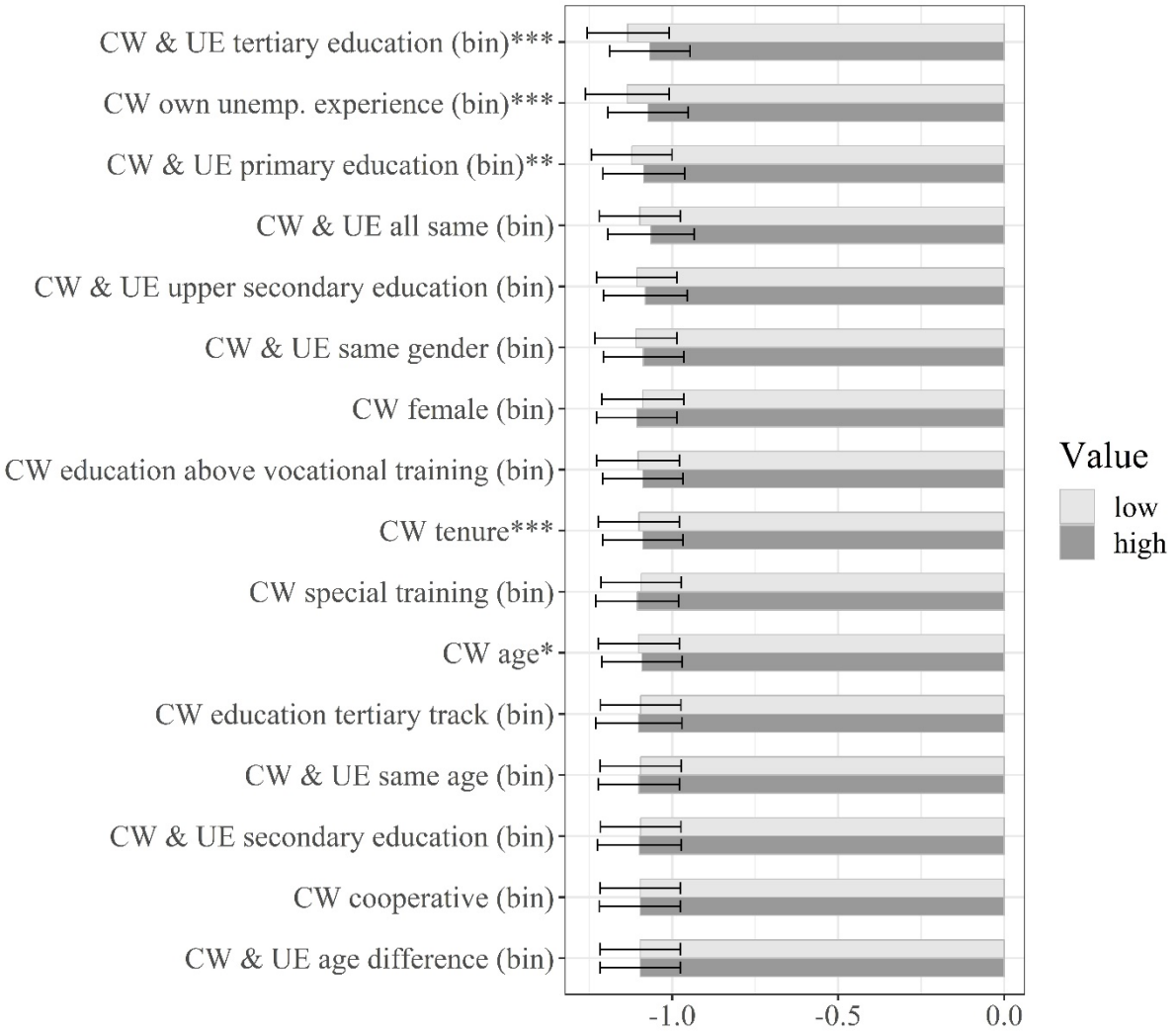
Note: Joint k-means clustering of AITEs of the four outcomes. Lower panel shows means of covariates.

Figure E.1: AITE on cumulated employment during the first 12 months after the start of JSP participation by characteristics of unemployed persons.



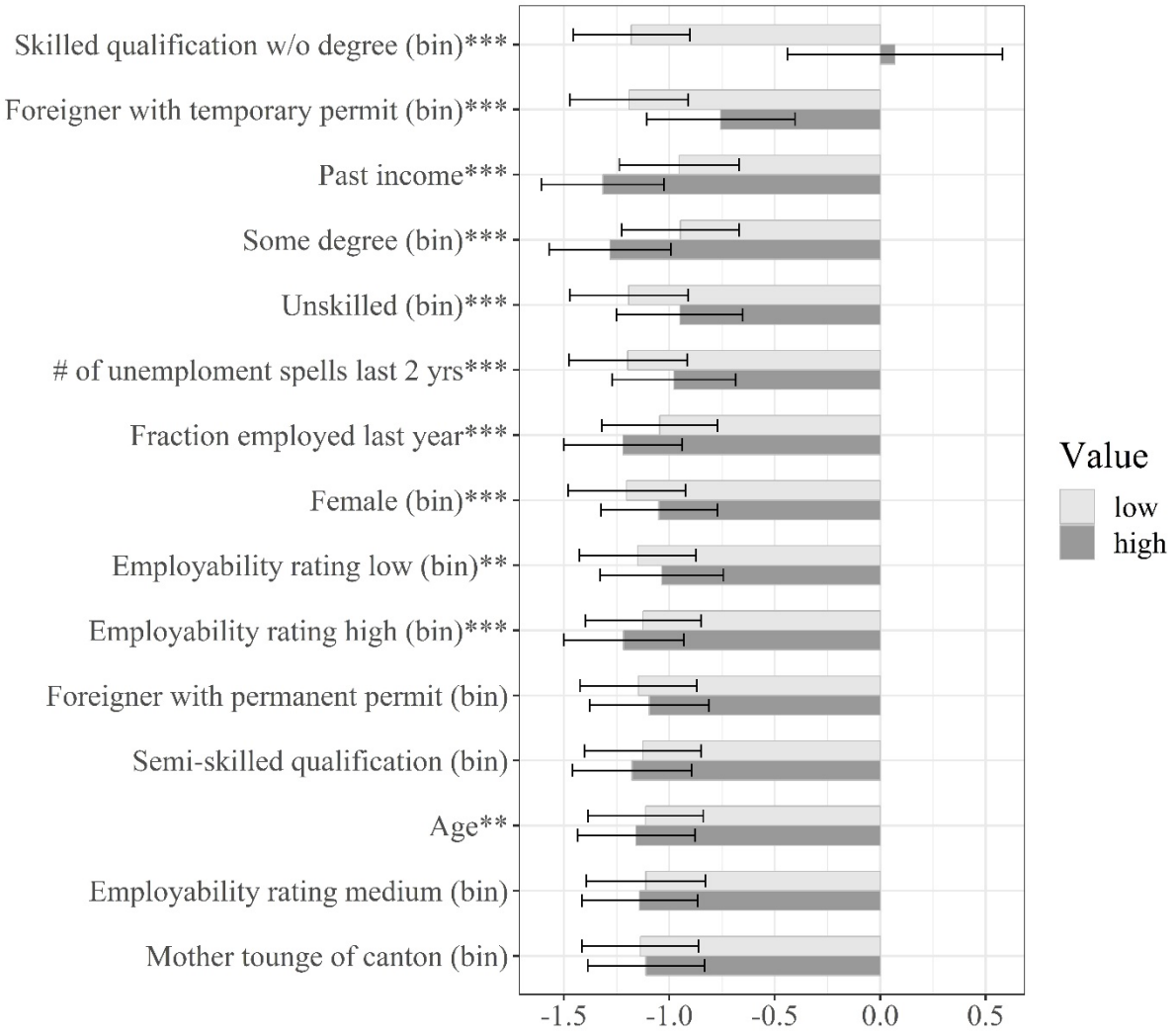
Note: AITE by low and high values of the respective characteristic of unemployed persons. A low value is zero when the variable is binary or below the median when the variable is non-binary. A high value is one when the variable is binary or not below the median when the variable is non-binary. We aggregate the CATEs over 30 random sample splits. For each partition, we choose the penalty term based on Post-LASSO RMSE, which we optimise with 10-fold cross-validation. We apply one-step efficiency augmentation. We report the 95%-confidence interval based on the bootstrap procedure described in section 4.6. *, **, *** mean statistically different from zero at the 10%, 5%, 1% level, respectively.

Figure E.2: AITE on cumulated employment during the first 12 months after the start of JSP participation by caseworker characteristics.



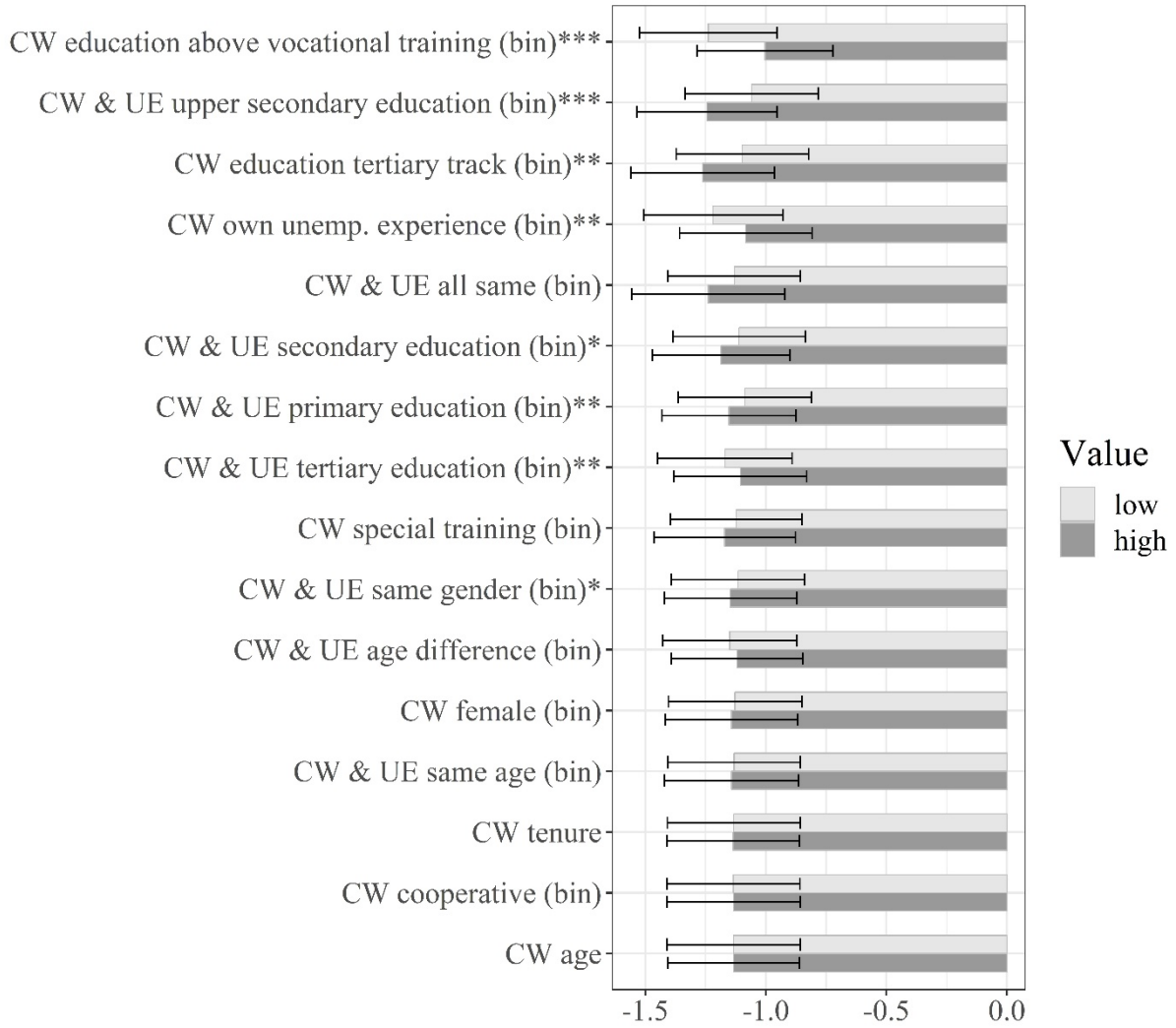
Note: AITE by low and high values of the respective caseworker characteristic. A low value is zero when the variable is binary or below the median when the variable is non-binary. A high value is one when the variable is binary or not below the median when the variable is non-binary. We aggregate the CATEs over 30 random sample splits. For each partition, we choose the penalty term based on Post-LASSO RMSE, which we optimise with 10-fold cross-validation. We apply one-step efficiency augmentation. We report the 95%-confidence interval based on the bootstrap procedure described in section 4.6. *, **, *** mean statistically different from zero at the 10%, 5%, 1% level, respectively. CW is the abbreviation for caseworker.

Figure E.3: AITE on cumulated employment during the first 31 months after the start of JSP participation by characteristics of unemployed persons.



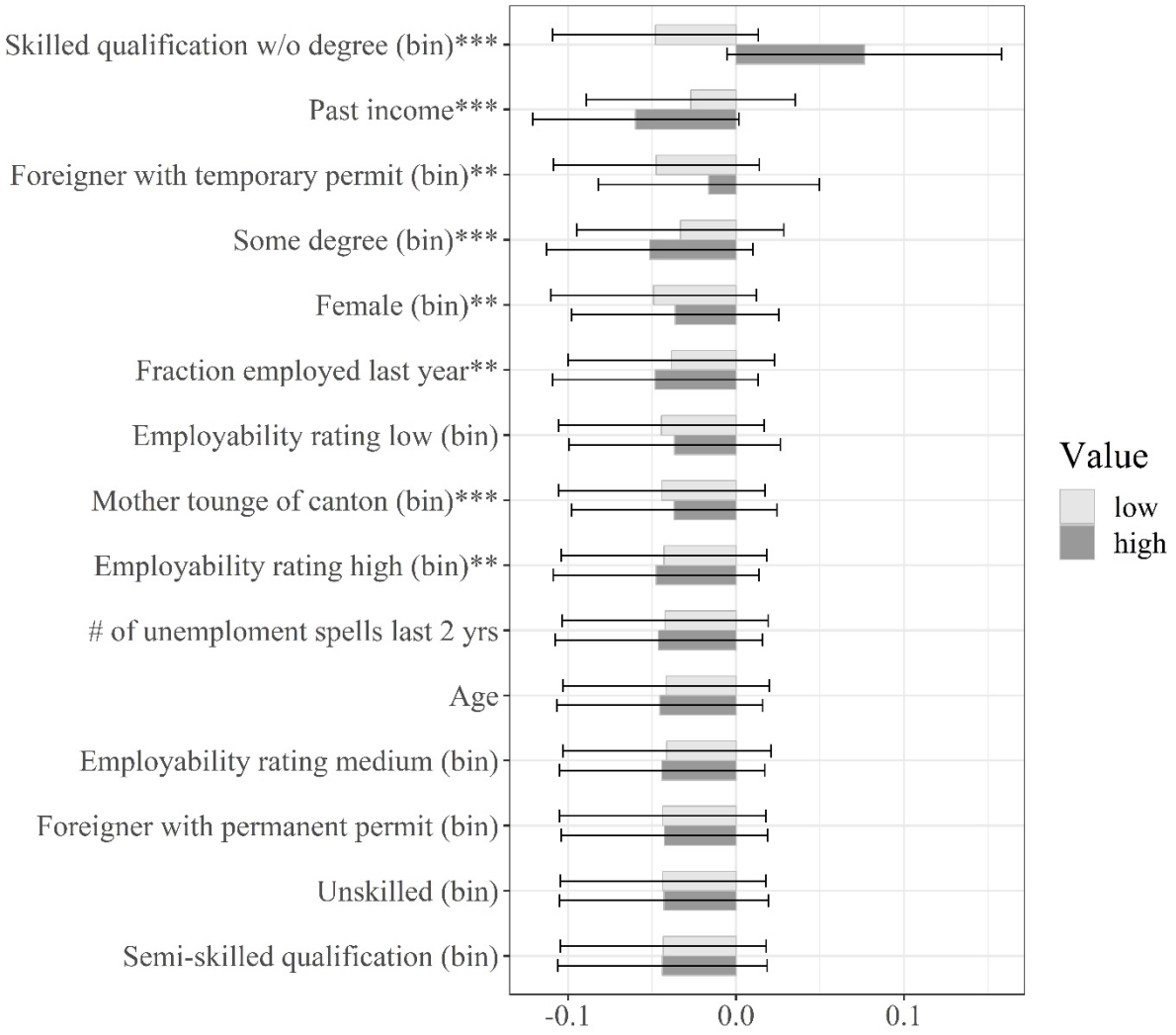
Note: AITE by low and high values of the respective characteristic of unemployed persons. A low value is zero when the variable is binary or below the median when the variable is non-binary. A high value is one when the variable is binary or not below the median when the variable is non-binary. We aggregate the CATEs over 30 random sample splits. For each partition, we choose the penalty term based on Post-LASSO RMSE, which we optimise with 10-fold cross-validation. We report the 95%-confidence interval based on the bootstrap procedure described in section 4.6. *, **, *** mean statistically different from zero at the 10%, 5%, 1% level, respectively.

Figure E.4: AITE on cumulated employment during the first 31 months after the start of JSP participation by caseworker characteristics.



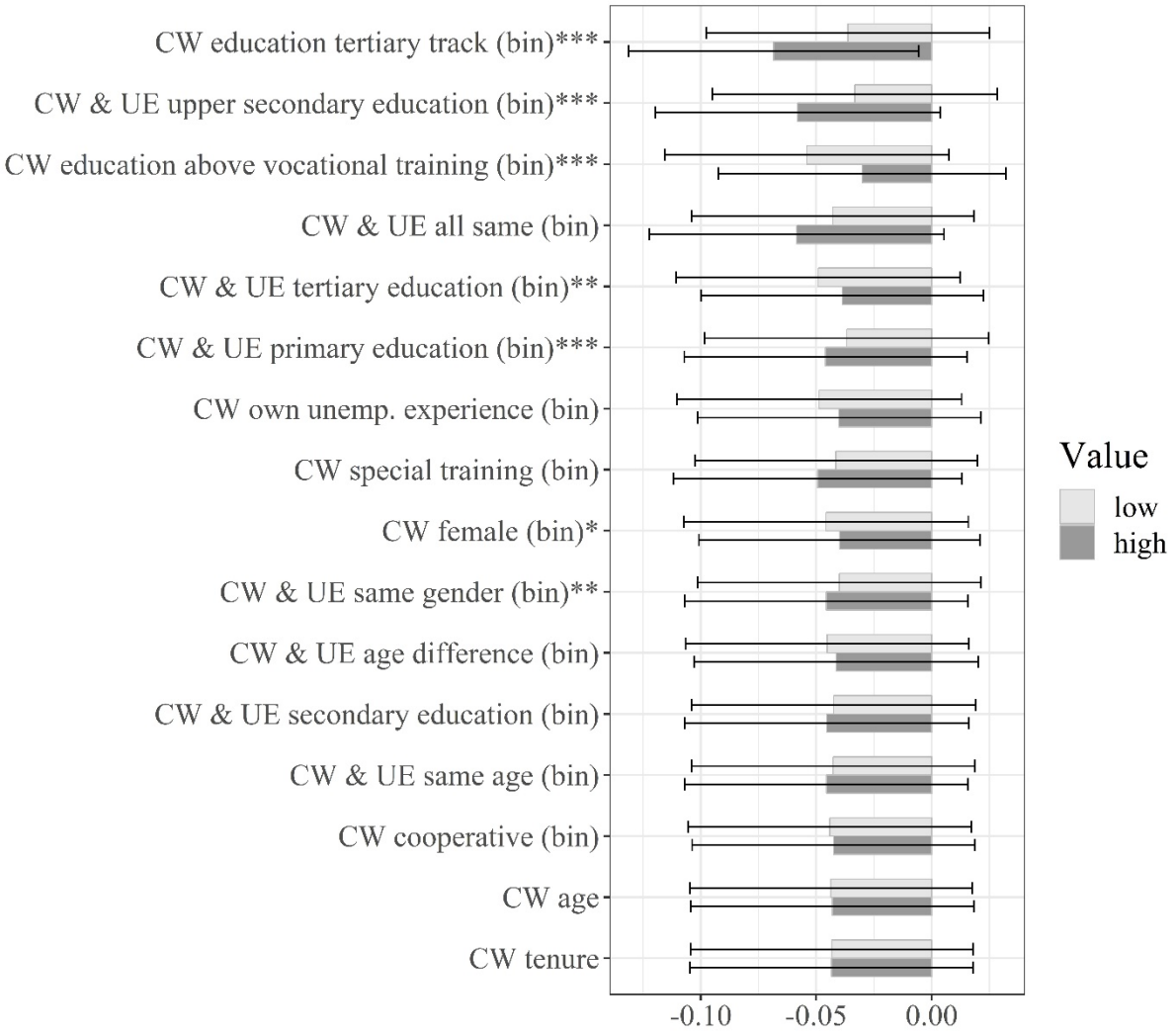
Note: CATE by low and high values of the respective caseworker characteristic. A low value is zero when the variable is binary or below the median when the variable is non-binary. A high value is one when the variable is binary or not below the median when the variable is non-binary. We aggregate the CATEs over 30 random sample splits. For each partition, we choose the penalty term based on Post-LASSO RMSE, which we optimise with 10-fold cross-validation.. We apply one-step efficiency augmentation. We report the 95%-confidence interval based on the bootstrap procedure described in section 4.6. *, **, *** mean statistically different from zero at the 10%, 5%, 1% level, respectively. CW is the abbreviation for caseworker.

Figure E.5: AITE on cumulated employment during the months 25-31 after the start of JSP participation by characteristics of unemployed persons.



Note: AITE by low and high values of the respective characteristic of unemployed persons. A low value is zero when the variable is binary or below the median when the variable is non-binary. A high value is one when the variable is binary or not below the median when the variable is non-binary. We aggregate the CATEs over 30 random sample splits. For each partition, we choose the penalty term based on Post-LASSO RMSE, which we optimise with 10-fold cross-validation. We apply one-step efficiency augmentation. We report the 95%-confidence interval based on the bootstrap procedure described in section 4.6. *, **, *** mean statistically different from zero at the 10%, 5%, 1% level, respectively.

Figure E.6: AITE on cumulated employment during the months 25-31 after the start of JSP participation by caseworker characteristics.



Note: AITE by low and high values of the respective caseworker characteristic. A low value is zero when the variable is binary or below the median when the variable is non-binary. A high value is one when the variable is binary or not below the median when the variable is non-binary. We aggregate the CATEs over 30 random sample splits. For each partition, we choose the penalty term based on Post-LASSO RMSE, which we optimise with 10-fold cross-validation. We apply one-step efficiency augmentation. We report the 95%-confidence interval based on the bootstrap procedure described in section 4.6. *, **, *** mean statistically different from zero at the 10%, 5%, 1% level, respectively. CW is the abbreviation for caseworker.

Figure E.7: Heterogeneous employment effects during months 20 to 25 by characteristics of unemployed.

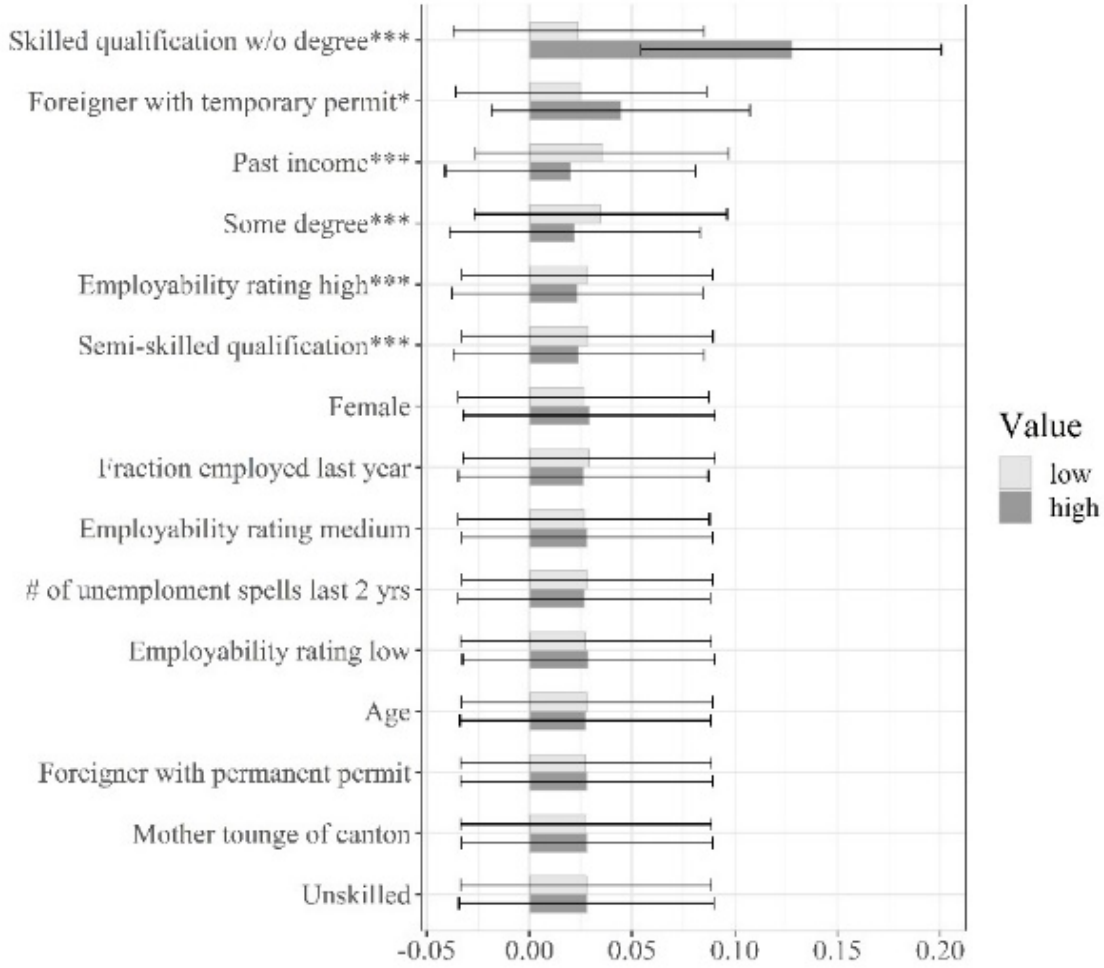


Figure E.8: Heterogeneous employment effects during months 20 to 25 by characteristics of caseworkers.

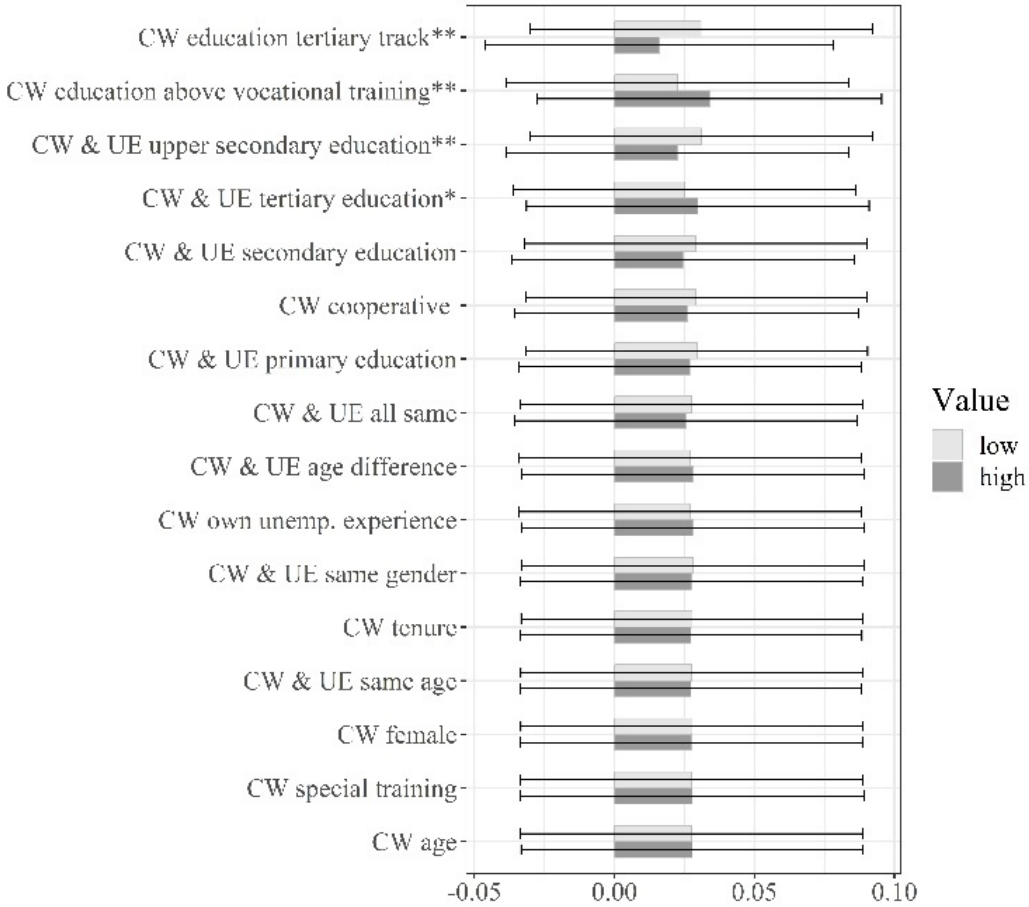


Figure E.9: Heterogeneous employment effects during months 20 to 31 by characteristics of unemployed.

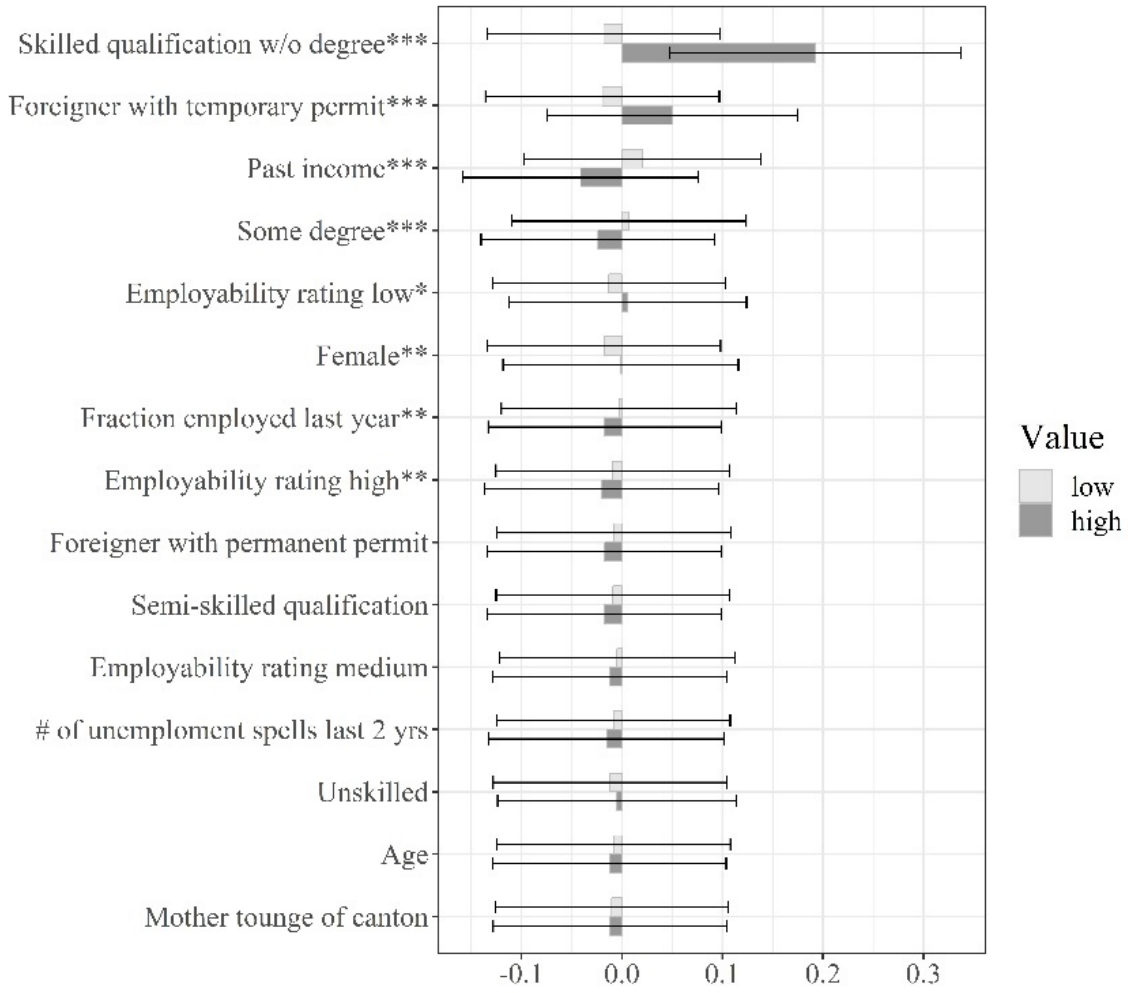
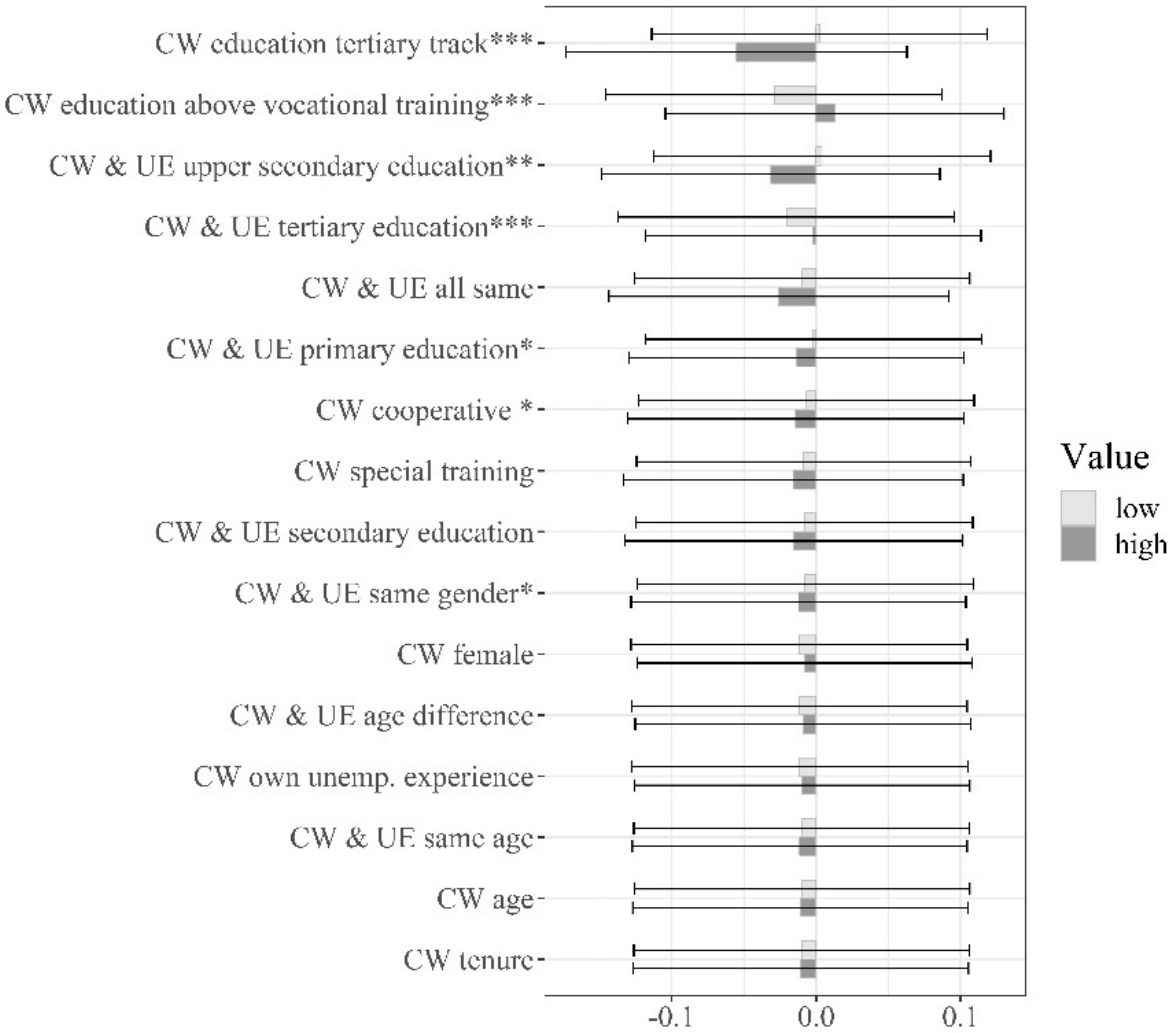


Figure E.10: Heterogeneous employment effects during months 20 to 31 by characteristics of caseworkers.



Appendix F: Robustness checks

F.1 Modified Outcome Method (MOM)

In the baseline estimations, we rely on the MCM, because it offers the possibility of efficiency augmentation described below. An alternative approach is the Modified Outcome Method (MOM), which modifies the outcome instead of the covariates. This procedure was proposed by Signorovich (2007) and extended to non-experimental studies by Zhang et al. (2012). We apply the MOM by minimising the objective function

$$\arg\min_{\hat{\delta}} \left[\sum_{i=1}^{N} \left(Y_i^* - Z_i \hat{\delta} \right)^2 + \lambda \sum_{j=1}^{p} \left| \hat{\delta}_j \right| \right],$$

where $Y_i^* = \hat{w}(D_i, X_i, Z_i) \cdot Y_i$ is the modified outcome.

F.2 Efficiency augmentation

Chen et al. (2017) propose two ways to account for the main effects, which might improve the efficiency of the selection procedure. First, the one-step procedure includes the main effects in the empirical model by solving

$$\arg\min_{\hat{\delta}, \hat{\beta}_t} \left[\sum_{i=1}^{N} \hat{w}(D_i, X_i, Z_i) T_i \left(Y_i - Z_i \hat{\beta}_t - \frac{Z_i T_i}{2} \hat{\delta} \right)^2 + \lambda \sum_{j=1}^{p} \left(\left| \hat{\beta}_{tj} \right| + \left| \hat{\delta}_j \right| \right) \right].$$

This specification is strongly related to the approach of Imai and Ratkovic (2013), but they consider only experimental research designs and propose to use a combination of LASSO und Support Vector Machines.

Second, the two-step procedure estimates a WOLS including only the main effects in the first step. Afterwards, the estimated residuals $\hat{u}$ of this auxiliary regression are used as a regressand when selecting the interaction effects

$$\arg\min_{\hat{\delta}}\left[\sum_{i=1}^{N}\hat{w}(D_i,X_i,Z_i)T_i\left(\hat{u}_i-\frac{Z_iT_i}{2}\hat{\delta}\right)^2+\lambda\sum_{j=1}^{p}\left|\hat{\delta}_j\right|\right].$$

We consider the one-step procedure in the main specifications and show sensitivity checks with the two-step procedure.

F.3 Adaptive LASSO

In the main results, we rely on the standard LASSO estimator. A potential disadvantage of this estimator is the inability to penalize the coefficients differentially. The adaptive LASSO is an alternative estimator that has received a lot of attention in the literature (see Zou, 2006). One way of specifying the adaptive LASSO in high-dimensional settings, is to minimise the objective function

$$\arg\min_{\hat{\delta}}\left[\sum_{i=1}^{N}\hat{w}(D_i,X_i,Z_i)T_i\left(Y_i-\frac{Z_iT_i}{2}\hat{\delta}\right)^2+\sum_{j=1}^{p}\frac{\lambda}{\left|\hat{\beta}_j\right|}\left|\hat{\delta}_j\right|\right], \tag{1}$$

where we obtain $\hat{\beta}_j$ from a first step Ridge estimator minimising

$$\arg\min_{\hat{\beta}}\left[\sum_{i=1}^{N}\hat{w}(D_i,X_i,Z_i)T_i\left(Y_i-\frac{Z_iT_i}{2}\hat{\beta}\right)^2+\lambda\sum_{j=1}^{p}\hat{\beta}_j^2\right].$$

The Ridge estimator penalises the sum of squared coefficients instead of the sum of the absolute coefficients (Hoerl and Kennard, 1970). Therefore, Ridge estimators shrink the coefficients to zero, but they do not reach zero, unless the penalty parameter is infinity. Accordingly, Ridge estimators do not select models. The penalty term of the adaptive LASSO in equation (1) decreases with the absolute size of the Ridge coefficients. Accordingly, variables with small Ridge coefficients have a larger penalty term in the adaptive LASSO. Zou (2006) shows that the adaptive LASSO can achieve under appropriate assumptions the oracle property. The oracle property implies that the adaptive LASSO selects the correct model at an asymptotically

appropriate rate, such that the selection step can be neglected. Wang and Leng (2007) discuss the properties of the adaptive LASSO.

F.4 Causal forest

We implement the approach suggested by Wager and Athey (2017). It is based on combining the causal tree approach by Athey and Imbens (2016) with the idea of random forests. In other words, deep trees are built and effects are estimated within the resulting leaves. Tree building is based on maximising estimated heterogeneity and using a randomly selected subset of features at each possible sample split. Then, these individual predictions of the CATEs are averaged over many bootstrap samples. So far, experience with these causal random forests are limited which is the reason why we use them only for the robustness analysis.

F.5 Additional confounders

In our main specifications, we use the propensity score specification of Huber, Lechner, and Mellace (2017). We consider two additional sets of confounding variables to check the sensitivity of our results with respect to misspecification of the propensity score. First, we estimate a LASSO model on the treatment equation

$$\underset{\hat{\alpha},\hat{\beta}}{\operatorname{argmin}}\left[\sum_{i=1}^{N}\left(D_i - X_i\hat{\alpha} - Z_i\hat{\beta}\right)^2\right] + \lambda\sum_{j=1}^{p}\left|\hat{\beta}_j\right|,$$

where the confounders X_i are not penalised. We consider all variables of Z_i with non-zero LASSO coefficients $\hat{\beta}$ as additional confounders (we call them 'additional confounders 1' in the following). We denote the additional confounders 1 by Z_i^1.

Second, we estimate a LASSO model on the outcome equation

$$\underset{\hat{\alpha},\hat{\beta}}{\operatorname{argmin}}\left[\sum_{i=1}^{N}\left(Y_i - X_i\hat{\alpha} - Z_i^1\hat{\beta} - Z_i\hat{\gamma}\right)^2\right] + \lambda\sum_{j=1}^{p}\left|\hat{\gamma}_j\right|,$$

where the confounders X_i and Z_i^1 are not penalised. We consider all variables of Z_i with non-zero LASSO coefficients $\hat{\gamma}$ as additional confounders (we call them 'additional confounders 2' in the following). This procedure to select additional confounders is in the spirit of double selection (see Belloni, Chernozhukov, and Hansen, 2013).

F.6 Results of additional robustness checks

We estimate CATEs, $\hat{\gamma}_s(Z_i)$, using different random sample splits. Table F.1 reports the average correlations between CATEs obtained from the different random sample splits. The CATEs are positively correlated between the different random sample splits. The positive correlations are particularly high when we consider the employment outcome during the first six months after the start of participation. After longer time periods, the positive correlations decrease, but remain decently positive. This suggests that our results are not sensitive to a particular random sample split. This finding is robust across the 12 different estimation procedures we consider.

Tables F.2 to F.4 documents the correlations of AITEs, $\bar{\gamma}(Z_i)$, obtained across different estimation procedures. These tables are similar to Table 7, but consider different employment outcomes. The AITEs obtained from the alternative methods are highly positively correlated, no matter which procedure we use. The smallest correlations are found for the outcome cumulated employment during months 25 to 31 after the start of participation. For this outcome the correlations are substantially lower. This is not surprising, because very little heterogeneity is found for this outcome in general.

Table F.5 provides the descriptive statistics of all AITEs by different estimation procedures and outcome variables. The means are close to the respective ATEs, which is expected under the law of iterative expectations. This reassures us that all estimation procedures are able to replicate the semi-parametric IPW estimates.

Interestingly, the differences in the standard deviations indicate that some estimation procedures detect more heterogeneity than others. We observe three striking patterns: First, the two-step efficiency augmentation detects less heterogeneity than the one-step efficiency augmentation or no efficiency augmentation. Second, the adaptive LASSO without efficiency augmentation finds the most effect heterogeneity. Third, the procedure with the weights obtained from radius-matching tends to detect slightly less heterogeneity than the estimation procedures using IPW weights. Table F.6 describes the number of selected variables in the 30 random sample splits we consider. We observe large differences over different estimation procedures. The adaptive LASSO selects substantially more variables than Post-LASSO estimators. This could be an explanation for why the adaptive LASSO detects the most effect heterogeneity.

Figure F.1: Scatter plot of estimated AITEs with and without repeated propensity score estimations.

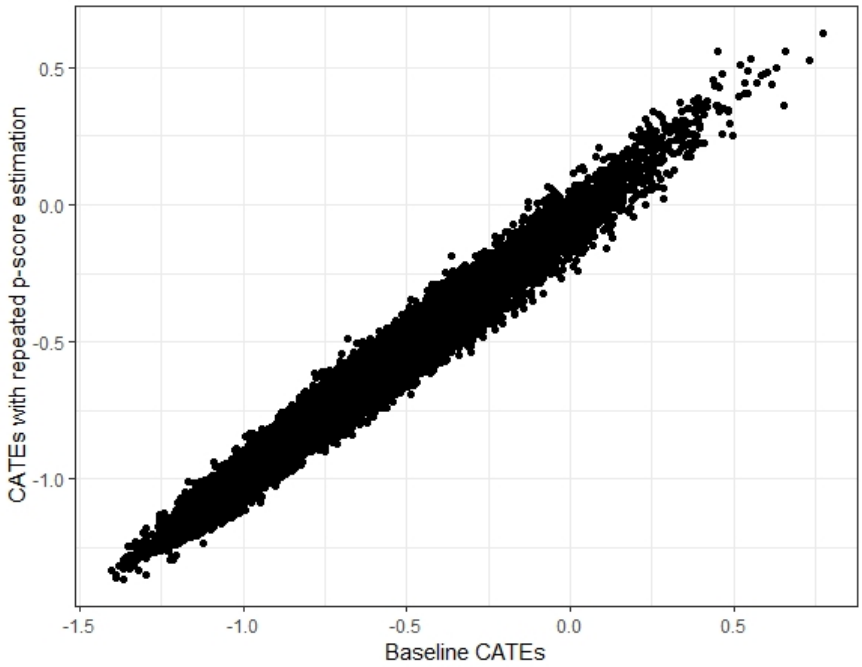


Figure F.2: Scatter plot of estimated AITEs with trimming on the 99.5% and 97.5% level.

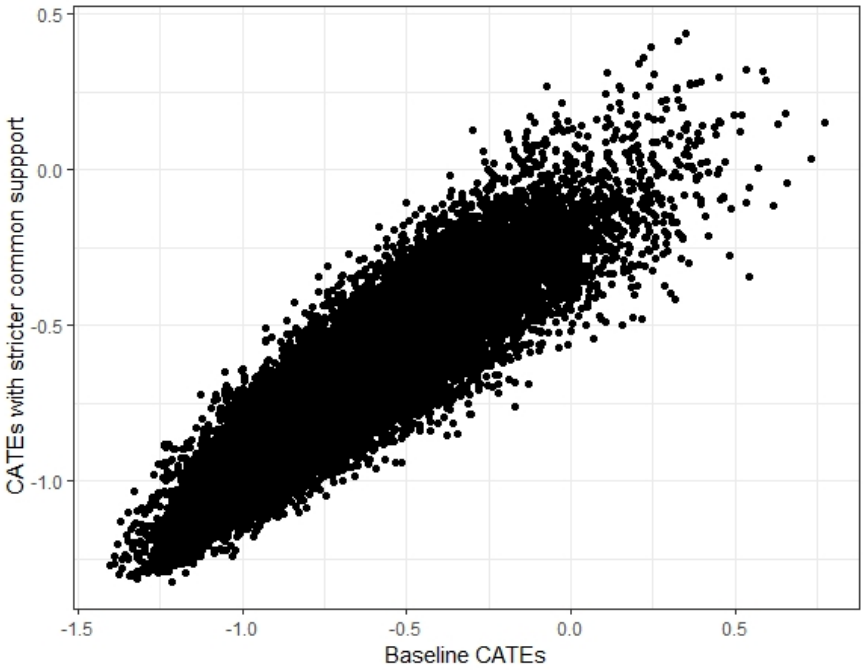


Table F.1: Average correlation between CATEs obtained from different random sample splits.

	Months employed since start participation			
	During first 6 months	During first 12 months	During first 31 months	During months 25-31
	(1)	(2)	(3)	(4)
(1) MCM, one-step EA, Post-LASSO	0.67	0.56	0.46	0.34
(2) MCM, two-step EA, Post-LASSO	0.66	0.57	0.24	0.22
(3) MCM, no EA, Post-LASSO	0.68	0.66	0.53	0.52
(4) MCM, one-step EA, adaptive LASSO	0.55	0.50	0.47	0.47
(5) MCM, two-step EA, adaptive LASSO	0.47	0.35	0.32	0.29
(6) MCM, no EA, adaptive LASSO	0.54	0.53	0.57	0.55
(7) MOM, Post-LASSO	0.63	0.64	0.46	0.44
(8) MOM, adaptive LASSO	0.50	0.47	0.44	0.41
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	0.64	0.52	0.30	0.24
(10) MCM, one-step EA, LASSO	0.59	0.53	0.48	0.51
(11) Procedure (1) + additional confounders 1	0.66	0.32	0.68	0.44
(12) Procedure (11) + additional confounders 2	0.66	0.47	0.54	0.40

Note: We estimate CATEs using different random sample splits and report the average correlation. We consider different methods of efficiency augmentation, variable selection, modifications and weights. EA is the abbreviation for efficiency augmentation. If not otherwise specified, IPW weights are used to balance the covariates. In procedure (9) we use radius-matching weights (Lechner Miquel, and Wunsch, 2011). See Online Appendix F for more details about the different procedures. In Online Appendix F.5, we describe how we select additional confounders for procedures (11) and (12).

Table F.2: Correlation between AITEs obtained from different empirical procedures.

Cumulated employment during first 12 months	(1)	(2)	(3)	(4)	(5)	(6)
(1) MCM, one-step EA, Post-LASSO	1.00					
(2) MCM, two-step EA, Post-LASSO	0.88	1.00				
(3) MCM, no EA, Post-LASSO	0.73	0.62	1.00			
(4) MCM, one-step EA, adaptive LASSO	0.71	0.49	0.57	1.00		
(5) MCM, two-step EA, adaptive LASSO	0.75	0.52	0.54	0.89	1.00	
(6) MCM, no EA, adaptive LASSO	0.66	0.48	0.84	0.70	0.66	1.00
(7) MOM, Post-LASSO	0.61	0.52	0.68	0.44	0.40	0.54
(8) MOM, adaptive LASSO	0.62	0.46	0.72	0.69	0.71	0.76
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	0.93	0.84	0.73	0.64	0.64	0.66
(10) MCM, one-step EA, LASSO	0.82	0.61	0.62	0.93	0.85	0.69
(11) Procedure (1) + additional confounders 1	0.74	0.74	0.52	0.48	0.49	0.45
(12) Procedure (11) + additional confounders 2	0.83	0.91	0.57	0.50	0.53	0.47
(13) Causal forest	0.45	0.37	0.36	0.38	0.36	0.43

Cumulated employment during first 12 months	(7)	(8)	(9)	(10)	(11)	(12)
(8) MOM, adaptive LASSO	0.52	1.00				
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	0.59	0.61	1.00			
(10) MCM, one-step EA, LASSO	0.49	0.63	0.73	1.00		
(11) Procedure (1) + additional confounders 1	0.59	0.44	0.69	0.57	1.00	
(12) Procedure (11) + additional confounders 2	0.61	0.45	0.82	0.59	0.82	1.00
(13) Causal forest	0.29	0.40	0.45	0.43	0.35	0.36

Note: Correlations of AITEs for different methods of efficiency augmentation, variable selection, modifications and weights. EA is the abbreviation for efficiency augmentation. If not otherwise specified, IPW weights are used to balance the covariates. In procedure (9) we use radius-matching weights (Lechner Miquel, and Wunsch, 2011). (11), (12) and (13) are estimated on different common support. Thus, the correlations are calculated for those observations being on support in both specifications. See Online Appendix F for more details about the different procedures. In Online Appendix F.5, we describe how we select additional confounders for procedures (11) and (12).

Table F.3: Correlation between AITEs obtained from different empirical procedures.

Cumulated employment during first 31 months	(1)	(2)	(3)	(4)	(5)	(6)
(1) MCM, one-step EA, Post-LASSO	1.00					
(2) MCM, two-step EA, Post-LASSO	0.88	1.00				
(3) MCM, no EA, Post-LASSO	0.54	0.41	1.00			
(4) MCM, one-step EA, adaptive LASSO	0.68	0.62	0.52	1.00		
(5) MCM, two-step EA, adaptive LASSO	0.82	0.80	0.46	0.79	1.00	
(6) MCM, no EA, adaptive LASSO	0.57	0.46	0.83	0.63	0.58	1.00
(7) MOM, Post-LASSO	0.46	0.33	0.65	0.37	0.34	0.50
(8) MOM, adaptive LASSO	0.57	0.47	0.69	0.70	0.61	0.72
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	0.88	0.67	0.57	0.59	0.61	0.59
(10) MCM, one-step EA, LASSO	0.82	0.71	0.58	0.90	0.80	0.64
(11) Procedure (1) + additional confounders 1	0.16	0.23	0.06	0.16	0.21	0.06
(12) Procedure (11) + additional confounders 2	0.24	0.26	0.13	0.21	0.23	0.12
(13) Causal forest	0.33	0.23	0.37	0.32	0.26	0.41
Cumulated employment during first 31 months	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
(8) MOM, adaptive LASSO	0.49	1.00				
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	0.49	0.53	1.00			
(10) MCM, one-step EA, LASSO	0.45	0.65	0.74	1.00		
(11) Procedure (1) + additional confounders 1	0.06	0.19	0.06	0.18	1.00	
(12) Procedure (11) + additional confounders 2	0.11	0.26	0.19	0.24	0.93	1.00
(13) Causal forest	0.26	0.41	0.39	0.36	0.04	0.12

Note: Correlations of AITEs for different methods of efficiency augmentation, variable selection, modifications and weights. EA is the abbreviation for efficiency augmentation. If not otherwise specified, IPW weights are used to balance the covariates. In procedure (9) we use radius-matching weights (Lechner Miquel, and Wunsch, 2011). (11), (12) and (13) are estimated on different common support. Thus, the correlations are calculated for those observations being on support in both specifications. See Online Appendix F for more details about the different procedures. In Online Appendix F.5, we describe how we select additional confounders for procedures (11) and (12).

Table F.4: Correlation between AITEs obtained from different empirical procedures.

Cumulated employment during months 25-31	(1)	(2)	(3)	(4)	(5)	(6)
(1) MCM, one-step EA, Post-LASSO	1.00					
(2) MCM, two-step EA, Post-LASSO	0.83	1.00				
(3) MCM, no EA, Post-LASSO	0.46	0.44	1.00			
(4) MCM, one-step EA, adaptive LASSO	0.69	0.75	0.47	1.00		
(5) MCM, two-step EA, adaptive LASSO	0.72	0.81	0.43	0.84	1.00	
(6) MCM, no EA, adaptive LASSO	0.48	0.51	0.82	0.61	0.56	1.00
(7) MOM, Post-LASSO	0.47	0.44	0.89	0.45	0.38	0.72
(8) MOM, adaptive LASSO	0.55	0.59	0.67	0.73	0.66	0.74
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	0.81	0.68	0.47	0.62	0.54	0.46
(10) MCM, one-step EA, LASSO	0.80	0.79	0.49	0.89	0.80	0.58
(11) Procedure (1) + additional confounders 1	0.20	0.27	0.09	0.24	0.28	0.10
(12) Procedure (11) + additional confounders 2	0.20	0.26	0.11	0.23	0.32	0.12
(13) Causal forest	0.26	0.29	0.37	0.33	0.32	0.41
Cumulated employment during months 25-31	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
(8) MOM, adaptive LASSO	0.65	1.00				
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	0.51	0.51	1.00			
(10) MCM, one-step EA, LASSO	0.48	0.65	0.76	1.00		
(11) Procedure (1) + additional confounders 1	0.11	0.25	0.23	0.29	1.00	
(12) Procedure (11) + additional confounders 2	0.09	0.28	0.24	0.27	0.87	1.00
(13) Causal forest	0.38	0.41	0.31	0.35	0.04	0.08

Note: Correlations of AITEs for different methods of efficiency augmentation, variable selection, modifications and weights. EA is the abbreviation for efficiency augmentation. If not otherwise specified, IPW weights are used to balance the covariates. In procedure (9) we use radius-matching weights (Lechner Miquel, and Wunsch, 2011). (11), (12) and (13) are estimated on different common support. Thus, the correlations are calculated for those observations being on support in both specifications. See Online Appendix F for more details about the different procedures. In Online Appendix F.5, we describe how we select additional confounders for procedures (11) and (12).

Table F.5: Descriptive statistics of AITEs by estimation procedure.

	Mean	Median	S.D.	Min	Max
	(1)	(2)	(3)	(4)	(5)
Cumulated employment during first 6 months					
(1) MCM, one-step EA, Post-LASSO	-0.78	-0.84	0.25	-1.41	0.77
(2) MCM, two-step EA, Post-LASSO	-0.76	-0.83	0.15	-0.90	0.28
(3) MCM, no EA, Post-LASSO	-0.80	-0.85	0.29	-1.95	0.93
(4) MCM, one-step EA, adaptive LASSO	-0.77	-0.81	0.32	-2.44	1.14
(5) MCM, two-step EA, adaptive LASSO	-0.77	-0.81	0.16	-1.46	0.67
(6) MCM, no EA, adaptive LASSO	-0.79	-0.81	0.26	-2.01	0.46
(7) MOM, Post-LASSO	-0.76	-0.83	0.27	-1.46	1.27
(8) MOM, adaptive LASSO	-0.77	-0.80	0.17	-1.39	1.47
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	-0.77	-0.83	0.23	-1.39	0.62
(10) MCM, one-step EA, LASSO	-0.77	-0.81	0.36	-2.33	1.26
(11) Procedure (1) + additional confounders 1	-0.80	-0.85	0.20	-1.30	1.13
(12) Procedure (11) + additional confounders 2	-0.77	-0.82	0.19	-1.22	0.71
(13) Causal forest	-0.82	-0.83	0.11	-1.36	0.15
Cumulated employment during first 12 months					
(1) MCM, one-step EA, Post-LASSO	-1.10	-1.20	0.32	-2.09	1.44
(2) MCM, two-step EA, Post-LASSO	-1.06	-1.13	0.14	-1.22	0.20
(3) MCM, no EA, Post-LASSO	-1.09	-1.14	0.54	-4.56	1.90
(4) MCM, one-step EA, adaptive LASSO	-1.06	-1.11	0.56	-4.70	3.12
(5) MCM, two-step EA, adaptive LASSO	-1.05	-1.10	0.27	-2.59	1.24
(6) MCM, no EA, adaptive LASSO	-1.08	-1.08	0.60	-4.46	1.74
(7) MOM, Post-LASSO	-0.98	-1.05	0.61	-3.59	4.64
(8) MOM, adaptive LASSO	-1.02	-1.07	0.40	-3.33	2.48
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	-1.09	-1.18	0.26	-1.93	0.87
(10) MCM, one-step EA, LASSO	-1.07	-1.14	0.60	-4.02	2.90
(11) Procedure (1) + additional confounders 1	-1.29	-1.34	0.14	-1.59	1.20
(12) Procedure (11) + additional confounders 2	-1.04	-1.10	0.18	-1.31	0.53
(13) Causal forest	-1.06	-1.06	0.22	-2.09	0.84

Table continues on next page >

Table F.5 continued.

	Mean	Median	S.D.	Min	Max
	(1)	(2)	(3)	(4)	(5)
Cumulated employment during first 31 months					
(1) MCM, one-step EA, Post-LASSO	-1.13	-1.25	0.60	-3.79	4.12
(2) MCM, two-step EA, Post-LASSO	-1.19	-1.23	0.20	-1.91	1.07
(3) MCM, no EA, Post-LASSO	-1.18	-1.19	1.39	-10.97	6.75
(4) MCM, one-step EA, adaptive LASSO	-1.06	-1.10	1.49	-11.57	11.33
(5) MCM, two-step EA, adaptive LASSO	-1.06	-1.08	0.52	-3.79	3.49
(6) MCM, no EA, adaptive LASSO	-1.19	-1.14	1.79	-9.60	10.03
(7) MOM, Post-LASSO	-0.88	-0.95	1.45	-7.02	16.17
(8) MOM, adaptive LASSO	-1.02	-1.11	1.35	-7.63	8.28
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	-1.14	-1.25	0.42	-2.82	1.62
(10) MCM, one-step EA, LASSO	-1.05	-1.11	1.26	-8.00	7.13
(11) Procedure (1) + additional confounders 1	-1.80	-1.83	0.38	-4.24	6.67
(12) Procedure (11) + additional confounders 2	-1.20	-1.26	0.32	-1.68	7.12
(13) Causal forest	-0.82	-0.83	0.59	-4.62	3.89
Cumulated employment during months 25-31					
(1) MCM, one-step EA, Post-LASSO	-0.04	-0.05	0.06	-0.32	0.48
(2) MCM, two-step EA, Post-LASSO	-0.05	-0.05	0.03	-0.27	0.26
(3) MCM, no EA, Post-LASSO	-0.03	-0.05	0.27	-1.45	1.67
(4) MCM, one-step EA, adaptive LASSO	0.00	0.00	0.29	-1.95	2.01
(5) MCM, two-step EA, adaptive LASSO	0.00	0.01	0.12	-0.77	1.53
(6) MCM, no EA, adaptive LASSO	-0.03	-0.03	0.38	-1.76	2.36
(7) MOM, Post-LASSO	0.01	-0.02	0.20	-0.77	1.53
(8) MOM, adaptive LASSO	0.00	-0.01	0.31	-1.64	2.57
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	-0.04	-0.05	0.03	-0.17	0.19
(10) MCM, one-step EA, LASSO	-0.03	-0.04	0.18	-1.26	1.27
(11) Procedure (1) + additional confounders 1	-0.03	-0.03	0.06	-1.18	0.63
(12) Procedure (11) + additional confounders 2	-0.08	-0.08	0.03	-0.16	0.44
(13) Causal forest	0.08	0.08	0.14	-0.84	0.93

Note: We obtain AITEs based on 30 different random sample splits. Standard deviations are abbreviated with S.D. in column (3). See Online Appendix F for more details about the different procedures. In Online Appendix F.5, we describe how we select additional confounders for procedures (11) and (12).

Table F.6: Number of selected variables in different estimation procedures.

	Mean	Median	S.D.	Min	Max
	(1)	(2)	(3)	(4)	(5)
Cumulated employment during first 6 months					
(1) MCM, one-step EA, Post-LASSO	34.9	32	18.2	5	87
(2) MCM, two-step EA, Post-LASSO	5.7	4	5.6	1	24
(3) MCM, no EA, Post-LASSO	52.1	48	30.8	9	113
(4) MCM, one-step EA, adaptive LASSO	114.3	112	40	47	187
(5) MCM, two-step EA, adaptive LASSO	46.8	41	28.9	2	111
(6) MCM, no EA, adaptive LASSO	84.8	87	36.5	13	156
(7) MOM, Post-LASSO	36.9	31	24.1	3	96
(8) MOM, adaptive LASSO	41.9	37	26.8	6	109
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	25.7	24	11.9	5	51
(10) MCM, one-step EA, LASSO	145.4	150	24.9	74	190
(11) Procedure (1) + additional confounders 1	21.9	19	14.1	0	60
(12) Procedure (11) + additional confounders 2	17.5	16	11.6	1	39
Cumulated employment during first 12 months					
(1) MCM, one-step EA, Post-LASSO	27.7	22	23.7	2	107
(2) MCM, two-step EA, Post-LASSO	3.3	2	3.8	0	16
(3) MCM, no EA, Post-LASSO	51.7	37	37.5	8	168
(4) MCM, one-step EA, adaptive LASSO	100	97	41	28	186
(5) MCM, two-step EA, adaptive LASSO	26.7	19	26.6	1	114
(6) MCM, no EA, adaptive LASSO	70.2	70	19.7	25	106
(7) MOM, Post-LASSO	34.5	30	18.8	9	100
(8) MOM, adaptive LASSO	52.6	47	29.2	13	121
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	13.2	11	9.7	0	47
(10) MCM, one-step EA, LASSO	107.1	102	19.5	70	150
(11) Procedure (1) + additional confounders 1	9.1	7	11	0	56
(12) Procedure (11) + additional confounders 2	6.3	6	5.3	0	19

Table continues on next page >

Table F.6 continued.

	Mean	Median	S.D.	Min	Max
	(1)	(2)	(3)	(4)	(5)
Cumulated employment during first 31 months					
(1) MCM, one-step EA, Post-LASSO	12.9	12	9.6	0	34
(2) MCM, two-step EA, Post-LASSO	3.1	2	3.8	0	14
(3) MCM, no EA, Post-LASSO	51.5	42	43	10	218
(4) MCM, one-step EA, adaptive LASSO	86.1	81	35.8	35	149
(5) MCM, two-step EA, adaptive LASSO	21.7	17	19.6	0	69
(6) MCM, no EA, adaptive LASSO	74	70	27.2	45	167
(7) MOM, Post-LASSO	35.6	29	22.6	12	107
(8) MOM, adaptive LASSO	72.3	66	32.3	26	144
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	7.9	4	14.4	0	75
(10) MCM, one-step EA, LASSO	73.6	79	19.4	0	101
(11) Procedure (1) + additional confounders 1	2.8	2	3.9	0	17
(12) Procedure (11) + additional confounders 2	3.6	2	5.6	0	27
Cumulated employment during months 25-31					
(1) MCM, one-step EA, Post-LASSO	4.8	2	6.8	0	29
(2) MCM, two-step EA, Post-LASSO	2.5	1	3.8	0	16
(3) MCM, no EA, Post-LASSO	40.4	27	36.5	10	160
(4) MCM, one-step EA, adaptive LASSO	78.6	72	36	22	176
(5) MCM, two-step EA, adaptive LASSO	13.7	11	12.8	1	46
(6) MCM, no EA, adaptive LASSO	61.4	59	19	35	137
(7) MOM, Post-LASSO	32.7	32	17.1	9	69
(8) MOM, adaptive LASSO	79.4	75	39.2	13	177
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	1	0	1.7	0	6
(10) MCM, one-step EA, LASSO	45.4	52	38.4	0	99
(11) Procedure (1) + additional confounders 1	2.2	1	3.7	0	17
(12) Procedure (11) + additional confounders 2	1.9	0	3.6	0	17

Note: Description of the number of selected heterogeneity variables in all 30 sample splits for all considered implementations and outcomes. Standard deviations are abbreviated with S.D. in column (3). See Online Appendix F for more details about the different procedures. In Online Appendix F.5, we describe how we select additional confounders for procedures (11) and (12).

Table F.7: Spearman rank correlation.

Cumulated employment during first 6 months	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) MCM, one-step EA, Post-LASSO	1.00											
(2) MCM, two-step EA, Post-LASSO	0,86	1.00										
(3) MCM, no EA, Post-LASSO	0.73	0.72	1.00									
(4) MCM, one-step EA, adaptive LASSO	0.78	0.57	0.63	1.00								
(5) MCM, two-step EA, adaptive LASSO	0.80	0.62	0.60	0.87	1.00							
(6) MCM, no EA, adaptive LASSO	0.67	0.58	0.82	0.68	0.65	1.00						
(7) MOM, Post-LASSO	0.74	0.71	0.86	0.62	0.61	0.71	1.00					
(8) MOM, adaptive LASSO	0.64	0.54	0.69	0.68	0.68	0.73	0.61	1.00				
(9) MCM, one-step EA, Post-LASSO with radius-matching weights	0.97	0.87	0.72	0.74	0.77	0.65	0.72	0.63	1.00			
(10) MCM, one-step EA, LASSO	0.84	0.63	0.63	0.92	0.80	0.63	0.62	0.62	0.81	1.00		
(11) (1) + additionally selected controls in propensity score	0.84	0.82	0.64	0.58	0.62	0.55	0.60	0.55	0.84	0.68	1.00	
(12) (11) + adding selected controls in outcome equation	0.88	0.86	0.67	0.60	0.65	0.57	0.63	0.57	0.88	0.69	0.94	1.00
(13) Causal forest	0.56	0.54	0.50	0.49	0.49	0.53	0.47	0.47	0.57	0.52	0.54	0.54

Additional References

Wang, H., C. Leng (2007): "Unified LASSO Estimation by Least Squares Approximation", *Journal of the American Statistical Association*, 102 (479), 1039-1048.