The net outcome of coaching and training for the self-employed

A statistical matching approach

Dr. Dirk Oberschachtsiek (Leuphana University of Lueneburg)
Patrycja Scioch (IAB)

Nürnberg, IAB; Nutzerkonferenz 08-09.04.2011

Motivation

- Self-employment as an increased employment option
- In Germany accompanied with an increase of market interventions (e.g. fostering self-employment and entrepreneurship)
- Different levels / types of political interventions:

  - e.g. taxes, subsidies, information,…. loans,… qualification,…

  - The Federal Employment Agency is one “big player” in this system of promotion activities:

    a) bridging allowance

    b) start-up-subsidy

    c) coaching

    d) training schemes
- The question is: what is the return related to these promotion activities?

Previous Research

Self-employment subsidies (*evaluation of financial support programs*)

Almus/Prantl (1999)

Pfeiffer/Reize (2000); Wiessner (2001); …

Baumgartner/Caliendo (2008)

Training schemes (*results related to non-financial support*)

Shutt/Sutherland (2003)

Eckl et. Al (2009)

Method

→ What is the net gain of a) training and b) coaching and c) other (flexible) promotion devices (focus: non-financial support programs)

methodological approach:

Estimate the effect of a promotion (D) on the survival chances (Y)

using a statistical matching approach framework.

SUTVA as the overall “identification”-assumption; CIA as the specific identification assumption

several challenges (clustering; unobserved substitutes due to multiple political actors ,…)

Interventions

Self-employment training

Part of the ESF-Funding program; 4 to 12 weeks of training: developing business plans,.. marketing strategies... bookkeeping,...

…. *enhancing qualification and establishing better learning capacities (prior start-up period)*

Founder coaching

Part of a ESF-Funding Program; unknown duration; quality varies across regional districts (heterogeneous suppliers and different regional strategies).

…. *ensuring better „information“ and improving learning capacities (post entry period)*

Other schemes

Part of the so called Discretionary Start-up Subsidies (Gründungshilfen; Freie Förderung; §10 SGB III); high degrees of freedom on the local level in managing related promotion schemes (not standard in Germany); across time self-employment became one of the most important subfields: §10 (discretionary) start-up support

…. *usually focused on qualification and substitutes training or coaching*

Data

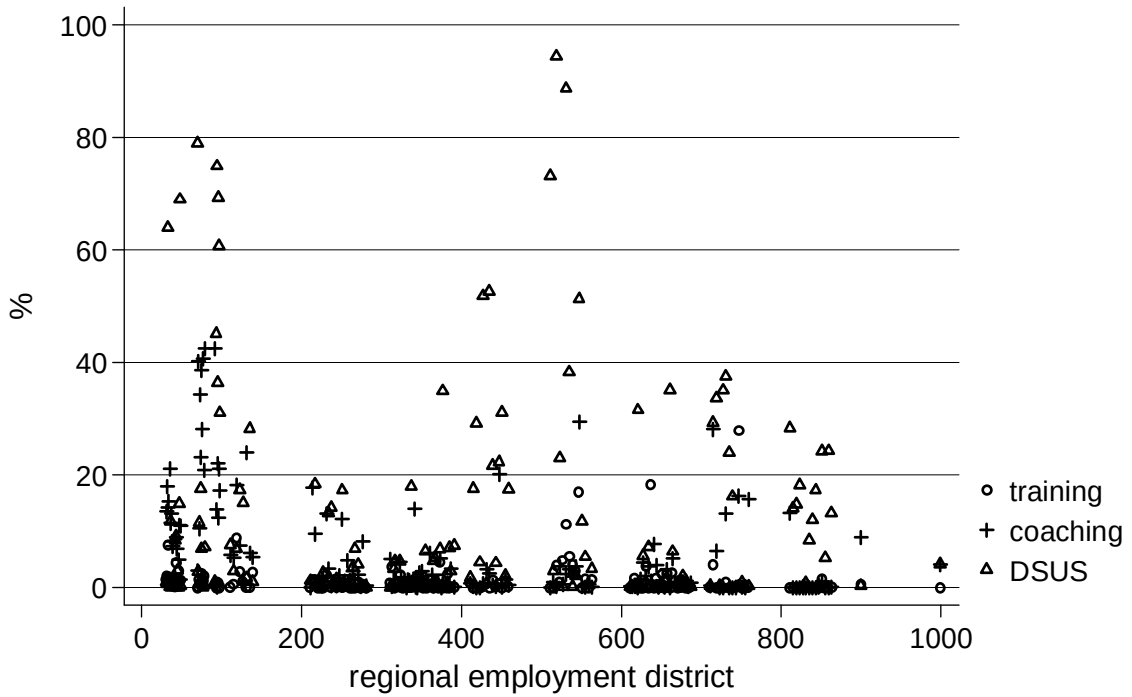
Integrated Employment Biographies: episodes of employment, unemployment, job search and participation in schemes of the active labor market policy; observation period: 1999 to 2005; entries 2000 to 2003

additional data taken from official data sources to include local labor market information (unemployment rate,… firm hazard, unemployment variance,…)

reference group: individuals who received a bridging allowance and no other self-employment promotion; not studied are combined promotions (e.g. coaching plus ..)

outcome: exit probability (Pr(T<36 months)) and survival chances (time depending)

Relative entries across regions



source: IEB, own calculations
the x-axis indicates the identifier of a local labor market district

share = ratio between no of entries in additional support and entries in bridging allowance in region j

Explaining Entries

	Training		Coaching		DSUS	
Block of variables	BIC	LR	BIC	LR	BIC	LR
model 1 (only b1)	40,459.61	1782.47***	171,601.50	7163.75***	200,113.40	1260.58***
model 2 (adding b2 to b1)	33,738.78	8204.86***	129,326.40	44134.18***	152,136.90	50014.96***
model 3 (adding b3 to model2)	33,057.17	950.84***	128,866.70	926.89***	150,720.80	1685.34***

Notes: the blocks of attributes are introduced sequentially in nested models.
The blocks of attributes contain: b1 (7 dummy variables for the # half-year of entry); b2 (regional information, 108 to 159 variables, including regional conditions and dummy variables for each local labor market district); b3 (individual information, 94-99 variables, including gender, age, qualification level, employment background and occupational background based on a two digit classification)

Matching procedure

1. Identify *j* and *i*.
2. Skip regions with no support (zero participants between 2000 and 2003).
3. Estimate three propensity scores *Ps(x): Pr(D=1|X^i), Pr(D=1|X^rc)* and *Pr(D=1|X^rd)*;[1] where *Pr(D=1|X=x) = 1 / (1 + e^{X'β})*.
4. Stratify the matching procedure into matching clusters (by annual quarter and type of region[2]).
5. Calculate the Mahalanobis distance based on $Ps^{i,rc,rd}(x)$ and selected *X* as the *B(x)*
6. Set a multiplier $m \in]0,1]$.
7. Run a pre-matching process to identify *h* based on the distance distribution of nearest neighbors in each matching cluster: a) Select a treated observation *i*. b) Use the nearest neighbor in terms of the Mahalanobis distance, given that *j* lies within the cluster *cl*; save the distances between the comparisons. c) Extract the 75th percentile of all distance values within cluster *cl*. d) Use the 90th percentile across all '*cl* p75-distance values' as the bandwidth *h*.
8. Run the clustered matching algorithm based on *h* taken from (7) which is multiplied by *m*.

   if the balancing property is not sufficient, re-run from (7) based on additional attributes that are added to the calculation of the Mahalanobis distance.

   if balancing is not sufficient based on the addition of attributes, re-run from (6) with a smaller multiplier.

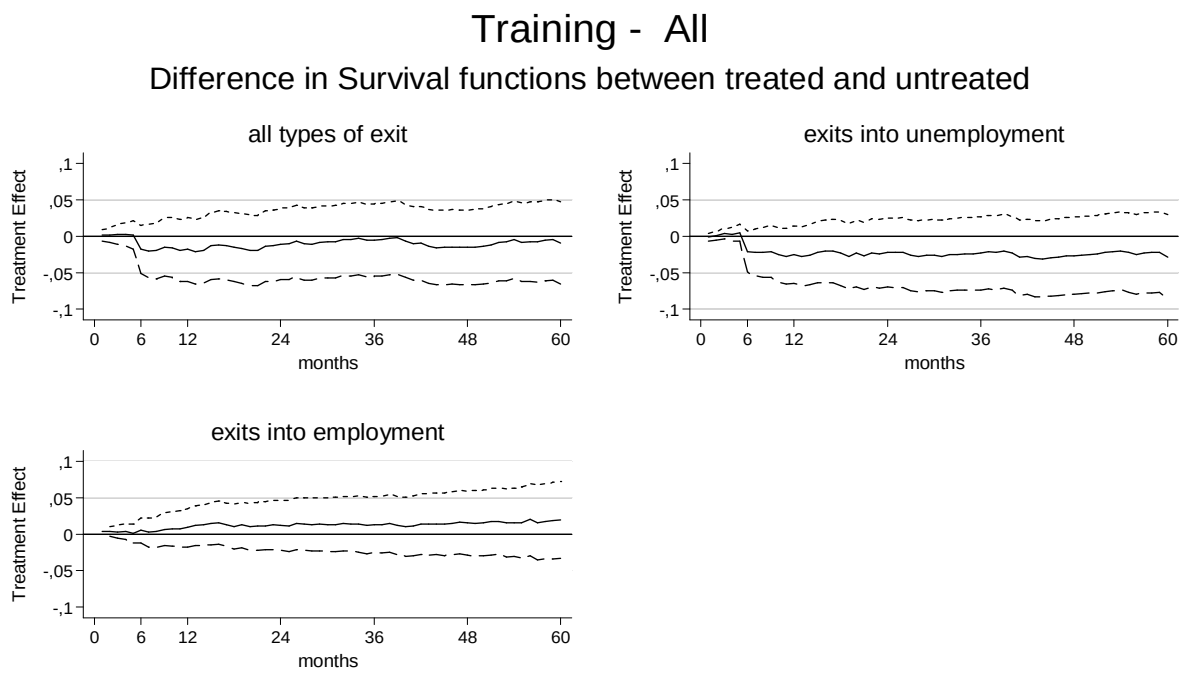
ATT; Prob(T<36 months)

Treatment / type of exit	on support[A] Nj	on support[A] Ni	matched[A] Nj	matched[A] Ni	ATT[B]	inference se	inference se^r/se, I	inference se^r/se, II	balance (MSB)[C] before	balance (MSB)[C] after	F-test[D] before	F-test[D] after
Training												
all types:	1555	118236	1555	32968	0.006	0.015	1.799	0.818	24.866	2.380	0.000	0.631
unempl.:	1555	118236	1555	32968	0.023+	0.014	1.364	1.031	24.866	2.380	0.000	0.631
employment:	1555	118236	1555	32968	-0.013	0.009	1.163	1.020	24.866	2.380	0.000	0.631
coaching												
all types:	7204	177573	7204	27529	0.002	0.008	2.237	1.623	28.573	0.970	0.000	0.823
unempl.:	7204	177573	7204	27529	0.007	0.007	2.166	1.179	28.573	0.970	0.000	0.823
employment:	7204	177573	7204	27529	-0.013*	0.005	1.392	1.060	28.573	0.970	0.000	0.823
discr. start-up support (DSUS)												
all types:	8942	206189	8942	22033	0.010	0.007	3.633	1.042	24.773	0.885	0.000	0.523
unempl.:	8942	206189	8942	22033	0.021*	0.007	2.329	0.888	24.773	0.885	0.000	0.523
employment:	8942	206189	8942	22033	-0.011*	0.005	1.942	1.358	24.773	0.885	0.000	0.523

[A] j and i are indicators for the population (i = treated population; j = untreated persons)
[B] ATT stands for the average treatment effect on the treated; the ATT is calculated on the basis of Formula (4): $Pr(T^k \leq 36)$
[C] the balancing property is calculated as the averaged mean standardized bias based on individual and regional variables as well as on the three propensity scores
[D] the test used is an F-test of the joint insignificance of all regressors before and after matching
+ indicates statistical significance at the 90% level; * indicates statistical significance at the 95% level

ATT; Survival

•



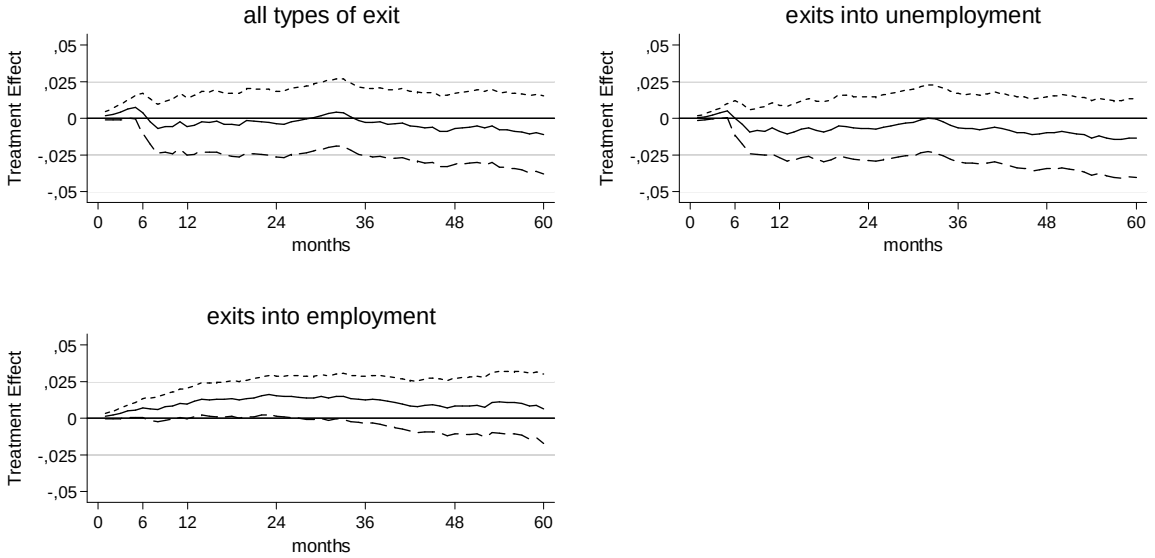
obs: 1555 treated, 32968 untreated
source: IEB, own calculations
bounds base on the Greenwood (1987) approximation of the standard errors

ATT; Survival

•

Coaching - All

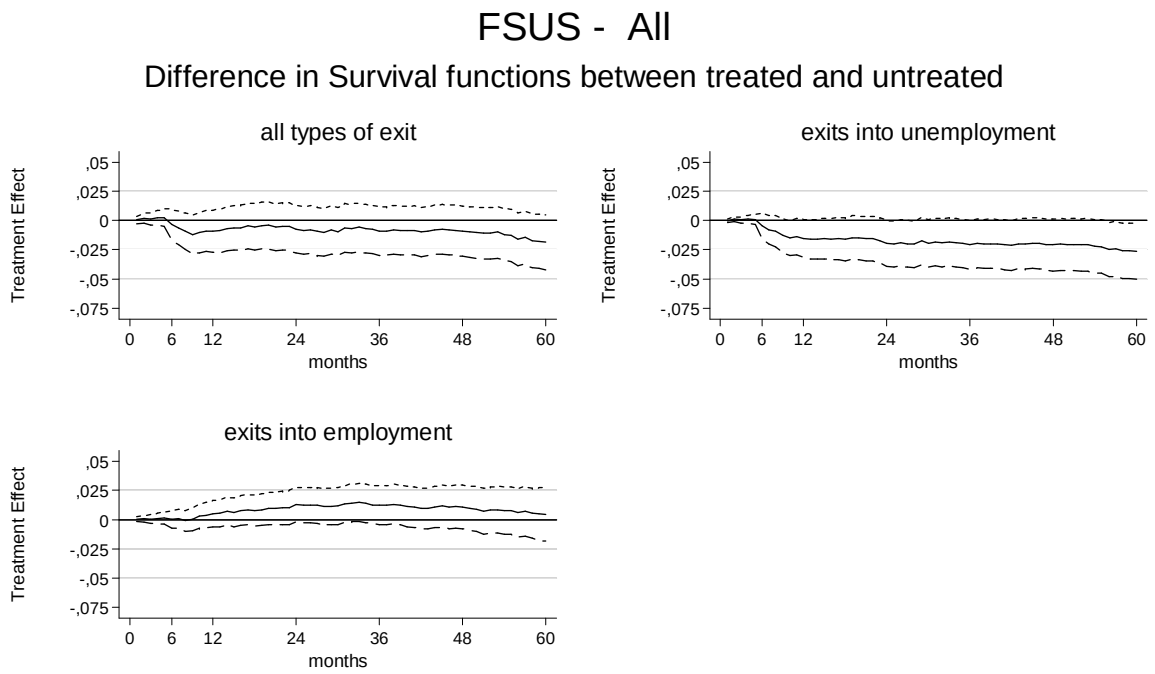
Difference in Survival functions between treated and untreated



obs: 7204 treated, 27529 untreated
source: IEB, own calculations
bounds base on the Greenwood (1987) approximation of the standard errors

ATT; Survival

•



obs: 8942 treated, 22033 untreated
source: IEB, own calculations
bounds base on the Greenwood (1987) approximation of the standard errors

Robustness checks

1) heterogeneous treatment effects across gender: *no substantial differences*
2) importance of unobserved heterogeneity in the treatment selection – rosenbaum-bounds: *no substantial differences*
3) presence of potential substitutes: exclude regions with high share of ESF-regional promotion activities (external data source): *no substantial differences*
4) Assume the presence of „negative creaming“ – focusing on regions with higher share of additional promotion should reduce the likelihood of conditioning on unpromising business projects: *no substantial differences*

Discussion I

on average additional support as identified with training, coaching and other schemes

….. does not reduce the likelihood of quitting self-employment (does not improve survival chances)….

learning is not improved ….

because the likelihood to quit into an employment state is not statistically higher for those with a promotion (partly: inverse effects)

However:

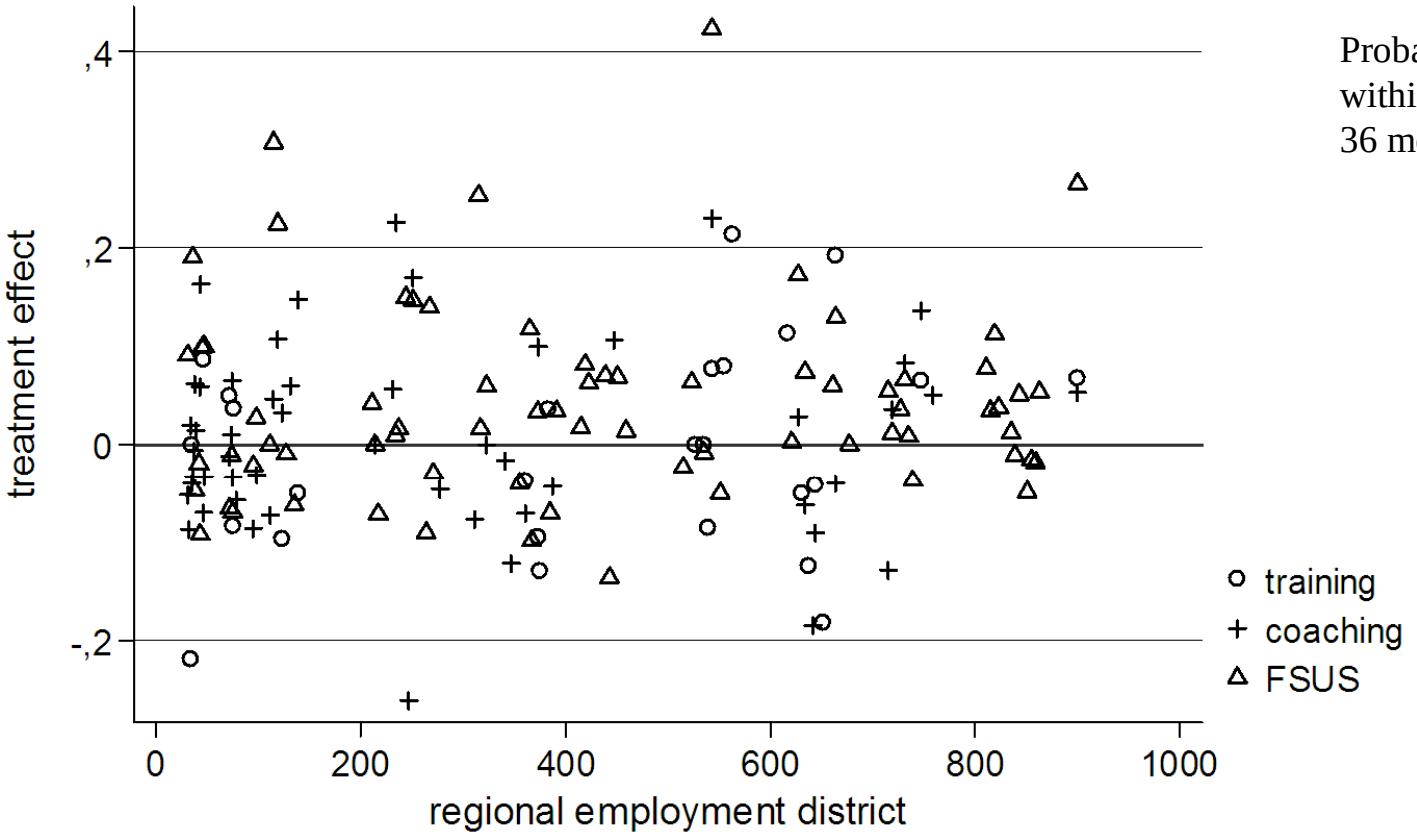
Further heterogeneous effects may be present (two sources:

1) real heterogeneous treatment effects across regions;

2) heterogeneous treatments across regions)

Regional variation so far unstudied

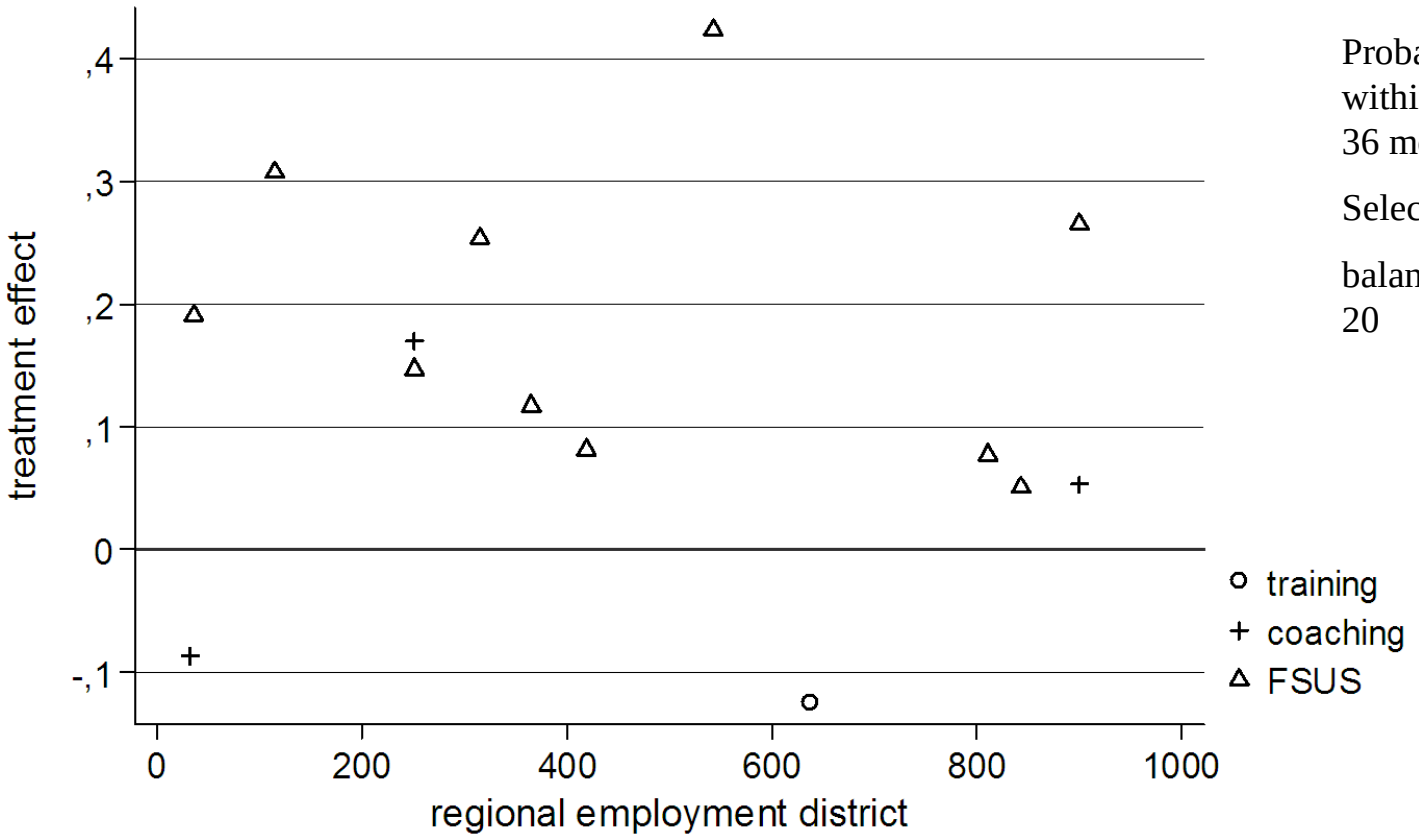
Heterogeneous treatment effects (all)



Probability to exit within a period of 36 months

source: IEB, own calculations

Heterogeneous treatment effects (selection)



Probability to exit within a period of 36 months

Selection:

balanced and n >= 20

source: IEB, own calculations; statistically significant treatment effects
the y-axis indicates the id of an employment office
effects are only reported if the F-test supports balance of the treated and untreated

Outlook

$ATT_C = f(alq_C; \textit{policy strategy})$; with C as an indicator for the Cluster

Weighting scheme: balance property and statistical significance