CEPA DP No. 30

MARCH 2021

Design and Effectiveness of Start-Up Subsidies: Evidence from a Policy Reform in Germany

Marco Caliendo
Stefan Tübbicke

Center for Economic Policy Analysis

CEPA Discussion Papers

Center for Economic Policy Analysis
https://www.uni-potsdam.de/cepa

University of Potsdam
August-Bebel-Straße 89, 14482 Potsdam
Tel.: +49 331 977-3225
Fax: +49 331 977-3210
E-Mail: dp-cepa@uni-potsdam.de

ISSN (online) 2628-653X

CEPA Discussion Papers can be downloaded from RePEc
https://ideas.repec.org/s/pot/cepadp.html

Published online at the Institutional Repository of the University of Potsdam
https://doi.org/10.25932/publishup-50005

CEPA Discussion Paper No. 30 March 2021

Design and Effectiveness of Start-Up Subsidies: Evidence from a Policy Reform in Germany*

Marco Caliendo
University of Potsdam, IZA Bonn, DIW Berlin, IAB Nuremberg

Stefan Tübbicke
IAB Nuremberg

ABSTRACT

While a growing body of literature finds positive impacts of Start-Up Subsidies (SUS) on labor market outcomes of participants, little is known about how the design of these programs shapes their effectiveness and hence how to improve policy. As experimental variation in program design is unavailable, we exploit the 2011 reform of the current German SUS program for the unemployed which strengthened case-workers' discretionary power, increased entry requirements and reduced monetary support. We estimate the impact of the reform on the program's effectiveness using samples of participants and non-participants from before and after the reform. To control for time-constant unobserved heterogeneity as well as differential selection patterns based on observable characteristics over time, we combine Difference-in-Differences with inverse probability weighting using covariate balancing propensity scores. Holding participants' observed characteristics as well as macroeconomic conditions constant, the results suggest that the reform was successful in raising employment effects on average. As these findings may be contaminated by changes in selection patterns based on unobserved characteristics, we assess our results using simulation-based sensitivity analyses and find that our estimates are highly robust to changes in unobserved characteristics. Hence, the reform most likely had a positive impact on the effectiveness of the program, suggesting that increasing entry requirements and reducing support increased the program's impacts while reducing the cost per participant.

Keywords: Start-Up Subsidies, Institutions, Policy Reform, Difference-in-Differences
JEL Codes: J68, H43, L26

Corresponding author:
Marco Caliendo
University of Potsdam
Chair of Empirical Economics
August-Bebel-Str. 89
14482 Potsdam
GERMANY
E-mail: caliendo@uni-potsdam.de

* The authors thank Lutz Bellmann for helpful comments and the Institute for Employment Research (IAB) for cooperation and institutional support within the research project no. 1755.

1 Introduction

Start-up subsidies (SUS) for the unemployed are a particular kind of active labor market program (ALMP) that help to re-integrate jobseekers into the first labor market by granting temporary monetary support to those who are willing to start their own business. SUS have been studied for a variety of countries, generally finding positive and relatively large effects on participants' employment rates and income.[1] While these studies paint a relatively clear picture of positive treatment effects, the influence of institutional factors on the effectiveness of SUS has been a black box thus far. However, improving our understanding of the interplay between institutional details of SUS programs and their effectiveness is paramount to enhancing policy design, potentially allowing for greater program impacts and/or higher cost-effectiveness.

Ideally, the question of how to improve SUS policy could be answered with experimental variation in the design of programs. As such variation is not available, it is necessary to apply quasi-experimental approaches. Germany provides an ideal case study for this as SUS have a long tradition and there have been multiple reforms in recent times. In this paper, we aim to shed light on this issue by exploiting the most recent reform of the current German SUS program ("Gründungszuschuss", dubbed New Start-Up Subsidy) from 2011. The reform touched the most important features of the program by increasing caseworkers' discretionary power to reject applicants, reducing monetary support and raising entry requirements.[2] Comparing the estimated long-term effects on employment

[1]See, e.g. Tokila (2009) for Finland, Duhautois *et al.* (2015) for France, Wolff *et al.* (2016) for Germany, O'Leary (1999) for Hungary and Poland, Perry (2006) for New Zealand, Rodríguez-Planas and Jacob (2010) for Romania, Behrenz *et al.* (2016) for Sweden and Caliendo (2016) for a general overview.

[2]In addition to the institutional reform, budget cuts were also put in place, drastically reducing participation in the program.

and earnings before and after the reform shows substantial differences. While Caliendo *et al.* (2016) estimate pre-reform effects for participants from 2009, Caliendo and Tübbicke (2020) use very similar data on post-reform participants from 2012 and find that the post-reform effects on employment and earnings (up to 40 months after entering the program) are substantially higher. While such a gap lacks causal interpretation as it may be confounded by differential characteristics of participants or local macroeconomic conditions, it raises the question of whether the changes in institutional details of the program had an impact on its effectiveness. In order to provide evidence on this question, we combine the two data sets and use a semi-parametric difference-in-differences design based on propensity score weighting to control for confounding. Purging the gap in estimated effects from its association with observed characteristics and macroeconomic indicators, estimates imply that the reform raised the employment effects of the program. As these findings may be contaminated by changes in unobserved characteristics over time, we assess the sensitivity of our estimates with respect to this threat using simulations similar to Ichino *et al.* (2008) and find only small differences to our baseline results. Hence, our findings do not appear to be driven by "hidden bias". Overall, our results suggest that increasing entry requirements and reducing support raised the program's impacts by about 7-10% while reducing the cost per participant by about 20%. Without additional sources of variation, however, we cannot credibly disentangle this overall effect and deduce the separate effects of increasing entry requirements and reducing support. Nonetheless, our results indicate a substantial role of program design in shaping the effectiveness of SUS.

The paper is organized as follows. Section 2 offers an overview of the current German SUS program, its reform and some theoretical considerations regarding potential reform effects and relevant mechanisms. Section 3 describes the data and provides some descriptive statistics. Section 4 discusses our empirical approach and presents the results of our empirical analysis. Section 5 concludes.

2 The current German SUS program and its reform

The New Start-Up Subsidy (NSUS) program for the unemployed has been in place since 2006, receiving a major make-over at the end of 2011. In general, for individuals to be eligible for the program, they are required to have a minimum remaining unemployment benefit entitlement (RBE), which they forgo upon entry into the program. The scheme supports participants in two phases: during the first period, participants receive a monthly transfer equivalent to their unemployment benefits, topped up with a fixed sum of €300, which is supposed to cover social security contributions. In the optional second period, only the fixed sum is paid if the participant is re-approved by the local caseworker. For this, the participant needs to prove that the business is still running and the support is indeed needed. While this basic structure of the program remained the same, the reform altered the required RBE, benefit period durations as well as entitlement rules. Table 1 provides an overview of these features and their changes due to the reform.

[Insert Table 1 about here]

Before the reform, individuals needed a minimum of 90 days RBE to be eligible and the first benefit period lasted nine months, while benefits in the second period were paid for six months. The reform increased the minimum RBE to 150 days. In addition, it shortened the first benefit period to six months and increased the length of the second benefit period to nine months, leaving the total potential duration of support the same but reducing monetary support conditional on the level of unemployment benefits. While the average total subsidy paid to pre-reform participants was about €13,200, post-reform participants received about €2,800 ($\sim$ 20%) less in transfers (Bellmann *et al.*, 2018). In addition, the reform also changed the program's entitlement rules. While eligible applicants were entitled to the program by law before the reform, since 2011 caseworkers have had the right to reject eligible applicants if they deem the subsidy unnecessary in light of the individual's re-employment probabilities.

Anticipated Effects of the Reform Even if changes in entry requirements and support do not have an impact on SUS effectiveness, post-reform program effects may be different due to effect heterogeneity if the reform altered selection patterns. Holding participants' characteristics constant, the increase in the required remaining benefit entitlement means that they forgo a larger part of their potential unemployment benefit duration, which makes re-entry into unemployment less attractive once the program is over. This is likely to increase the program's effectiveness. By contrast, the reduction in the duration of the first benefit period may have a positive or negative effect overall. On the one hand, the shorter duration results in lower monetary support, potentially making the program less effective at reducing capital constraints for the unemployed. On the other hand, the shorter first benefit period may result in lower rates of moral hazard among participants, leading to ambiguous predictions on the effects of these institutional changes overall.

3 Data and Descriptives

Merging the datasets used by Caliendo *et al.* (2016) and Caliendo and Tübbicke (2020), we have access to rich data on participants and non-participants of the NSUS before and after the reform. Both datasets combine administrative data from the Integrated Employment Biographies (IEB) of the Federal Employment Agency and survey data collected up to 40 months after (hypothetical) entry into the program, providing us with very detailed information on socio-demographics, education, qualification, intergenerational transmission mechanisms, the entire labor market and ALMP treatment history of individuals as well as the necessary outcome data. Pre-reform participants entered the program in the first quarter of 2009, while post-reform participants joined between February and June 2012. Comparison individuals were chosen by a pre-matching strategy based on basic socio-demographic characteristics among other individuals who were unemployed for at least part of the respective time frame but did not join the SUS program during this period (see Caliendo *et al.*, 2016; Caliendo and Tübbicke, 2020, for more details on the data). Overall, our estimation sample comprises 5,039 individuals.

Descriptives Table 2 provides some descriptive statistics for our sample. Comparing covariate means shows that post-reform participants are older, more educated, more likely to be female and they reside in East Germany more often compared with pre-reform participants. Moreover, post-reform participants are less likely to enter unemployment from self-employment and they display a worse short-term labor market history relative to pre-reform participants. These differences in characteristics themselves may be interpreted as effects of the entitlement reform on selection into treatment to some degree. This view is supported by Bernhard and Grüttner (2015) who show that in contrast to the changes in participant characteristics there was no change of similar magnitude in basic characteristics of the unemployed population around the time of the reform.

[Insert Table 2 about here]

To analyze SUS effects on employment, we make use of three outcome measures: First, we use cumulated months in self- or regular employment over the entire observed 40 months after (hypothetical) entry. This allows us to estimate whether there was any impact of the reform on employment effects for participants overall. As such a finding may be driven by short-run positive lock-in effects during the subsidy period[3], we also use cumulated months in self- or regular employment after the subsidy has run out, i.e. between month 16 and month 40 after entry into the program. Doing so enables us to estimate reform effects net of possible lock-in effects during subsidy receipt. Moreover, we examine an indicator for self- or regular employment at month 40 after entry to analyze whether the reform had an impact until the very end of our observation period. Lastly, we use net monthly labor earnings 40 months after entry in order to also estimate effects on earnings in the next Section. The descriptive statistics in Table 2 reveal largely insignificant raw gaps between SUS participants' outcomes and significantly negative gaps for non-participants across the two time periods for all outcomes used. However, due to the pre-matching, the latter difference is not informative regarding the general unemployed population.

[3]While lock-in effects usually correspond to a negative effect for participants during program participation, the opposite is true here. Both participants and non-participants are unemployed in the month before the treatment starts, then participants join the program and change immediately to the successful state. That is, they leave unemployment and become self-employed. Hence, one should not overemphasize this large effect at the start of the self-employment spell.

4 Empirical Analysis

In this section, we aim to estimate the *change* in average treatment effects on the treated (ATT) of the SUS program due to the reform by holding participants' characteristics and the local macroeconomic environment constant. Using the standard notation of the potential outcomes framework (Roy, 1951; Rubin, 1974), the ATT in period t can be written as $\Delta_t = E(Y_t^1 - Y_t^0 \mid D = 1)$, where Y^1 and Y^0 denote potential outcomes with (without) subsidy receipt, i.e. $D = 1$ and $D = 0$. The subscript t takes the value of one for the post-reform period and zero otherwise. As is standard in the SUS literature, Caliendo *et al.* (2016) and Caliendo and Tübbicke (2020) estimate Δ_0 and Δ_1 using a selection-on-observables approach as

$$\Delta_t = E_{X_t}\{E(Y_t^1 \mid D = 1, X) - E(Y_t^0 \mid D = 0, X)\},$$

where the outer expectation re-weighs non-participants' outcomes according to the covariate distribution of participants, denoted by X_t. To set the stage for our main analysis, we re-estimate post- and pre-reform effects (Δ_1 and Δ_0) using inverse probability weighting based on the exactly-identified covariate balancing propensity scores (CBPS) by Imai and Ratkovic (2014), as their approach leads to an exact finite sample balance if a solution to the balancing problem exists. The CBPS methodology assumes a standard logistic distribution of the balancing weights and estimates them as a function of a linear index of covariates, where the parameters are chosen via the Generalized Method of Moments in order to minimize differences in covariate distributions across groups.[4] Trimming ob-

[4] Alternatively, a normal distribution may be specified. Similar to standard propensity score methods, this choice is largely unimportant.

servations with excessive weights as recommended by Imbens (2004) is not necessary as the largest weights obtained are much below the suggested 5 % threshold. The specification that we employ largely follows Caliendo and Tübbicke (2020) and uses a rich set of covariates including detailed information on socio-demographics, human capital, ALMP and labor market history as well as parental information. To obtain standard errors, we bootstrap the estimates using 999 replications (MacKinnon, 2006).

4.1 Main Analysis

The resulting estimates of Δ_1, Δ_0 and their difference are provided in columns (1) to (3) in Table 3. Our results support the finding that post-reform effects appear to be substantially larger, with the differences being statistically significant at all common levels for all outcomes. Estimated effects on cumulated employment increased by 3.6 months for the post-entry period and 2.4 months for the post-subsidy period. Effects on self- or regular employment after 40 months have increased by 8 percentage points and effects on earnings are larger by €255. While these estimates remove bias due to covariate differences between participants and non-participants *within* samples before and after the reform, there remain substantial differences in observed characteristics *between* participants over time of almost 8% in terms of mean absolute standardized bias (see bottom of Table 3).

[Insert Table (3) about here]

To control for these differences, we estimate the conditional gap in ATTs – denoted by $\Delta_1 - \Delta_0 \mid X$ – by re-weighing pre-reform conditional outcomes by the *post-reform* covariate distribution X_1, essentially imputing the ATT before the reform had participants had the same observed characteristics as participants after the reform. This approach can be interpreted as a semi-parametric DiD design (Abadie, 2005). Hence, we avoid functional form restrictions on the way in which observed covariates affect the outcomes of interest and we control for time-constant unobserved confounding. The resulting DiD estimates from column (4) suggest that the reform indeed influenced the SUS program's effectiveness in terms of employment outcomes, while no significant effect can be detected regarding net monthly labor income. More precisely, the estimates imply an increase in employment effects by 3.39 months over the post-entry period and an increase of 2.31 months when focusing on the post-subsidy period. Relative to the mean outcomes of post-reform participants, this represents a sizable impact of almost 10%. Lastly, the estimates suggest that effects on the likelihood of being in self- or regular employment 40 months after entry were also increased by seven percentage points, or about 7.5% relative to post-reform participants' mean. Hence, the results imply sizable effects of the reform on the program's effectiveness in terms of employment outcomes. A comparison of relative effect sizes for the different employment outcomes reveals that the effects are not driven by a stronger positive lock-in effect during the subsidy-period but by rather permanent effects. Even though the relative effect size is somewhat smaller at the end of the observation period, effects remain highly significant in economic terms. As these estimates may be biased if there are interaction effects between macroeconomic conditions and ATTs, we additionally control for regional fixed effects, the local unemployment rate, the vacancy rate as well as GDP per capita via parametric outcome regression on the re-weighed sample. Adding these controls – denoted by R – in column (5) leaves the estimates essentially unchanged, further supporting our findings of a positive effect of the reform. Thus, it appears that the effectiveness of the SUS program was improved by increasing the required benefit entitlement (RBE) and reducing monetary support. Even though we cannot disentangle the mechanism based on the data at hand, the results suggest that the positive effect of raising the RBE outweighed the potentially negative effects of reducing support.

4.2 Sensitivity Analysis

While our empirical approach is robust to time-constant unobserved heterogeneity due to the DiD methodology, estimates may be inconsistent if the reform affected not just the selection patterns regarding on observable characteristics but also regarding unobserved characteristics. As we cannot rule out this possibility, we run simulations similar to Ichino *et al.* (2008) in order to assess how effect estimates change under different assumptions with respect to the distribution of a relevant unobserved confounder. For simplicity, the existence of an unobserved binary confounder U is assumed. In order to perform the simulation, one has to take a stance on the set of probabilities p^t_{dy} defined as

$$p^t_{dy} = Pr\{U = 1 \mid D = d, Y^* = y, T = t\},$$

where t, d and $y \in \{0, 1\}$ and Y^* is an indicator the outcome Y being larger than the median.[5] This means the researcher has to specify the probability that an individual of treatment group D in period t with Y being above or below the median of the outcome distribution possess the characteristic U. In our set-up, this means that we have to choose $2^3 = 8$ parameters for each outcome to be inspected. To be concise, we limit the pre-

[5] Y^* is used in order to avoid having to specify a complete probability density function which would arise due to the continuous nature of Y. See Nannicini (2007) for details.

sentation of the simulations results to our preferred employment measure, i.e. cumulated months in self- or regular employment in the post-subsidy period. Results for the other outcome measures are similar and can be found in Table A.1 in the Appendix. As no prior knowledge about the proper choice of p^t_{dy} exists, we follow Ichino *et al.* (2008) and simulate unobserved confounders such that they mimic the distribution of important observed confounders. Based on the choice of parameters, we run $S = 1,000$ simulations. In each simulation replication, we generate a draw of the unobserved confounder and include it in our estimation of the CBPS. Resulting effect estimates are saved and mean simulated effects are displayed in Table 4. Standard errors are obtained based on the *within-* and *between-* imputation variance of the simulation (again, see Ichino *et al.*, 2008, for details). Moreover, in order to judge the influence of U on selection into treatment as well as the outcome, Table 4 also provides information on the average odds-ratio coefficients from a logistic regression of D and Y^* on X and U.

[Insert Table (4) about here]

We run six different simulation scenarios with parameters p^t_{dy} chosen as to reflect the distribution of indicators for being female, having an upper high school degree, parental self-employment, having been self-employed before entering unemployment as well as indicators for below median prior earnings and employment in the year before entering the program. Looking at average odds-ratio coefficients in columns three to six, one can say that these simulated variables have a strong influence on selection into treatment and/or success in terms of outcomes. Turning to average effect estimates displayed in column one, we see that none of the estimated effects taking the simulated confounder into account differ in any substantial way from our main employment effect estimate of 2.3 months.

Effects found for simulations based on the female indicator, parental self-employment as well as having low previous income essentially yield identical estimates. Evidence based on the distribution of self-employment before unemployment gives only a slightly smaller average effect even though the simulated confounder increases the odds of receiving substantially in both samples but has an opposite association with the outcome. The smallest effect obtained is found for the simulation based on the upper high school degree indicator. Here, the simulated confounder exerts a strongly positive effect on the odds of receiving treatment in the post-reform population, whereas having this characteristic lowers the odds of receiving treatment in the pre-reform era. Nonetheless, the estimates still suggest a positive and statistically significant impact of the reform on employment effects of SUS of over two months. While the simulation exercise does not provide evidence on how robust our estimates are to unobserved confounders that are distributed differently, the general lack of variation in our estimates based on simulations that mimic important observed confounders suggests that the estimates are not sensitive to "hidden bias".

5 Conclusion

This paper exploits the 2011 reform of the German NSUS for the unemployed to estimate the impact of changes in the institutional setup of SUS programs on their effectiveness, making a first contribution to filling an important research gap. The reform tightened entry requirements and reduced monetary support, leading to ambiguous predictions on the overall sign of reform effects. Based on rich administrative data combined with information gathered from surveys for both pre- and post-reform participants and non-participants, we estimate these effects on participants' employment and earnings using a semi-parametric difference-in-differences design. While our findings indicate that the reform increased employment effects for participants, we find no significant change for effects on earnings. Hence, it seems that the hypothesized positive effect of raising entry requirements outweighed the potentially negative effect of reducing support. Overall, our results suggest that increasing entry requirements and reducing support raised the program's impacts by about 7-10% while reducing the cost per participant by about 20%. In order to further advance our understanding of how SUS work, future experimental analysis of the effects of design features is paramount. The random assignment of institutional details would allow credibly disentangling effects, improving program design and participants' outcomes.

Declaration of competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *Review of Economic Studies*, **72** (1), 1–19.

Behrenz, L., Delander, L. and Månsson, J. (2016). Is Starting a Business a Sustainable way out of Unemployment? Treatment Effects of the Swedish Start-up Subsidy. *Journal of Labor Research*, **37** (4), 389–411.

Bellmann, L., Caliendo, M. and Tübbicke, S. (2018). The Post-Reform Effectiveness of the New German Start-Up Subsidy for the Unemployed. *LABOUR*, **32** (3), 293–319.

Bernhard, S. and Grüttner, M. (2015). *Der Gründungszuschuss nach der Reform: Eine qualitative Implementationsstudie zur Umsetzung der Reform in den Agenturen.* Forschungsbericht 4/2015, IAB, Nürnberg.

Caliendo, M. (2016). Start-up subsidies for the unemployed: Opportunities and limitations. *IZA World of Labor*, **200**, 1–11.

—, Künn, S. and Weissenberger, M. (2016). Personality traits and the evaluation of start-up subsidies. *European Economic Review*, **86**, 87 – 108.

— and Tübbicke, S. (2020). New evidence on long-term effects of start-up subsidies: Matching estimates and their robustness. *Empirical Economics*, **59** (4), 1605–1631.

Duhautois, R., Redor, D. and Desiage, L. (2015). Long Term Effect of Public Subsidies on Start-up Survival and Economic Performance: An Empirical Study with French Data. *Revue d'Économie Industrielle*, **149** (1), 11–41.

Ichino, A., Mealli, F. and Nannicini, T. (2008). From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity. *Journal of Applied Econometrics*, **23**, 305–327.

Imai, K. and Ratkovic, M. (2014). Covariate balancing propensity score. *Journal of the Royal Statistical Society Series B*, **76** (1), 243–263.

Imbens, G. (2004). Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *The Review of Economics and Statistics*, **86** (1), 4–29.

MacKinnon, J. G. (2006). Bootstrap Methods in Econometrics. *Economic Record*, **82**, 2–18.

Nannicini, T. (2007). A simulation-based sensitivity analysis for matching estimators. *Stata Journal*, **7** (3), 334–350.

O'Leary, C. J. (1999). *Promoting Self Employment Among the Unemployed in Hungary and Poland*. Working Paper 99-55, W.E. Upjohn Institute for Employment Research.

Perry, G. (2006). *Are Business Start-Up Subsidies Effective for the Unemployed: Evaluation of Enterprise Allowance*. Working paper, Auckland University of Technology.

Rodríguez-Planas, N. and Jacob, B. (2010). Evaluating active labor market programs in romania. *Empirical Economics*, **38** (1), 65–84.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, **3** (2), 135–146.

Rubin, D. (1974). Estimating Causal Effects of Treatments in Randomised and Nonrandomised Studies. *Journal of Educational Psychology*, **66** (5), 688–701.

Tokila, A. (2009). *Start-up grants and self-employment duration*. Working paper, School of Business and Economics, University of Jyväskylä.

Wolff, J., Nivorozhkin, A. and Bernhard, S. (2016). You can go your own way! The long-term effectiveness of a self-employment programme for welfare recipients in Germany. *International Journal of Social Welfare*, **25** (2), 136–148.

Tables and Figures

Table 1: Institutional Details of the New SUS Program

	Pre-Reform Program (8/1/2006-12/27/2011)	Post-Reform Program (12/28/2011-present)
Eligibility	Unemployed individuals with at least **90** days of UB remaining	Unemployed individuals with at least **150** days of UB remaining
Support	1st benefit period: **9** months of UB+€ 300	1st benefit period: **6** months of UB+€ 300
	2nd benefit period: **6** months € 300	2nd benefit period: **9** months € 300
Entitlement	1st period: **yes**	1st period: **no**
	2nd period: no	2nd period: no

Note: This table summarizes the institutional set-up of the New Start-Up Subsidy program in Germany before and after its reform at the end of 2011. The reform altered eligibility rules, the length of the two benefit periods as well as entitlement rules by giving caseworkers additional discretionary power. Abbreviations: UB = unemployment benefits.

Table 2: Selected Descriptives Statistics

	Post-reform		Pre-reform		*p*-values	
	Part.	Non-part.	Part.	Non-part.		
	(1)	(2)	(3)	(4)	(1)v(3)	(2)v(4)
	Part.	Non-part.	Part.	Non-part.		
A. *Covariate descriptives*						
Age in years	42.80	44.02	41.25	40.91	0.000	0.000
Share female	0.426	0.512	0.364	0.392	0.002	0.000
Share living in eastern Germany	0.333	0.369	0.220	0.261	0.000	0.000
Share with an upper high school degree	0.605	0.407	0.533	0.583	0.000	0.000
Share with self-employed parents	0.351	0.252	0.343	0.280	0.679	0.116
Share self-employed before unempl.	0.056	0.011	0.140	0.029	0.000	0.002
Months in employment before unempl.	7.768	6.686	9.406	9.380	0.000	0.000
Last daily real income (Euro)	84.04	63.36	77.12	77.88	0.003	0.000
B. *Outcome descriptives*						
Cumulated months in self- or regular employment						
Post-entry period $\sum_{m=1}^{40}$	38.48	23.16	38.33	27.25	0.457	0.000
Post-subsidy period $\sum_{m=16}^{40}$	23.81	15.50	23.53	18.32	0.107	0.000
Outcomes after 40 months						
Share in self- or reg. employment	0.936	0.673	0.924	0.735	0.221	0.001
Net monthly labor income (Euro)	2,257	1,038	2,239	1,397	0.815	0.000
Number of observations[a]	1,257	1,248	1,237	1,297		

Note: Reported are sample shares and covariate means with *p*-values based on *t*-tests of equal means. Due to the pre-matching of non-participants, their samples are not representative of the underlying unemployed population.
[a] The number of observations for the earnings variable is slightly lower due to item non-response.

Table 3: Estimation Results

	Preliminary Analysis			Main Analysis	
	Δ_1	Δ_0	$\Delta_1 - \Delta_0$	$\Delta_1 - \Delta_0 \mid X$	$\Delta_1 - \Delta_0 \mid X, R$
	(1)	(2)	(3)	(4)	(5)
Cumulated months in self- or regular employment					
Post-entry period $\sum_{m=1}^{40}$					
Estimate	12.33***	8.74***	3.59***	3.39***	3.43***
S.E.	(0.55)	(0.46)	(0.72)	(0.76)	(0.75)
Post-subsidy period $\sum_{m=16}^{40}$					
Estimate	6.39***	3.97***	2.42***	2.31***	2.32***
S.E.	(0.38)	(0.31)	(0.49)	(0.53)	(0.52)
Outcomes after 40 months					
Self- or regular employment					
Estimate	0.22***	0.15***	0.08***	0.07**	0.07**
S.E.	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Net monthly labor income (Euro)					
Estimate	966.6***	711.5***	255.1***	112.0	153.2
S.E.	(70.2)	(71.0)	(98.3)	(109.5)	(109.9)
Number of observations[a]	2,505	2,534	5,039	5,039	5,039
Specification					
Weights $\hat{E}_{X_t}$	X_1	X_0	X_1/X_0	X_1	X_1
Macro controls					✓
Covariate balance					
Mean standardized absolute bias in %					
SUS^{post} vs. NP^{post}	0.00	16.7	0.00	0.00	0.00
SUS^{pre} vs. NP^{pre}	7.22	0.00	0.00	0.00	0.00
SUS^{post} vs. SUS^{pre}	7.90	7.90	7.90	0.00	0.00

Note: The table shows the estimated effects. Estimates are obtained via a semi-parametric DiD approach based on covariate balancing propensity scores (Imai and Ratkovic, 2014). Standard errors are displayed underneath the point estimates. Standard errors are bootstrapped using 999 replications. ***/**/* indicate significance at the 1/5/10% level. Columns (1) and (2) re-estimate program effects for the post- and pre-reform program individually. Column (3) compares these two estimates. Column (4) adjusts this comparison for differential characteristics across participants. Column (5) adds parametric controls for regional macroeconomic factors.

[a] The number of observations for the earnings variable is slightly lower due to item non-response.

Table 4: Simulation Results

Outcome / Simulated confounder U	$\Delta_1 - \Delta_0 \mid X, R, U$	S.E.	Y^*_{post}	Y^*_{pre}	D_{post}	D_{pre}
	(1)	(2)	(3)	(4)	(5)	(6)
Cumulated months in self- or regular employment						
Post-subsidy period $\sum_{m=16}^{40}$						
Simulated confounder U to mimic ...						
Female	2.32***	0.40	0.93	0.76	0.71	0.89
Upper high school degree	2.13***	0.42	1.51	1.86	2.23	0.80
Self-employed parents	2.32***	0.40	1.11	1.23	1.61	1.35
Self-employed before unempl.	2.26***	0.43	0.75	1.68	6.16	6.01
Low employment in the year prior	2.16***	0.44	0.64	0.57	0.57	0.93
Low last income	2.31***	0.42	0.75	0.43	0.49	0.89

Note: The table shows estimation results based on the simulation set-up by Ichino *et al.* (2008). Average effect estimates and estimates on the odds of scoring above the median in terms of the outcome or receiving treatment are shown. Results on six scenarios are presented. Each scenario performs $S = 1,000$ simulations, generating an unobserved binary confounder U that mimics the distribution of an observed variable. Standard errors are obtained based on the *within-* and *between-* imputation variance of the simulation (see Ichino *et al.*, 2008, for details).

Appendix

Table A.1: Additional Simulation Results

Outcome / Simulated confounder U	$\Delta_1 - \Delta_0 \mid X, R, U$ (1)	S.E. (2)	Y^*_{post} (3)	Y^*_{pre} (4)	D_{post} (5)	D_{pre} (6)
Cumulated months in self- or regular employment						
Post-treatment period $\sum_{m=1}^{40}$						
Simulated confounder U to mimic ...						
Female	3.40***	0.58	0.78	0.71	0.71	0.90
Upper high school degree	3.37***	0.60	1.18	2.08	2.27	0.80
Self-employed parents	3.43***	0.58	1.10	1.09	1.62	1.36
Self-employed before unempl.	3.28***	0.61	2.45	3.01	5.98	5.97
Low employment in the year prior	3.35***	0.62	0.63	0.65	0.57	0.93
Low last income	3.47***	0.60	0.85	0.44	0.49	0.90
Self- or regular employment after 40 months						
Simulated confounder U to mimic ...						
Female	0.07***	0.02	1.05	0.56	0.70	0.90
Upper high school degree	0.06***	0.02	1.39	1.45	2.27	0.80
Self-employed parents	0.07***	0.02	0.87	1.19	1.62	1.35
Self-employed before unempl.	0.07***	0.02	0.80	2.00	6.10	5.97
Low employment in the year prior	0.05**	0.02	0.68	0.51	0.57	0.94
Low last income	0.06***	0.02	0.77	0.55	0.49	0.89
Net monthly labor income (Euro) after 40 months						
Simulated confounder U to mimic ...						
Female	156.8*	95.02	0.45	0.38	0.78	0.92
Upper high school degree	94.1	95.03	2.57	2.46	2.13	0.79
Self-employed parents	144.3	93.8	1.26	1.15	1.58	1.35
Self-employed before unempl.	124.2	96.43	1.30	1.51	6.35	6.12
Low employment in the year prior	164.1	100.7	0.66	0.61	0.60	0.94
Low last income	96.7	96.4	0.17	0.14	0.57	0.95

Note: The table shows estimation results based on the simulation set-up by Ichino *et al.* (2008). Shown are average effect estimates and estimates on the odds of scoring above the median in terms of the outcome or receiving treatment. Results on six scenarios for each of the remaining outcomes are presented. Each scenario performs $S = 1,000$ simulations, generating an unobserved binary confounder U that mimics the distribution of an observed variable. Standard errors are obtained based on the *within-* and *between-* imputation variance of the simulation (see Ichino *et al.*, 2008, for details).