Evaluating training measures using the Generalized Propensity Score

Jochen Kluve, RWI Essen and IZA

Hilmar Schneider, IZA

Arne Uhlendorff, DIW Berlin and IZA

Zhong Zhao, IZA

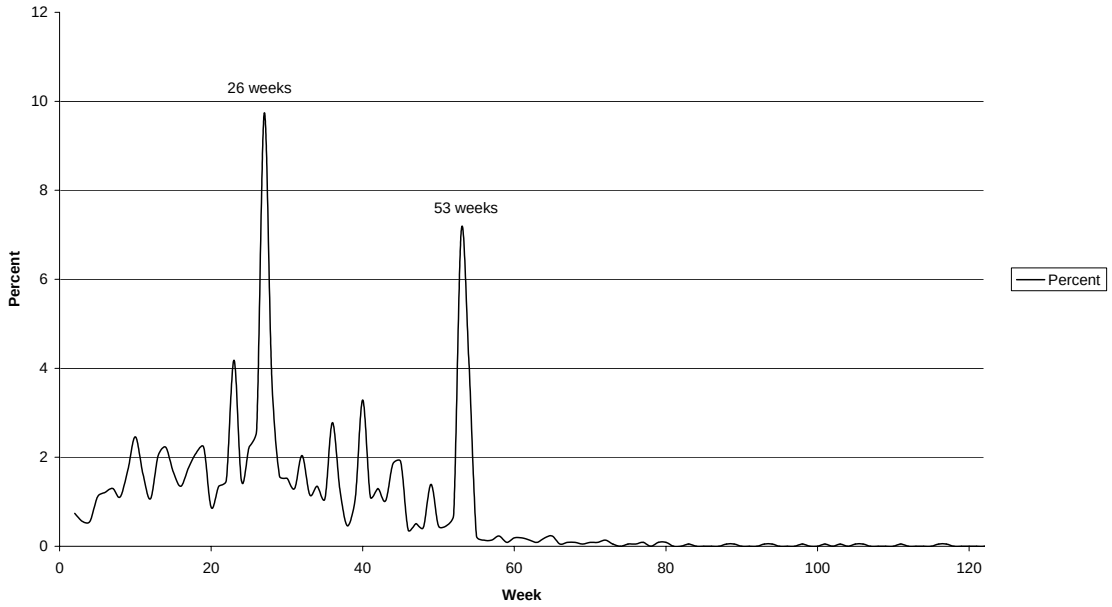
St Gallen, COST meeting

Motivation I

- High amount of money spent on training programs
- A lot of studies on effectiveness of training programs
- Differentiation by the type of program, region, gender, duration (long vs. short programs, which probably goes along with a different type)
- Mixed results for Germany, but the recent studies seem to find at least for some sub-groups significantly positive effects
- In Germany the duration of programs varies within one program type

Motivation II

Figure 1. Distribution of Treatment Duration



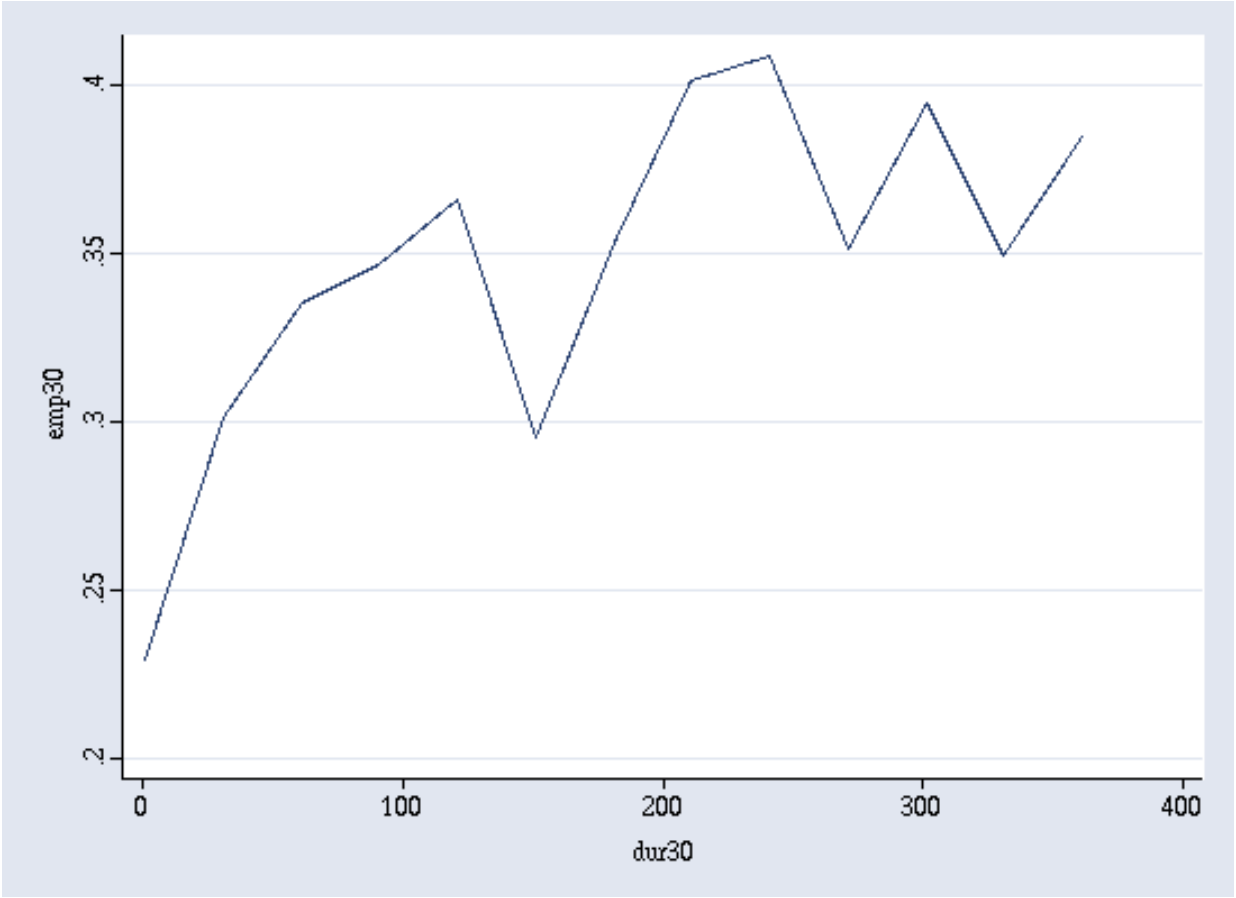
- Do the outcomes vary with the program duration?
- Data: participants in two relatively short program types in Germany
- Method: dose response function applying generalized propensity score following Hirano and Imbens (2004)

Data I

- Integrated Employment Biographies (IEB) of the German Federal Employment Agency
- detailed information on employment subject to social security contribution, unemployment, education, disability, nationality and participation in programs of ALMP $\longrightarrow$ richness of the data important for the matching approach
- program types: occupation-specific training programs and general training programs, which do not lead to the acquisition of a degree, with a focus on class room training or with a focus on practical experience, both with similar distributions of duration and similar overall treatment effects
- random sample of around 250 male participants per quartal, 2039 participants in 2000/2001
- we include duration ranging from 10 days up to 13 months
- outcome variables: employment status and earnings if employed
- both outcomes measured (i) three years after program entry and (ii) one year after program end

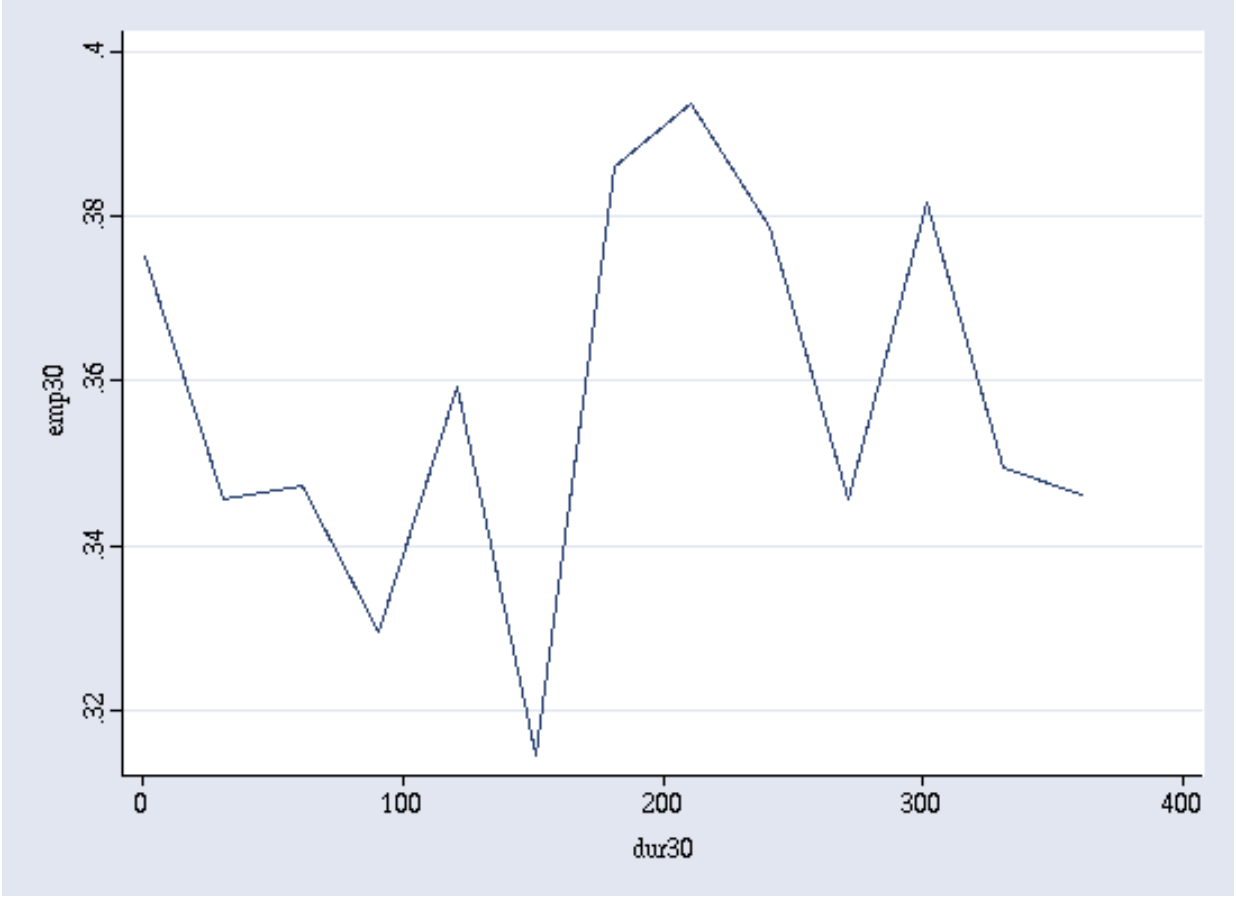
Data II

Employment Probability three years after program start



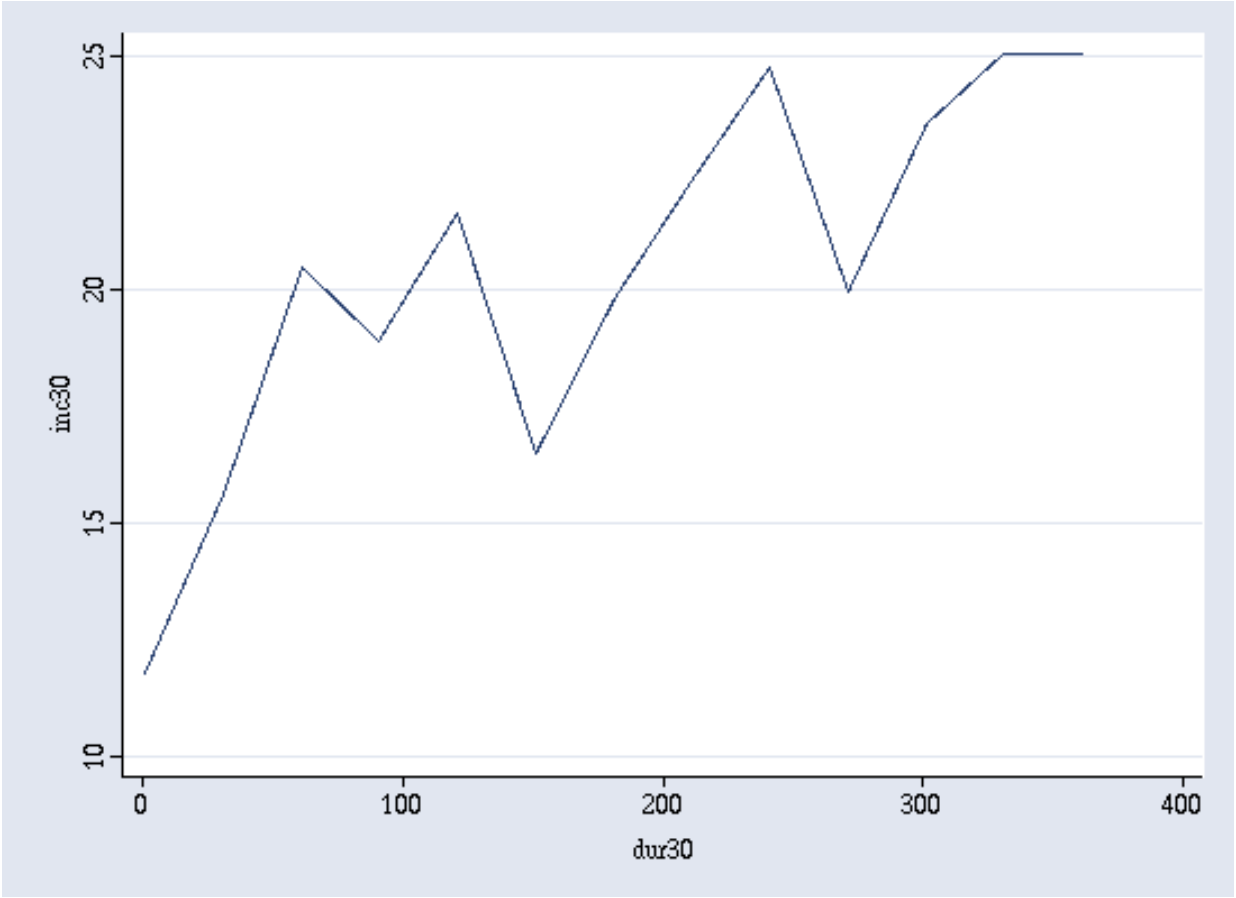
Data III

Employment Probability one year after program exit



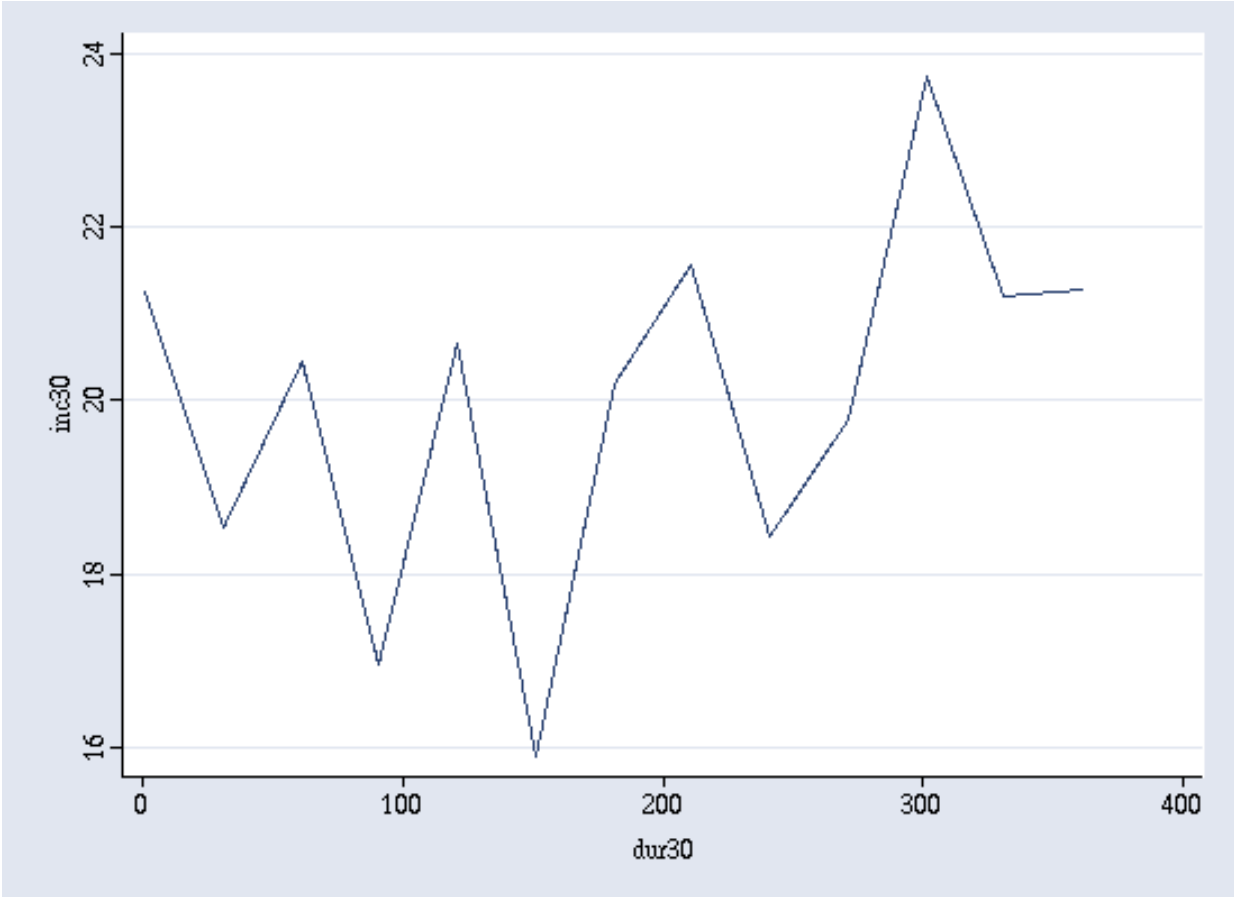
Data IV

Earnings three years after program start



Data V

Earnings one year after program end



Method

Instead of binary treatment (Rosenbaum and Rubin 1983) and multi-valued treatment (Imbens 2000 and Lechner 2001) we have a continuous treatment.

Dose-response function

$$\mu(t) = E[Y_t(t)] \tag{1}$$

Unconfoundness assumption

$$Y(t) \perp T|X \tag{2}$$

Generalized Propensity Score (GPS)

$$R = r(T, X) \tag{3}$$

$X \perp 1\{T = 1\}|r(t, X) \rightarrow$ within strata with the same value of r(t,X) the probability that T=t does not depend on $X \rightarrow$ with (2) this implies that assignment is unconfounded given the GPS

$r(t, X) = f_{T|X}(t|X)$: conditional density of the treatment given the covariates

Implementation

1. Generalized Propensity Score

Normal distribution for the treatment given the covariates

$$T_i|X_i \sim N(\beta_0 + \beta_1' X_i, \sigma^2)$$

$$\widehat{R_i} = \frac{1}{\sqrt{2\pi\widehat{\sigma}}} \exp(-\frac{1}{2\widehat{\sigma}^2}(T_i - \widehat{\beta_0} - \widehat{\beta_1}' X_i)) \tag{4}$$

2. Dose Response Function

Estimating a flexible conditional expectation function of Y_i given T_i and R_i

$$\begin{aligned} E[Y_i|T_i, R_i] = \quad & \alpha_0 + \alpha_1 T_i + \alpha_2 T_i^2 + \alpha_3 T_i^3 + \alpha_4 R_i + \alpha_5 R_i^2 + \alpha_6 R_i^3 \\ & + \alpha_7 T_i R_i + \alpha_8 T_i^2 R_i + \alpha_9 T_i R_i^2 \end{aligned} \tag{5}$$

Given the estimated parameters we estimate the average potential outcome at treatment level t as:

$$\widehat{E[Y(t)]} = \frac{1}{N} \sum_{i=1}^{N} (\widehat{\alpha_0} + \widehat{\alpha_1} t + \widehat{\alpha_2} t^2 + ... + \widehat{\alpha_9} t \widehat{r}(t, X_i)^2) \tag{6}$$

Testing for balancing of covariates

In contrast to the binary case - comparison of covariate means for treated and untreated before and after matching - and the multi-valued treatments (Imbens 2000, Lechner 2001) - pairwise comparison of means - testing for covariate balancing is more difficult with continuous treatment. We follow Hirano and Imbens (2004):

1. Dividing the sample into three groups according to the distribution of the treatment length (cutting at the 30^{th} and the 70^{th} percentile)

2. Within each group we evaluate the GPS at the median of the treatment variable

3. We divide each group into five blocks by the quintiles of the GPS evaluated at the median (distribution of the individuals in that group)

4. Within each block, we compare individuals who are "treated", i.e. belonging to the group (according to step 1), with individuals who are in the same block but belong to another group.

5. Differences before adjustment: Means of individuals in group 1 are compared to means of individuals in the other groups

6. Differences after adjustment: Means of individuals in group 1 are compared to means of individuals who belong to the same block. A weighted average over the five blocks in each treatment-level group is calculated.

If adjustment for the GPS properly balances the covariates, we would expect all differences to be statistically insignificant

Balancing of covariates I

	(1)	(2)	(3)	(4)	(5)	(6)
		Unadjusted			Adjusted	
Covariate	[10,152]	[153,271]	[272,390]	[10,152]	[153,271]	[272,390]
Age	**2.50**	0.63	**-3.19**	0.63	0.23	0.03
No disability	**-2.90**	-0.36	**3.30**	-0.77	0.06	0.03
Disab. low degree	**2.08**	-0.19	-1.89	0.52	-0.17	0.07
Disab. medium degree	1.67	0.45	**-2.16**	0.49	-0.08	0.10
Disab. high degree	1.57	1.20	**-2.87**	0.47	0.38	-0.25
German	0.74	1.21	**-2.04**	-0.22	0.16	-0.35
Foreigner EU	**-2.74**	0.32	**2.41**	-0.53	-0.01	0.39
Foreigner Non-EU	0.31	-1.41	1.19	0.45	-0.16	0.23
No graduation	**-2.84**	-1.21	**4.17**	-0.37	-0.07	0.38
First state of sec. level	**-3.48**	**-5.15**	**9.14**	0.82	-1.06	0.32
Second stage of sec. level	**3.22**	**1.98**	**-5.39**	-0.51	-0.06	-0.41
Technical college entrance qualification	0.76	**2.43**	**-3.37**	-0.60	0.79	-0.14
Qualification for university entrance	**3.96**	**5.66**	**-10.22**	0.27	1.56	-0.20
No vocational degree	**-4.62**	**-3.60**	**8.60**	-0.10	-0.65	1.06
In-plant training	1.74	-0.09	-1.65	-0.16	-0.27	-0.58
Off-the-job training, vocational school	1.16	**2.45**	**-3.79**	-0.31	0.73	-0.10
University, advanced technical college	**4.47**	**5.04**	**-10.06**	0.90	1.32	-0.52

t-statistics for Equality of Means, n=2,039

Balancing of covariates II

	(1)	(2)	(3)	(4)	(5)	(6)
		Unadjusted			Adjusted	
Covariate	[10,152]	[153,271]	[272,390]	[10,152]	[153,271]	[272,390]
Previous unemployment duration	0.94	**-2.04**	-1.24	0.87	-0.71	0.95
Duration of last employment	1.38	0.28	-1.69	0.24	0.13	-0.21
log(wage) of last employment	-0.65	0.46	0.17	-0.33	0.19	0.17
no last employment observed	-1.71	0.16	-1.56	0.63	-0.07	-0.06
share in emp., 1. year bef. prog. start	-0.78	0.00	0.78	0.03	-0.11	-0.37
share in emp., 2. year bef. prog. start	-0.43	**2.00**	-1.71	-0.66	-0.53	-0.55
share in emp., 3. year bef. prog. start	-0.18	1.10	-0.99	-0.45	0.16	-0.09
share in emp., 4. year bef. prog. start	1.65	-0.77	-0.82	0.14	-0.63	0.10
share in unemp., 1. year bef. prog. start	0.34	-1.33	1.07	0.57	-0.50	1.05
share in unemp., 2. year bef. prog. start	**-2.06**	**-2.62**	**4.89**	-0.12	-0.64	1.61
share in unemp., 3. year bef. prog. start	-1.21	**-2.34**	**3.72**	0.06	-0.53	1.23
share in unemp., 4. year bef. prog. start	**-3.19**	-0.85	**4.12**	-0.74	-0.04	1.67

t-statistics for Equality of Means, n=2,039

Results I

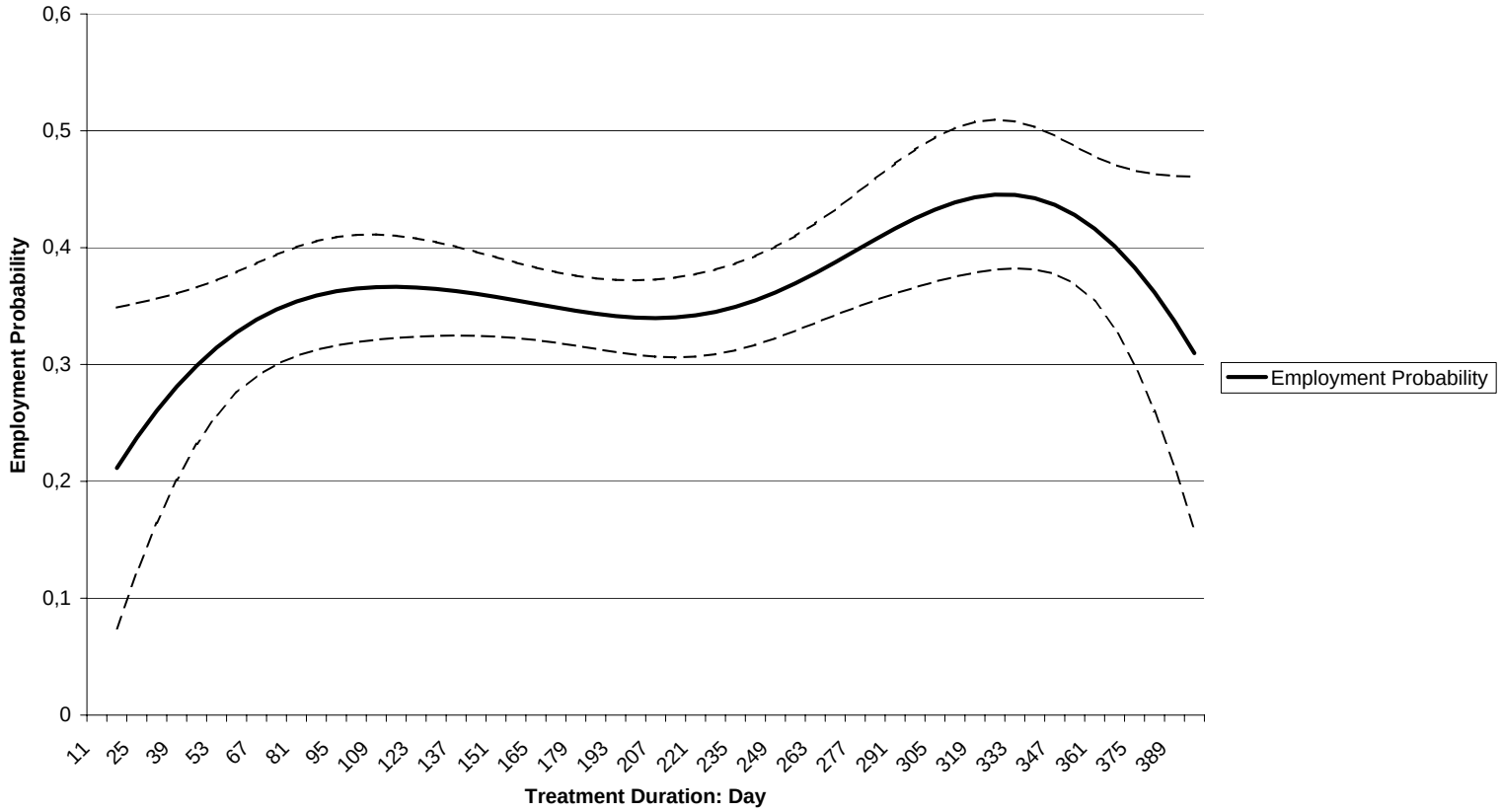
	Employment 3 years after entry		Employment 1 year after exit	
Days of treatment	Mean	Std. Error	Mean	Std. Error
11	0.21	0.07	0.31	0.07
32	0.28	0.04	0.33	0.04
53	0.33	0.03	0.35	0.03
74	0.35	0.02	0.36	0.02
95	0.37	0.02	0.36	0.02
116	0.37	0.02	0.36	0.02
137	0.36	0.02	0.35	0.02
158	0.35	0.02	0.35	0.02
200	0.34	0.02	0.36	0.02
221	0.35	0.02	0.37	0.02
242	0.36	0.02	0.38	0.02
263	0.39	0.02	0.38	0.02
284	0.42	0.03	0.38	0.03
305	0.44	0.03	0.38	0.03
326	0.45	0.03	0.37	0.03
347	0.43	0.03	0.35	0.03
368	0.38	0.04	0.34	0.04
389	0.31	0.08	0.32	0.08

Results II

	Earnings 3 years after entry		Earnings 1 year after exit	
Days of treatment	Mean	Std. Error	Mean	Std. Error
11	8.29	4.81	14.95	4.85
32	14.73	2.76	18.89	2.73
53	18.84	1.84	21.03	1.78
74	20.96	1.75	21.61	1.68
95	21.65	1.74	21.10	1.66
116	21.53	1.57	20.13	1.50
137	21.02	1.30	19.22	1.26
158	20.32	1.08	18.66	1.08
200	19.21	1.12	18.77	1.12
221	19.45	1.25	19.26	1.22
242	20.55	1.38	19.85	1.33
263	22.39	1.63	20.39	1.57
284	24.52	2.01	20.74	1.95
305	26.31	2.31	20.83	2.23
326	27.18	2.31	20.69	2.21
347	26.75	2.24	20.44	2.06
368	24.94	3.13	20.28	2.87
389	21.95	5.50	20.46	5.22

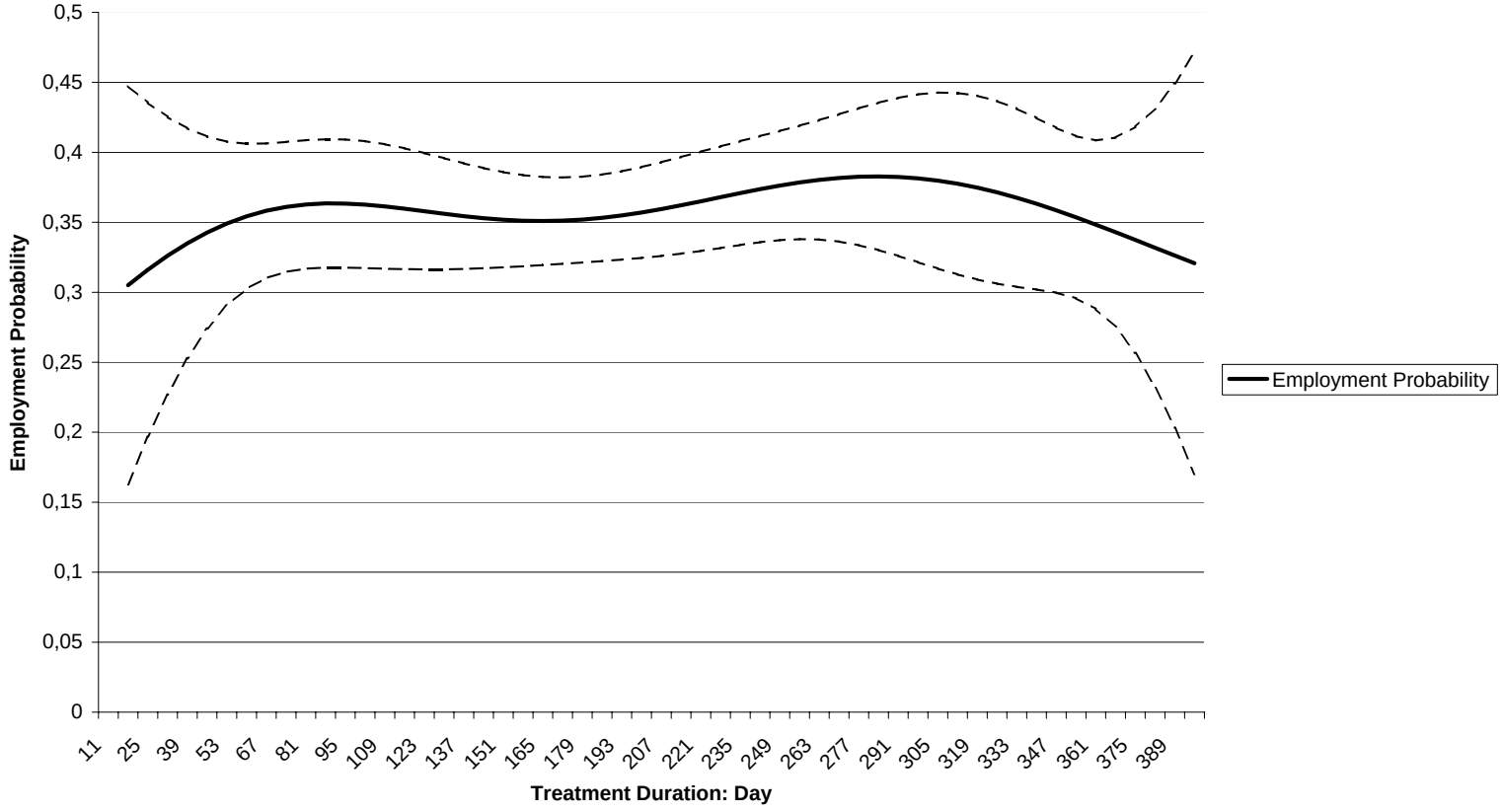
Results III

Figure 2. Employment Probability at Time 3 Years after Entry into the Program



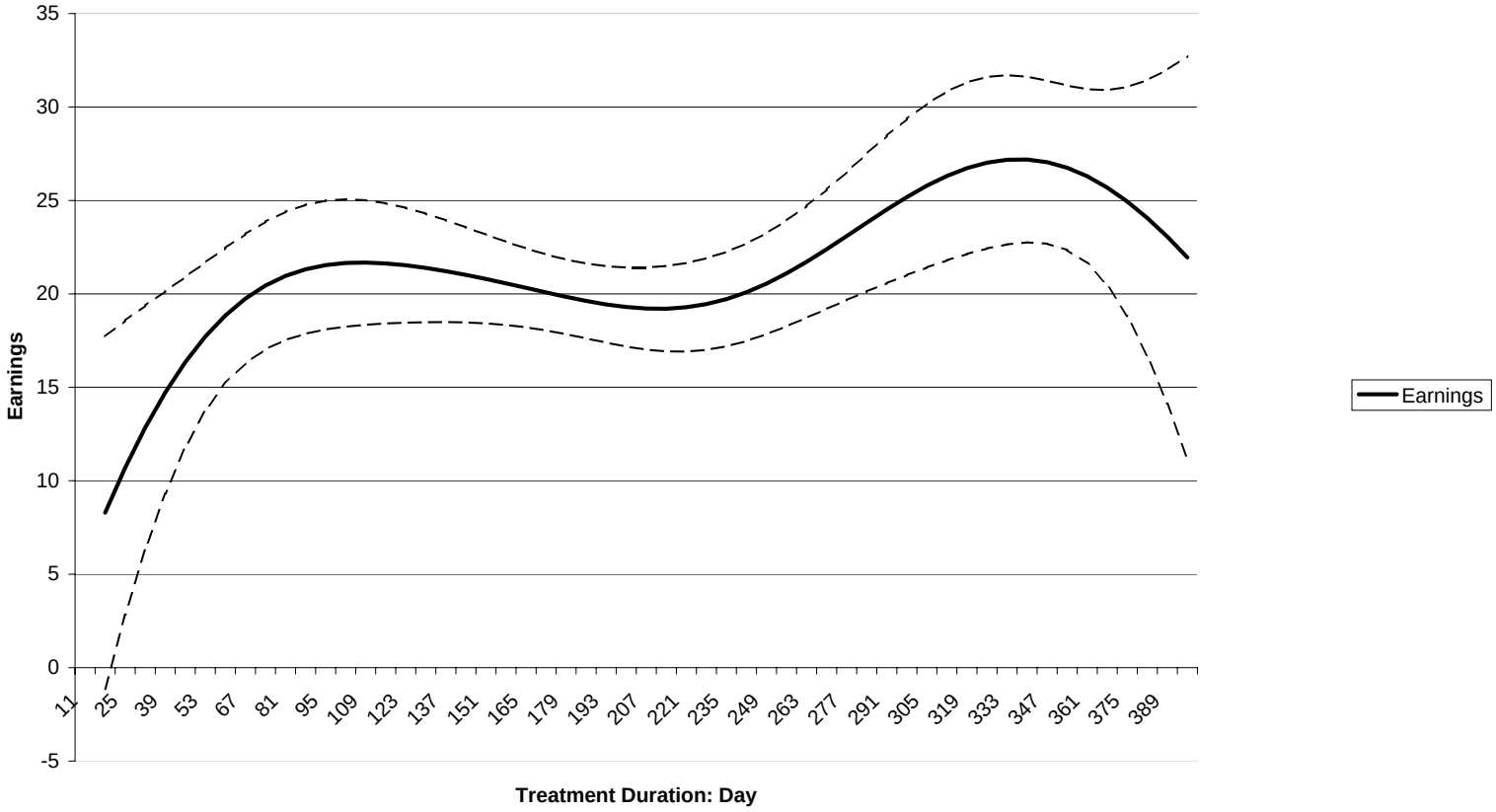
Results IV

Figure 3. Employment Probability at Time 1 Years after Exit from the Program



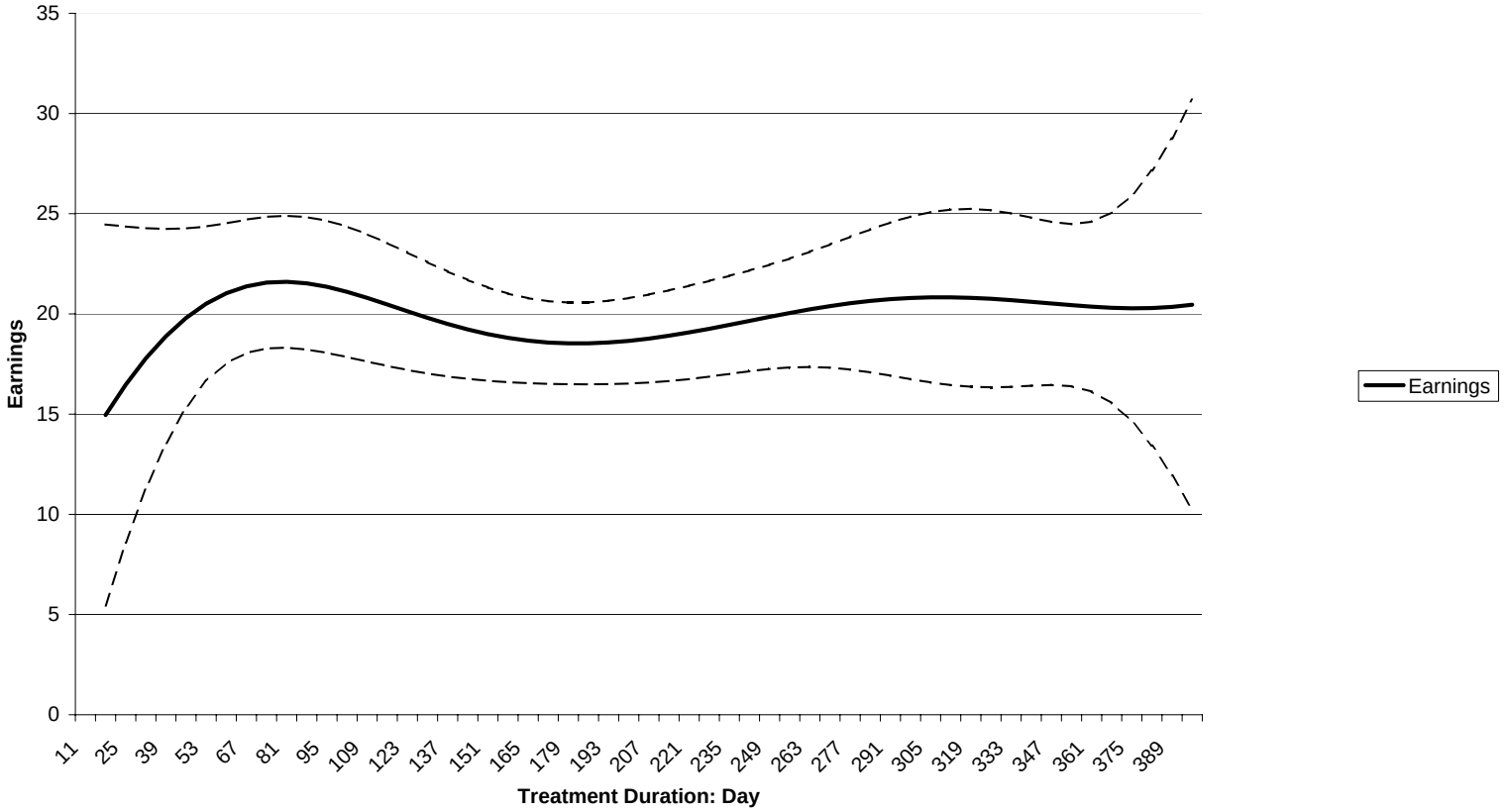
Results V

Figure 4. Earnings at Time 3 Years after Entry into the Program



Results VI

Figure 5. Earnings at Time 1 Years after Exit from the Program



Conclusions

- Implementation of the GPS seems to work, no significant differences in covariates after adjustment
- The descriptive positive association of treatment level and employment probability clearly decreases after adjustment
- Only small differences - if at all - between different treatment levels
- Preliminary policy conclusion: If the treatment level has only a small impact on the employment probability, it could be efficient to reduce the program duration of the analyzed program types

Further steps

- Inclusion of additional covariates like marital status and local unemployment rate
- Potential endogeneity of the program duration: Same analysis with the planned duration
- Different number of groups: are the covariates still balanced?
- Comparison to non-participants
- Measuring the outcomes at more points in time