Turning Unemployment into Self-Employment: Effectiveness and Efficiency of Two Start-Up Programmes*

Hans J. Baumgartner Marco Caliendo

DIW Berlin

Working Paper

This draft: May 31, 2007

Abstract

Turning unemployment into self-employment has become a major focus of German active labour market policy (ALMP) in recent years. If effective, this would not only reduce Germany's persistently high unemployment rate, but also increase its notoriously low self-employment rate. Empirical evidence on the effectiveness of such programmes is scarce. The contribution of the present paper is twofold: first, we evaluate the effectiveness of two start-up programmes for the unemployed. Our outcome variables include the probability of being employed, the probability of being unemployed, and personal income. Second, based on the results of this analysis, we conduct an efficiency analysis, i.e., we estimate whether the Federal Employment Agency has saved money by placing unemployed individuals in these programmes. Our results show that at the end of the observation period, both programmes are effective and one is also efficient. The considerable positive effects present a stark contrast to findings from evaluations of other German ALMP programmes in recent years. Hence, ALMP programmes aimed at moving the unemployed into self-employment may prove to be among the most effective, both in Germany and elsewhere.

Keywords: Start-Up Subsidies, Evaluation, Effectiveness, Efficiency, Self-Employment
JEL Classification: J68, C14, H43, M13

*The authors thank Arne Uhlendorff and Viktor Steiner for valuable comments and Steffen Künn for excellent research assistance. Corresponding author: Marco Caliendo, DIW Berlin, Dep. of Public Economics, Königin-Luise-Str. 5, 14195 Berlin, phone: +49-30-89789-154, fax: +49-30-89789-9154, e-mail: mcaliendo@diw.de. Marco Caliendo is also affiliated with IZA, Bonn, and IAB, Nuremberg.

1 Introduction

Turning unemployment into self-employment has become a major focus of German active labour market policy (ALMP) in recent years. Whereas the Federal Employment Agency (FEA) funded only 37,000 business start-ups by formerly unemployed individuals in 1994, the number was already above 350,000 in 2004 (approximately 250,000 in West Germany). This increase was driven, among other things, by a new programme known as the 'start-up subsidy' (SUS, *Existenzgründungszuschuss*), which was introduced in 2003 as part of the Hartz reforms. Unemployed individuals can now choose between this and a second programme, the 'bridging allowance' (BA, *Überbrückungsgeld*), which was already implemented in the late 1980s. The two programmes differ in their design, most importantly regarding the amount and duration of the subsidy. Whereas the BA pays recipients the same amount that they would have received in unemployment benefits for a period of six months (plus a lump sum to cover social security contributions), the SUS runs for three years, paying a lump sum of €600/month for the first year, €360/month for the second, and €240/month for the third. If successful, these programmes could potentially not only decrease Germany's persistently high unemployment rate, but increase its notoriously low self-employment rate as well. Looking at the FEA's spending on ALMP, we clearly see the increasing priority assigned to these programmes within the overall ALMP strategy. Whereas in 1994 only 0.6% of ALMP resources were allocated to these measures, in 2004 this number was 17.2%. This corresponds to annual spending of over €2.7 billion.

For all the aforementioned reasons, the high research interest in evaluating these programmes is unsurprising. However, empirical evidence on start-up aid is very rare, not only in Germany but also internationally. Meager (1996) summarises findings for five countries (Denmark, France, West Germany, UK and US) and concludes that the evidence presented does not allow a conclusive assessment of the overall effectiveness of such schemes. Existing papers usually focus either on survival rates of subsidised businesses, e.g., Cueto and Mato (2006), or compare start-ups by formerly unemployed people with start-ups which were not created out of unemployment (see, e.g., Pfeiffer and Reize, 2000). The present paper takes a different approach. Instead of comparing business start-ups by formerly unemployed individuals with other start-ups, we compare the labour market outcomes of the formerly unemployed entrepreneurs with other unemployed individuals. This approach is driven by the consideration that start-up subsidies form one component of ALMP, and their effectiveness should thus be compared to other ALMP programmes. In recent years, empirical evidence on the effectiveness of German ALMP has been constantly growing. Following the introduction of new legislation at the end of the 1990s (*Sozialgesetzbuch III*, Social Code III) and especially the Hartz reforms in 2002, the FEA was required to evaluate the effectiveness of its ALMP programmes. To fulfil this obligation, researchers were provided access to the FEA's administrative data and several programmes were evaluated. For example, Lechner, Miquel, and Wunsch (2005) and Biewen, Fitzenberger, Osikominu, and Völter (2006) evaluate the effectiveness of vocational training (VT) programmes, whereas Caliendo, Hujer, and Thomsen (2005) concentrate on job-creation schemes. The findings are negative for job-creation schemes and mixed for vocational training programmes, where due to high locking-in effects at the beginning of VT, positive effects appear only after some time.

The contribution of this paper is twofold: first, we evaluate the effectiveness of the two start-up programmes. Since the major goal of German ALMP is to avoid future unemployment and integrate unemployed individuals into the primary labour market, we concentrate on the outcome variables 'not unemployed' and 'in paid or self-employment', and in addition, we analyse the programme's effects on personal income. While most evaluations of ALMP stop at that point, we want to take the analysis a step further. Thus, in the second step, we conduct an efficiency analysis based on the results of the effectiveness analysis. This analysis is designed to answer the question of whether the FEA has saved money by helping people get out of unemployment and into self-employment (in contrast to financing their continued unemployment). It should be clear that the aim of this paper is not to compare the relative success of the two programmes, e.g., with respect to the success of the businesses themselves (number of employees, etc.). This is left to future studies.

Our analysis is based on a combination of administrative data from the FEA and a follow-up survey. The follow-up survey was necessary because 1) administrative data are only available with a certain time lag and 2) more importantly, they only contain information about employment for which social security contributions are compulsory, which is not the case for self-employment. The data contain approximately 3,100 participants in both programmes who founded a business in the third quarter of 2003 in West Germany.[1] The interviews took place at the beginning of 2005 and 2006, such that we observe individuals at least 28 months after programmes started. Whereas for BA this means we can monitor the employment paths of individuals for at least 22 months after the programme has ended, SUS was still ongoing at the end of our observation period. At this stage, participants in SUS were in their third year of participation and were receiving a reduced transfer payment. Hence, results for this programme are only preliminary and interpretation hinges on this drawback. Additionally, we have a group of unemployed individuals (approx. 2,300) who were eligible for either programme but did not choose to participate in the third quarter of 2003. This nonparticipant group will function as our comparison group.

Given this informative data set, we base our analysis on the conditional independence assumption and use a kernel matching estimators to estimate the treatment effects. To test the sensitivity of the results with respect to unobserved differences we also use a conditional difference-in-differences strategy as suggested by Heckman, Ichimura, Smith, and Todd (1998). The results show that at the end of our observation period both programmes are effective in terms of the above-mentioned outcome variables. Unemployment rates of participants are lower, and employment rates and personal income are higher when compared to nonparticipants. However, only one of the programmes—the bridging allowance—is also efficient in terms of the cost-benefit analysis.

The paper proceeds as follows. Section 2 gives a brief overview of the German labour market in the last decade, focusing on self-employment, unemployment, and active labour market policies, whereas Section 3 summarises previous empirical findings. Section 4 outlines our evaluation ap-

[1]We concentrate on West Germany in this paper because the labour market and especially self-employment dynamics in East Germany are quite different and have to be analysed separately.

proach, while Section 5 describes the data used for the analysis and discusses some implementation issues. Section 6 presents the results and Section 7 concludes.

2 Unemployment, Self-Employment and Start-Up Subsidies in Germany

Table 1 contains some summary statistics of the West German labour market. It can be seen that the self-employment rate has remained relatively stable over the last decade, fluctuating between 10 and 11% (relative to the workforce). Compared with other OECD countries, this is relatively low. Blanchflower (2000) refers to numbers for 1996 and shows that only Denmark, Luxembourg, Norway, and the United States have lower rates. On the other hand, the unemployment rate is persistently high, fluctuating between 7.3 and 9.1%.

To overcome this unemployment problem, the German government spends significant amounts on ALMP (approximately €12 billion in West Germany in 2004), including measures like vocational training programmes, job creation schemes, employment subsidies, and self-employment of formerly unemployed individuals.[2]

Table 1: Self-employment, Unemployment and Start-Up Subsidies in West Germany, 1994-2004

	1994	1998	1999	2000	2001	2002	2003	2004
Self-employed^a (in %)	10.4	10.6	10.4	10.2	10.3	10.4	10.6	11.0
Unemployed^a (in %)	8.1	9.1	8.4	7.5	7.3	7.9	8.8	8.8
Supported self-employment (Entries)								
BA (in thousand)	22.2	66.2	65.9	59.3	62.0	86.9	115.5	137.4
SUS (in thousand)	–	–	–	–	–	–	68.0	113.8
Total (in thousand)	22.2	66.2	65.9	59.3	62.0	86.9	183.5	251.1
Total^b (in %)	0.9	2.4	2.5	2.5	2.7	3.5	6.7	9.0
ALMP expenditure (in bn Euro)								
ALMP - Total	9.84	9.86	11.75	12.23	12.42	12.15	12.28	11.89
BA^c	0.06	0.43	0.55	0.53	0.58	0.73	1.09	1.37
SUS	–	–	–	–	–	–	0.18	0.67
Sup. self-empl. (total)	0.06	0.43	0.55	0.53	0.58	0.73	1.27	2.05
Sup. self-empl. (in %)	0.6	4.4	4.7	4.4	4.6	6.0	10.3	17.2

[a] Relative to the workforce.
[b] Relative to all unemployed.
[c] The figures for the years 1994-1998 are approximated.
Source: Bundesagentur für Arbeit, various issues.

From 1986 to 2002, the bridging allowance was the only programme providing support to unemployed individuals who wanted to start their own business. Its main goal is to cover basic costs of living and social security contributions during the initial stage of self-employment. BA supports

[2] For a recent overview of German active labour market policy see Caliendo and Steiner (2005).

the first six months of self-employment by providing the same amount that the recipient of a BA would have received if he or she had remained unemployed. Since the unemployment scheme also covers social security contributions including health insurance, retirement insurance, etc., a lump sum for social security is granted, equal to 68.5% of the unemployment support that would have been received in 2003, adjusted annually. Unemployed people are entitled to BA conditional on their business plan being approved externally, usually by the regional chamber of commerce. Thus, approval of an individual's application does not depend on the case manager at the local labour office.

In January 2003, an additional programme was introduced to support unemployed people in starting a new business. This 'start-up subsidy' was introduced as part of a large package of ALMP programmes introduced through the Hartz reforms.[3] The main goal of SUS is to secure the initial phase of self-employment. It focuses on the provision of social security to the newly self-employed person. The support is a lump sum of €600/month in the first year. A growth barrier is implemented in SUS such that the support is only granted if income is not expected to exceed €25,000 per year. The support shrinks to €360/month in the second year and €240/month in the third. In contrast to the BA, SUS recipients are obligated to pay into the legal pension insurance fund, and may claim a reduced rate for national health insurance (Koch and Wießner, 2003). When the SUS was introduced in 2003, applicants did not have to submit business plans for prior approval, but have been required to do so since November 2004, as is the case with the BA as well. See Table 2 for more details on both programmes.

Table 2: Design of the Programmes

	Bridging Allowance	**Start-Up Subsidy**
Entry conditions:	Unemployment benefit *entitlement* Approval of the business plan by an external source (e.g. chamber of commerce)	Unemployment benefit *receipt* Approval of the business required since November 2004
Support:	Participant receives UB for six months To cover social security liabilities, an additional lump sum of approx. 70% is granted	Participants receive a fixed sum of €600/month in the first year, €360/month (€240/month) in the second (third) year Claim has to be renewed every year, income is not allowed to exceed €25,000 per year
Other:	Social security is left at the individual's discretion	Participants are required to join the legal pension insurance and receive a reduced rate on the legal health insurance
Details:	§57(1) Social Code III.	§421 l Social Code III.

Hence, unemployed individuals can now choose between two programmes for help in starting their own business. Table 1 contains some information on participants and spending in measures promoting self-employment from 1994 to 2004. In 1994, about 1% of all unemployed individuals participated in BA, and the FEA spent 0.6% of their total resources for ALMP on BA. Due to a

[3]Wunderlich (2004) provides a thorough overview of the Hartz reforms.

legal change in 1995 that made it easier to receive a BA, these numbers increased steadily up to 2002, when 3.5% of the unemployed received a BA (6.0% of the spending). Table 1 also shows that the introduction of the SUS did not replace the BA, but did make self-employment significantly more attractive for the unemployed. In 2004, as much as 9% of Germany's unemployed participated in these two programmes together, thus absorbing a share of 17.2% of the total spending for ALMP.

Individuals planning to exit unemployment by entering self-employment can now choose between two alternative forms of start-up aid. One supports the first six months of self-employment by providing what the individual would have received in unemployment benefits plus a lump sum for social security contributions (BA), and the other provides a fixed and declining amount for the first three years of self-employment with the risk of losing the support if the growth barrier is exceeded (SUS). In this institutional framework, rational programme choice favours a BA if the unemployment benefits would be fairly high, and/or if the income generated through the start-up firm is expected to exceed €25,000.

3 Previous Empirical Findings

In contrast to other ALMP programmes such as vocational training or job-creation schemes, the empirical evidence on the effectiveness of start-up subsidies for the unemployed is rather scarce. This might be explained by the fact that in most countries start-up subsidies usually only form one small component of ALMP. In 2003, the EU-15 countries spent an average of 0.697% of their GDP on ALMP, but only 0.034% of GDP on start-up subsidies. That is, out of the total spending on ALMP only 4.8% was used for these incentives (European Commission, 2005). The numbers in the last section have shown that this has changed substantially in Germany.

The main indicators used for evaluating self-employment programmes are the survival rate, the number of jobs created directly by the new business, and the employability and income of participants. Additionally, it is usually of interest whether there have been deadweight losses or displacement effects.[4] Additionally, one has to define the comparison group. Some studies do not have a comparison group at all (and focus, e.g., solely on survival rates); others use start-ups by those who were not previously unemployed as a benchmark to compare the income of self-employed programme participants with the income of individuals in paid employment. We have already pointed out that we use a different approach in this paper, comparing the outcomes of participants with other unemployed individuals. In the following we give a brief overview of the findings in the literature on start-up subsidies for the unemployed, starting with some international evidence before turning to the results for Germany.

Meager, Bates, and Cowling (2003) evaluate business start-up subsidies to young people (18-30 years) in the UK. They not only look at the characteristics and survival of the start-ups but also compare the labour market outcomes of the participants with those of a comparison group. The comparison group is chosen to be in the same age category and then matched on three criteria

[4] A deadweight loss occurs when behaviour is not changed due to the programme, e.g., when unemployed individuals would also have entered self-employment in the absence of the subsidy. Displacement effects take place, e.g., when the businesses set up by the participants drive other existing (unsubsidised) businesses out of the market.

(gender, region and employment status immediately before the date when the matched person in the participant sample entered self-employment). Based on multinomial and standard logistic regressions the authors conclude that participating in the programme does not have any significant impacts on subsequent employment or earnings chances.

Perry (2006) uses difference-in-differences propensity score matching to evaluate the impact on males receiving an Enterprise Allowance grant in New Zealand between 1993 and 1995. This programme has been providing start-up subsidies since 1990 and can be seen as an integrated programme that provides business skills training as well as financial aid (for at least 26 weeks). The author's results (measured up to two years after participation) indicate statistically significant beneficial effects for the participants, where the outcome variable is 'not registered unemployed'.

Cueto and Mato (2006) analyse the success of self-employment subsidies in one region of Spain using a Cox Proportional Hazards Model. They look at the determinants of survival (duration) of self-employment and also estimate a competing risk model to distinguish between business failures and other reasons why businesses that had received support were closed, e.g., the because the individual had take a job and moved out of self-employment. Their study is based on data for individuals who received the subsidy between 1996 and 2000 and their labour market outcomes (still self-employed, unemployed, in paid employment) measured in December 2001. Hence, survival for 2-5 years can be observed and the survival is approximately 93% after two and 76% after five years.

Comparisons are difficult due to the heterogeneity of the institutional settings of the different programmes, the economic circumstances in the respective countries, and the indicators used. The assumed deadweight losses range from low to high and are usually based on survey information of the participants (Meager, 1996). What should be kept in mind here is that even if a participant would have started a business anyway—even without a subsidy—it is unclear whether it would have been equally successful. Displacement effects are hardly ever analysed and would require a macroeconomic framework.

Conclusive evidence for Germany is even harder to find. Pfeiffer and Reize (2000) use the ZEW Firm Start-Up Panel in their study to compare a group of start-ups founded between 1993 and 1995 by formerly unemployed recipients of a BA to a group of start-ups not subsidised by a BA. Assessing business survival and employment growth, they find different effects for West and East Germany. Whereas start-ups by the unemployed in the East German regions have a 6% lower one-year survival probability, no significant differences can be detected in West Germany. In terms of employment growth, subsidised start-ups by the unemployed are no different from non-subsidised start-ups. Reize (2004) uses the German Socio-Economic Panel (SOEP) and estimates competing risk models to model the paths out of unemployment. Comparing individuals moving into self-employment with those moving into paid employment shows that after four years, the unemployment risk is lower for the self-employed than for the other group. Both studies focus on the BA and have the problem of a rather small group of participants. Empirical evidence on the effectiveness of the SUS has not yet been produced since the programme is relatively new. In the next section, we turn to a description of our evaluation approach.

4 Identifying Average Treatment Effects

4.1 Fundamental Evaluation Problem and Selection Bias

We base our analysis on the potential outcome framework, also known as the Roy(1951)-Rubin(1974) model. The two potential outcomes are Y^1 (individual receives treatment, $D = 1$) and Y^0 (individual does not receive treatment, $D = 0$). The actually observed outcome for any individual i can be written as: $Y_i = Y_i^1 \cdot D_i + (1 - D_i) \cdot Y_i^0$. The treatment effect for each individual i is then defined as the difference between her potential outcomes: $\tau_i = Y_i^1 - Y_i^0$. Since we can never observe both potential outcomes for the same individual at the same time, the fundamental evaluation problem arises. We will focus on the most prominent evaluation parameter, which is the average treatment effect on the treated (ATT), and is given by:

$$E(Y^1 - Y^0 \mid D = 1). \tag{1}$$

To see how selection bias might arise, we cast the discussion in familiar econometric notation and write the potential outcomes as a function of observed (X) and unobserved (U^0, U^1) variables:

$$Y_{it}^1 = g_t^1(X_i) + U_{it}^1 \quad \text{and} \quad Y_{it}^0 = g_t^0(X_i) + U_{it}^0, \tag{2}$$

where the subscript t identifies the time period. The functions g^0 and g^1 represent the relationship between potential outcomes and the set of observable characteristics. U^0 and U^1 are error terms which have zero mean and are assumed to be uncorrelated with regressors X. For the familiar case of linear regression, the g functions specialise to $g^1(X) = X\beta_1$, and $g^0(X) = X\beta_0$ (see, e.g., (Heckman, Ichimura, and Todd, 1997)).

Heckman and Robb (1985a) note that the decision to participate in treatment may be determined by a prospective trainee, by a programme administrator, or both. Whatever the specific content of the rule, it can be described in terms of an index function framework. Let IN_i be an index of benefits to the relevant decision-maker from participating in the programme. It is a function of observed (Z_i) and unobserved (V_i) variables. Therefore:

$$IN_i = f(Z_i) + V_i. \tag{3}$$

In terms of this function $D_i = 1$ if $IN_i > 0$ and 0 otherwise. Except in case of randomised experiments, the assignment process to treatment is most probably not random. Consequently, the assignment process will lead to non-zero correlation between enrolment (D_i) and the outcome's error term (U^1, U^0). This may occur because of stochastic dependence between (U^1, U^0) and V_i or because of stochastic dependence between (U^1, U^0) and Z_i. In the former case we have selection on unobservables, and in the latter selection on observables (Heckman and Robb, 1985b). We will combine two evaluation methods—matching and difference-in-differences—to cover both possible sources of selection bias.

4.2 Matching under Unconfoundedness

Matching is based on the conditional independence (or unconfoundedness) assumption, which states that conditional on some covariates $W = (X, Z)$, the potential outcomes (Y^1, Y^0) are independent of D.[5] Since we are interested in ATT only, we only need to assume that Y^0 is independent of D, because the moments of the distribution of Y^1 for the treatment group are directly estimable. That is:

Assumption 1 *Unconfoundedness for Comparison Group:*

$$Y^0 \amalg D|W,$$

where $\amalg$ denotes independence. Clearly, this assumption may be a very strong one and has to be justified on a case-by-case basis, since the researcher needs to observe all variables that simultaneously influence participation and outcomes. We will do so in Section 5.2. Additionally, it has to be assumed that:

Assumption 2 *Weak Overlap:*

$$Pr(D = 1 \mid W) < 1,$$

for all W. This implies that there is a positive probability for all W of not participating, i.e., that there are no perfect predictors which determine participation. These assumptions are sufficient for identification of the ATT, which can be written as:

$$\tau_{ATT}^{MAT} = E(Y^1|W, D = 1) - E_W[E(Y^0|W, D = 0)|D = 1], \tag{4}$$

where the first term can be estimated from the treatment group and the second term from the mean outcomes of the matched comparison group. The outer expectation is taken over the distribution of W in the treatment group.

As matching on W can become hazardous when W is of high dimension ('curse of dimensionality'), Rosenbaum and Rubin (1983) suggest the use of balancing scores $b(W)$. These are functions of the relevant observed covariates W such that the conditional distribution of W given $b(W)$ is independent of the assignment to treatment, that is, $W \amalg D|b(W)$. The propensity score $P(W)$, i.e., the probability of participating in a programme, is one possible balancing score. For participants and nonparticipants with the same balancing score, the distributions of the covariates W are the same, i.e., they are balanced across the groups. Hence, assumption 1 can be re-written as $Y^0 \amalg D|P(W)$ and the new overlap condition is given by $Pr(D = 1 \mid P(W)) < 1$.

[5]See Imbens (2004) or Smith and Todd (2005) for recent overviews regarding matching methods.

4.3 Combining Matching with Difference-in-Differences

Even though we will argue in Section 5.2 that the CIA is most likely to hold in our setting, we will test the sensitivity of our results with respect to unobserved heterogeneity. The matching estimator described so far assumes that after conditioning on a set of observable characteristics, (mean) outcomes are independent of programme participation. The conditional DID or DID matching estimator relaxes this assumption and allows for unobservable but temporally invariant differences in outcomes between participants and nonparticipants. This is achieved by comparing the conditional before/after outcomes of participants with those of nonparticipants. DID matching was first suggested by Heckman, Ichimura, Smith, and Todd (1998). It extends the conventional DID estimator by defining outcomes conditional on the propensity score and using semiparametric methods to construct the differences. Therefore it is superior to DID as it does not impose linear functional form restrictions in estimating the conditional expectations of the outcome variable, and it re-weights the observations according to the weighting function of the matching estimator (Smith and Todd, 2005). If the parameter of interest is ATT, the DID propensity score matching estimator is based on the following identifying assumption:

$$E[Y_t^0 - Y_{t'}^0 | P(W), D = 1] = E[Y_t^0 - Y_{t'}^0 | P(W), D = 0], \tag{5}$$

where (t) is the post-treatment and (t') the pre-treatment period. It also requires the common support condition to hold and can be written as:

$$\tau_{ATT}^{CDID} = E(Y_t^1 - Y_{t'}^0 | P(W), D = 1) - E(Y_t^0 - Y_{t'}^0 | P(W), D = 0). \tag{6}$$

5 Implementing the Estimators

Having discussed our evaluation approach in the previous section, we now present details on the implementation of the propensity score matching estimator. Caliendo and Kopeinig (2005) provide an extensive overview of the issues arising when implementing matching estimators. They point out that a crucial step is to discuss the likely validity of the underlying CIA. Hence, we deal with this issue in Section 5.2, after having presented the data and some sample characteristics in Section 5.1. This will be followed by an estimation of the propensity score in 5.3, the choice of the matching algorithm in 5.4, and a discussion of matching quality in 5.5.

5.1 Data and Some Descriptives

We use a unique data set which combines administrative data from the FEA with survey data. For the administrative part we use data based on the 'Integrated Labour Market Biographies' (ILMB, *Integrierte Erwerbs-Biographien*) of the FEA, containing relevant register data from four sources: employment history, unemployment support recipience, participation in active labour market measures, and job seeker history. One drawback of the ILMB data is that employment history covers only employment that is subject to social security contributions. Since this is not the case for self-employment, the register data does not provide any information on the employment status and/or income of self-employed individuals. A second drawback is that the ILMB data is usually only available with a certain time lag. Hence, to get information about the success in self-employment for a reasonable time period, we enriched the ILMB data with information from a computer-assisted telephone interview.

To do so, we randomly drew participants from each programme who became self-employed in the third quarter of 2003. Since we wanted to compare them with nonparticipants, we had to choose a comparison group. Choosing such a group is a heavily discussed topic in the recent evaluation literature. Although participation in ALMP programmes is not mandatory in Germany, the majority of unemployed persons participate at some point in time. Thus, comparing participants to individuals who never participate is inadequate, since it can be assumed that the latter group is particularly selective.[6] Sianesi (2004) discusses this problem for Sweden and argues that those who never participate did not enter a programme because they had already found a job. Additionally, since we did not know the future employment/participation status of the comparison group before the interviews took place, we restricted this comparison group to those who were unemployed in the third quarter of 2003, eligible for participation in either of the two programmes, but did not join a programme in this quarter. What should be kept in mind is that these comparison group members might participate in some ALMP programme after this quarter.[7]

To minimise the survey costs we used a crude propensity score matching approach to select somewhat similar unemployed individuals.[8] These individuals were interviewed twice. The first interview took place in January/February 2005 and the second in January/February 2006. This enables us to observe the labour market activity of individuals for at least 28 months after programmes started. We compiled a sample of 3,100 individuals who had started a new business out of unemployment. Of these, 1,082 individuals received a SUS and 2,018 received BA. Additionally, a control group of 2,296 nonparticipants was assembled.

Table A.1 in the Appendix contains detailed descriptive statistics for all the available variables, differentiated by treatment status and gender. To abbreviate the discussion, we focus here on the most relevant variables and discuss differences between participants in both programmes and nonparticipants. What should be kept in mind is the non-random sample of nonparticipants. Since we used a crude matching approach to make individuals similar, the nonparticipant sample does not represent a random sample of unemployed individuals. Clearly, this does not affect our estimation and interpretation strategy but should be kept in mind when interpreting the differences. Table 3 contains sample means of selected variables and in addition results from a t-test of mean equality between participants and nonparticipants, where p_1 (p_2) refers to a test between nonparticipants and participants in SUS (BA).

[6] Furthermore, it should be noted that using individuals who are observed to never participate in the programmes as the comparison group may invalidate the conditional independence assumption due to conditioning on future outcomes (see discussion in Fredriksson and Johansson, 2004).

[7] The actual number of nonparticipants who participated in any ALMP programme after this quarter is rather low. It is approximately 5% after 12, 7% after 18 and around 10% after 24 months.

[8] For details on this pre-matching approach and the construction of the data see Caliendo, Steiner, and Baumgartner (2005).

Table 3: Selected Descriptives and Results of t-Tests

	Men				Women			
	NP	SUS	BA		NP	SUS	BA	
	Mean	Mean	Mean	p_1	Mean	Mean	Mean	p_1
Variable	SD	SD	SD	p_2	SD	SD	SD	p_2
Number of observations	1,448	811	1,207		848	704	378	
Qualificational Variables								
School Degree								
No Degree	0.02	0.04	0.01	0.04	0.01	0.01	0.01	0.47
	(0.14)	(0.19)	(0.12)	0.22	(0.10)	(0.08)	(0.07)	0.36
Upper Secondary Schooling	0.24	0.18	0.26	0.00	0.32	0.27	0.40	0.04
	(0.43)	(0.38)	(0.44)	0.16	(0.47)	(0.44)	(0.49)	0.01
Job Qualification								
High-Qualified	0.20	0.12	0.24	0.00	0.22	0.17	0.33	0.01
	(0.40)	(0.32)	(0.42)	0.01	(0.41)	(0.37)	(0.47)	0.00
Unskilled	0.18	0.27	0.14	0.00	0.15	0.20	0.08	0.02
	(0.38)	(0.44)	(0.35)	0.02	(0.36)	(0.40)	(0.27)	0.00
Labour Market History								
Previous Unemployment Duration								
$<$ 3 months	0.24	0.30	0.32	0.00	0.24	0.34	0.61	0.00
	(0.42)	(0.46)	(0.47)	0.00	(0.43)	(0.47)	(0.47)	0.00
$>$ 12 months	0.17	0.21	0.13	0.03	0.18	0.16	0.12	0.24
	(0.38)	(0.41)	(0.33)	0.00	(0.39)	(0.37)	(0.32)	0.00
No. of months in employment in 2002	6.69	5.52	7.79	0.00	6.35	6.02	7.65	0.20
	(5.03)	(4.93)	(4.66)	0.00	(5.15)	(5.04)	(4.73)	0.00
Avgerage daily earnings in 2002 (in €)	46.02	27.39	64.07	0.00	30.75	22.25	50.12	0.00
	(43.85)	(29.69)	(47.77)	0.00	(34.27)	(25.13)	(42.69)	0.00
Daily Unemployment Transfer (in €)	31.92	23.33	38.82	0.00	21.53	17.25	29.76	0.00
	(14.03)	(10.99)	(14.97)	0.00	(11.45)	(8.97)	(13.16)	0.00
Remaining Time of UB (in months)	6.32	4.72	7.31	0.00	5.57	5.02	6.83	0.07
	(6.34)	(5.55)	(6.24)	0.00	(5.99)	(5.88)	(6.07)	0.00

Note: All variables are measured one month before program start. Standard deviations are in parentheses. p-values refer to t-tests of mean equality in the variables between participants in the start-up subsidy (SUS) and nonparticipants (p_1) and participants in bridging allowance (BA) and nonparticipants (p_2).

A first glance at the number of observations reveals clear gender differences in participation in both programmes. Whereas the male-female ratio is about 3:1 for BA, it is nearly 1:1 for the SUS. Further differences arise when looking at qualifications. Comparing the participants' qualifications either by highest school-leaving degree or the variable 'job qualifications', an assessment by the placement officer in the local labour office, we see that BA participants are more highly qualified. For example, the share of individuals who had completed upper secondary schooling is quite high for participants in BA (26% of men / 40% of women) and rather low for participants in SUS (18% of men / 27% of women). Job qualifications show a similar picture. Here, 24% of the male and 33% of the female participants in BA are ranked as highly qualified, whereas this is only true for 12% (17%) of the male (female) participants in SUS.

Based on that, it is hardly surprising that participants in BA programmes also have a more favourable labour market history. Not only were they less frequently found among the long-term unemployed before starting a programme; they also had higher and longer claims for unemployment benefits. Differences are substantial: for example, male BA recipients received unemployment support amounting to €38.80/day before starting a programme whereas SUS recipients received only €23.30/day. It is also worth mentioning that the remaining period of benefit entitlement differed significantly between the two groups (approximately seven months for BA recipients and five for SUS recipients).

Given the relatively stable participant structure in the BA programme since the introduction of the SUS, one can argue that the SUS attracts a different 'clientele' for self-employment. In general it can be stated that participants in SUS are less qualified (when compared to BA participants), and that this programme is frequently used by women. We will discuss the available variables in more detail in the next section, where we also discuss the validity of the CIA.

5.2 Validity of the CIA

The CIA is in general a very strong assumption and the applicability of the matching estimator depends crucially on its plausibility. Blundell, Dearden, and Sianesi (2005) argue that the plausibility of such an assumption should always be discussed on a case-by-case basis. Only variables that influence the participation decision and the outcome variable simultaneously should be included in the matching procedure. Hence, economic theory, a sound knowledge of previous research, and information about the institutional setting should guide the researcher in specifying the model (see, e.g., Smith and Todd, 2005 or Sianesi, 2004). Both economic theory and previous empirical findings highlight the importance of socio-demographic and qualificational variables. Regarding the first category we can use variables such as age, marital status, number of children, nationality (German or foreigner), and health restrictions. Additionally, we also use information whether individuals want to work full-time or part-time, and hence we might be able to approximate the labour market flexibility of these individuals.

A second class of variables (qualification variables) refers to the human capital of the individual, which is also a crucially important determinant of labour market prospects. The attributes available are school degree, job qualification, and work experience. Furthermore, as pointed out by Heckman and Smith (1999), unemployment dynamics and labour market history play a major role in driving outcomes and programme participation. Hence, we use career variables describing the individual's labour market history. The available data in this regard is quite extensive. We have a nearly complete seven-year labour market history including information about the months spent in employment or unemployment. Additionally we know the daily earnings from employment and the amount of daily unemployment benefits. Furthermore, we can draw on the duration of the last unemployment spell, the number of (unsuccessful) placement propositions, the employment status before unemployment, and the previous profession.

Heckman, Ichimura, Smith, and Todd (1998) also emphasise the importance of drawing treatment and comparison groups from the same local labour market and giving them the same questionnaire. Since we use administrative data from the same sources for participants and nonparticipants, the latter point is not a problem in our data. To account for the situation on the local labour market, we use a classification of similar and comparable labour office districts derived by the FEA.

Nine different clusters can be identified for West Germany.[9]

Finally, the institutional structure and the selection process into programmes provide further guidance in selecting the relevant variables. As we have seen from the discussion in Section 2, the two programmes differ among other things in the size of the subsidy. Whereas the SUS is a lump sum, the BA depends on the amount of the unemployment benefits. Hence, we include the daily unemployment transfer payment before the start of the programme as an explanatory variable. In contrast to many other studies we are also able to include the remaining duration of unemployment benefits, which probably plays a determining role in these individuals' decision.[10] Based on this exhaustive data, we argue that the CIA holds in our application. However, we also test the sensitivity of the results with respect to time-invariant unobserved differences between participants and nonparticipants.

5.3 Estimation of the Propensity Score and Common Support

Since the choice probabilities are not known a priori, we have to replace them with an estimate. To do so, we estimate binary conditional probabilities for both programmes versus nonparticipation. Since we estimate the effects separately for men and women, we are left with four logit estimations. The results can be found in Table A.2 in the Appendix. To ensure the comparability between the estimates we choose the same covariates for each combination and both genders. We do not interpret the results of the propensity score estimation, since we only use this estimation to reduce the dimensionality problem. One has to remember that the group of participants and nonparticipants are already quite similar due to the construction of the data (see Section 5).

The distribution of the propensity score is depicted in Figure 1. A visual analysis already suggests that the overlap between the group of participants and nonparticipants is sufficient in general. Nevertheless, there are some parts of the distribution (starting approximately at a propensity score value of 0.7) where the mass of comparison individuals is quite thin. This is especially true for female participants in BA. However, by using the usual 'Minmax' criterion, where treated individuals are excluded from the sample whose propensity score lies above the highest propensity score in the comparison group, only 13 individuals are dropped overall.[11]

5.4 Matching Details

Several matching procedures have been suggested in the literature, such as nearest-neighbour or kernel matching.[12] To introduce them, a more general notation is needed: let I_0, and I_1 denote the

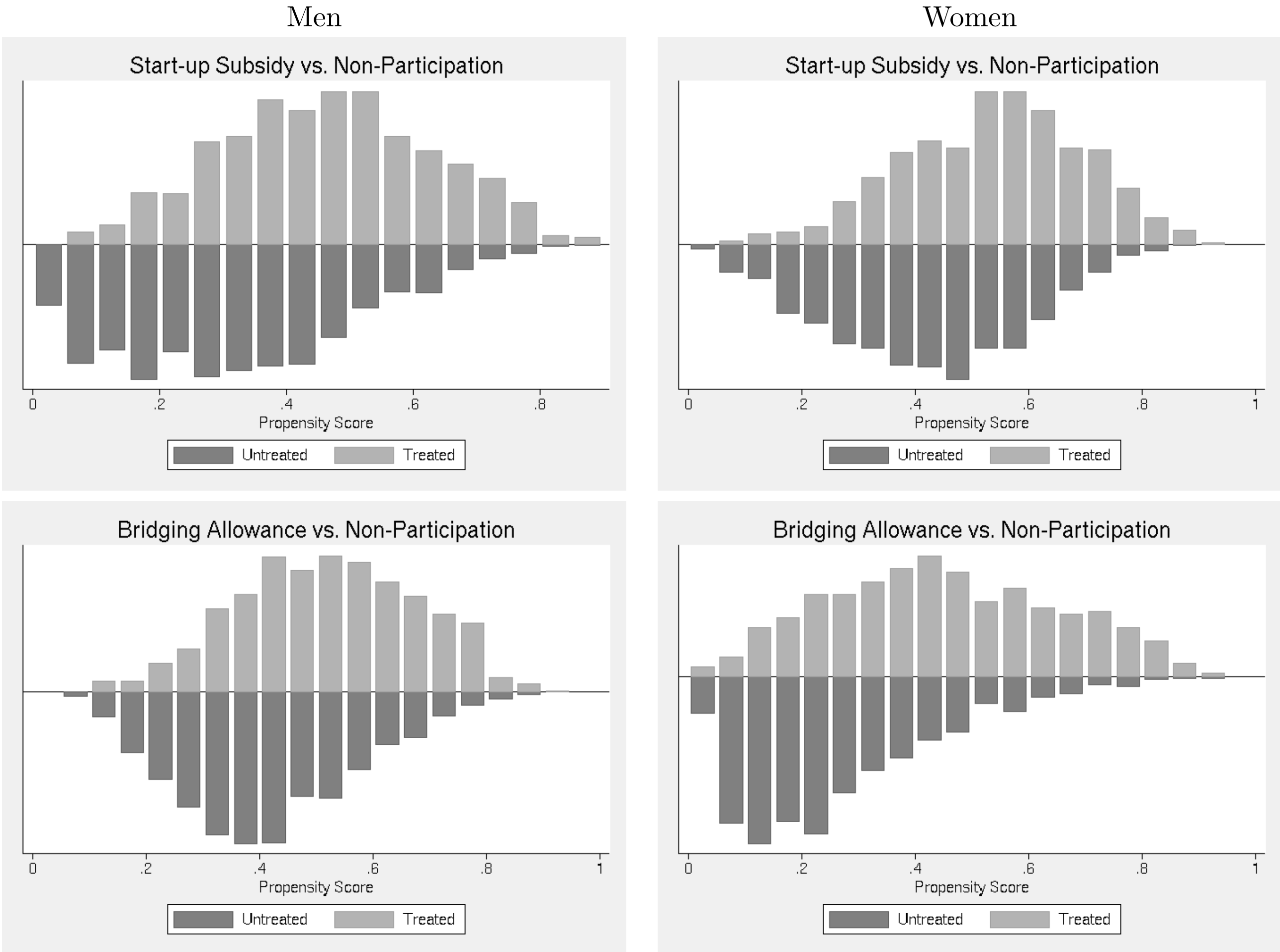
[9] This classification was undertaken by a project group of the FEA (Blien *et al.*, 2004) whose aim was to enhance the comparability of the labour office districts for a more efficient allocation of funds. It categorises the 181 German labour office districts into twelve comparable clusters. The comparability of the labour office districts is built upon several labour market characteristics, where the most important criteria are the underemployment rate and the corrected population density.

[10] Lechner and Wunsch (2006) evaluate the effectiveness of ALMP (excluding start-up subsidies) in East Germany using a very similar set of variables.

[11] We also test the sensitivity of the results with respect to more strict imposition of the common support requirement, e.g., by dropping 5%(10%) of the individuals where the overlap between participants and nonparticipants is especially low. It turns out that the results are not sensitive.

[12] See Heckman, Ichimura, Smith, and Todd (1998), Smith and Todd (2005), and Imbens (2004) for overviews.

Figure 1: Distribution of the Propensity Scores – Common Support[1]



Note: Propensity score is estimated according to the specification in Table A.2. Participants are depicted in the upper half, nonparticipants in the lower half of each figure.

set of indices for nonparticipants and participants. We estimate the effect of treatment for each treated observation $i \in I_1$ in the treatment group by contrasting her outcome with treatment with a weighted average of control group observations $j \in I_0$ in the following way:

$$\Delta^{MAT} = \frac{1}{N_1} \sum_{i \in I_1} [Y_i^1 - \sum_{j \in I_0} W_{N_0}(i,j) Y_j^0], \tag{7}$$

where N_0 is the number of observations in the control group I_0 and N_1 is the number of observations in the treatment group I_1. Matching estimators differ in the weights attached to the members of the comparison group (Heckman, Ichimura, Smith, and Todd, 1998), where $W_{N_0}(i,j)$ is the weight placed on the j-th individual from the comparison group in constructing the counterfactual for the i-th individual of the treatment group. The weights always satisfy $\sum_j W_{N_0}(i,j) = 1, \forall i$, that is, the total weight of all controls sums up to one for each treated individual. Matching estimators differ in how the neighbourhood is defined and the weights are constructed, e.g., with nearest-neighbour matching, only the closest neighbour is used to construct the counterfactual outcome. Kernel matching (KM), on the other hand, is a non-parametric matching estimator that uses (nearly) all units in the control group to construct a match for each programme participant. One major advantage of these approaches is the lower variance which is achieved because more information is used for constructing counterfactual outcomes. Since our treatment and comparison groups are rather small, we will focus now and in the later empirical application on this method.[13]

Heckman, Ichimura, and Todd (1998) derive the asymptotic distribution of these estimators and show that bootstrapping is valid to draw inference for this matching method. This is an additional advantage since it allows us to circumvent the issues raised by Abadie and Imbens (2006), pointing out that bootstrap methods are invalid for NN matching. It is worth noting that if weights from a symmetric, nonnegative, unimodal kernel are used, the average places higher weight on persons close in terms of P_i and lower weight on more distant observations. Kernel matching sets $A_i = I_0$ and uses the following weights:

$$W_{N_0}^{KM}(i,j) = \frac{G_{ij}}{\sum_{k \in I_0} G_{ik}}, \tag{8}$$

where $G_{ik} = G[(P_i - P_k)/h]$ is a kernel that downweights distant observations from P_i and h is a bandwidth parameter (Heckman, Ichimura, Smith, and Todd, 1998).[14] Before applying kernel matching, assumptions have to be made regarding the choice of the kernel function and the bandwidth parameter h. The choice of the kernel appears to be relatively unimportant in practice (see, e.g., DiNardo and Tobias (2001) or Jones, Marron, and Sheather (1996)).

What is seen as more important in the non-parametric literature is the choice of the bandwidth parameter h. Silverman (1986) and Pagan and Ullah (1999) note that there is little to choose between various kernel functions, whereas results depend more on h with the following trade-off arising: high values of h yield a smoother estimated density function, producing a better fit and a decreasing variance between the estimated and the true underlying density function. On the other hand, underlying features may be smoothed away by a large h, leading to a biased estimate. The choice of h is therefore a compromise between a small variance and an unbiased estimate of the true density function. Instead of using a 'rule of thumb' as proposed by Silverman (1986), we use cross-validation (CV) as suggested in Black and Smith (2004) and Galdo (2005) to choose h. CV methods are based on the principle of optimizing the out-of-sample predictive ability of the selected estimator. Here, we use a leave-one-out CV principle that drops the jth unit in the comparison group and forms the counterfactual $\widehat{Y}_{0j}$ for that unit using the $N_0 - 1$ observations left in the comparison group (Stone, 1974). Repeating the process for all comparison units, and given the fact that each estimation does not include the jth unit, this represents an out-of-sample forecast. Then, the bandwidth is chosen which minimises the mean square error (Galdo, 2005). More details and most importantly, the chosen bandwidth parameters can be found in Table A.3 in the Appendix. We will use these bandwidth parameters for the further empirical analysis.[15]

[13]However, we will also show that our results are not sensitive to the matching algorithm chosen.

[14]h satisfies $\lim_{N_0 \to \infty} h = 0$. See Heckman, Ichimura, and Todd (1998) for precise conditions on the rate of convergence needed for consistency and asymptotic normality of the kernel matching estimator.

[15]Estimations are done using the `PSMATCH2` Stata ado-package by Leuven and Sianesi (2003).

5.5 Matching Quality

To test if the matching procedure is able to balance all the covariates we ran a standardised difference (SD) test (Rosenbaum and Rubin, 1985). This is a suitable indicator to assess the distance in marginal distributions of the W-variables. For each covariate W it is defined as the difference of sample means in the treated and matched control subsamples as a percentage of the square root of the average of sample variances in both groups. This is a common approach used in many evaluation studies, including those by Lechner (1999), Sianesi (2004) and Caliendo, Hujer, and Thomsen (2005). Table 4 shows the mean standardised difference (MSD), i.e., the mean of the SD over all covariates before and after the matching took place.

Table 4: Matching Quality — Some Indicators

Variable	Start-up Subsidy		Bridging Allowance	
	Men	Women	Men	Women
MSD - Before Matching	13.049	7.780	12.658	18.577
MSD - After Matching	1.375	2.133	1.303	2.612
R^2 - Before Matching	0.127	0.094	0.082	0.150
R^2 - After Matching	0.003	0.007	0.003	0.008
χ^2 - Before Matching	0.000	0.000	0.000	0.000
χ^2 - After Matching	1.000	1.000	1.000	1.000
Participants off support	1	7	4	1

Note: Mean standardised difference (MSD) has been calculated as an unweighted average of the standardised difference of all covariates. Standardised difference before matching calculated as: $100 \cdot (\overline{W}_1 - \overline{W}_0)/\{\sqrt{(V_1(W) + V_0(W))/2}\}$ and standardised difference after matching calculated as: $100 \cdot (\overline{W}_{1M} - \overline{W}_{0M})/\{\sqrt{(V_1(W) + V_0(W))/2}\}$.

It can be seen that the MSD before matching lies between 7.8% for women and 13.0% for men in SUS and even between 12.7% (men) and 18.6% (women) in BA. The matching procedure is able to balance the distribution of the covariates very well, especially for men, where the MSD after matching lies around 1.3%. For women in SUS, the MSD after matching is 2.1%; for women in BA it is 2.6%. In general, it is not sufficient to look at the MSD if one wants to judge the quality of the matching procedure. Instead a careful look at the SD for each variable is necessary, which, in our case, showed very satisfying results.[16]

Additionally Sianesi (2004) suggests re-estimating the propensity score on the matched sample (i.e., on the participants and matched nonparticipants) and comparing the pseudo-R^2's before and after matching. After matching there should be no systematic differences in the distribution of the covariates between the two groups. Therefore, the pseudo-R^2 after matching should be fairly low. As the results from Table 4 show, this is true for our estimation. The results of the F-tests point in the same direction, indicating a joint significance of all regressors before, but not after matching. Overall, these are satisfying results and show that the matching procedure was successful

[16] Detailed results are available on request by the authors. The highest SD after matching in a single variable lies at 4.0% for men in SUS and 4.0% for men in BA. For women matching quality is slightly worse and the highest SD after matching lies around 7.5% for women in SUS and 7.4% for women in BA.

in balancing the covariates between treated individuals and members from the comparison group. Hence, we move on to the presentation of the results.

6 Results

The presentation of the results will be split in two parts. First we discuss the effectiveness of the two programmes in relation to nonparticipation. Three potential outcome variables are of crucial interest here. First, we want to know if programme participation lowers the risk of returning to unemployment. To this end, we construct a variable that treats registered unemployment as a failure and all possible other states as a success (outcome variable A). Since avoiding unemployment is one of the two major goals of German ALMP, this allows us to compare the effectiveness of the programmes in reaching this goal. A second aim is integration into regular, stable employment. Hence, we construct a second outcome variable which treats ongoing self-employment and regular paid employment as a success (outcome variable B). The combination of these two labour market states is important for comparing outcomes between participants in either of the two programmes (who are mainly self-employed at the time of the interview) and nonparticipants (who are more likely in regular employment). Finally, we also assess the effects of the programmes on the personal income of participants. In the second step, we use our results on effectiveness to conduct a cost-benefit analysis. The question to be answered here is whether the FEA saved money by assigning individuals to either of the two programmes (in contrast to providing continued unemployment support).

6.1 Effectiveness in Terms of Employment Status and Income

Effects on the Employment Status over Time: Figure 2 presents the treatment effects over time, where the upper panel relates to outcome variable A (not unemployed) and the lower part to outcome variable B (self-employed or in regular employment). Effects for men (women) are depicted on the left (right) side of each row. Rows 1 and 3 show the effects of participating in SUS vs. nonparticipation, whereas rows 2 and 4 show the effects of BA (vs. nonparticipation).

Effects start in the first month after the treatment has begun. Before starting the interpretation one has to note the following: a look at both figures shows a strong positive effect at the beginning of our observation period. This can be seen as a 'positive locking-in effect'. Whereas a locking-in effect usually corresponds to a negative effect during participation in a programme—for example, vocational training—the findings for our programmes are the opposite. Both participants and nonparticipants are unemployed in the month before the treatment starts, then participants join the programme and change immediately to the 'hoped-for' state. That is, they leave unemployment and become self-employed, which is viewed as a success for both outcome variables. Hence, one should not overemphasise this large effect at the start of the self-employment spell. BA runs out after six months, and a reasonable interpretation should start there. Clearly, for the three-year-long SUS, the problem is that participants may receive aid during the complete observation period, interfering with interpretation. However, after 12 months, the transfer payment is reduced from €600 to €360 and after 24 months it is reduced further to €240. Since this reduced payment is hardly sufficient to cover social security contributions, it gives us us an initial idea of the success of the newly self-employed.

Let us start the discussion with the first outcome variable, that is, the probability of not being unemployed. In the first months after treatment starts, we have very high positive effects for both programmes, lying well above 60 percentage points, irrespective of programme and gender. This means, for example, that the unemployment probability of participants in SUS or BA is about 60 percentage points lower than the unemployment probability of nonparticipants. Clearly, results at that point have to be interpreted with care, since both programmes are still ongoing. The effects show a negative time trend, where the paths of the programmes are very similar up to month six. After that, the transfer payment for participants in BA terminates and the effects plunge. The downward trend continues but the rate of decrease is much lower. At the end of our observation period, that is, 28 months after programmes have started, we get an effect of 16.8 percentage points for male and 16.8 percentage points for female participants in BA. If we look at the effect of SUS versus nonparticipation, the downward trend is much smoother, spiking somewhat in month 12, but decreasing relatively constantly to an effect of 28.2 percentage points for males and 17.6 percentage points for females in month 28.[17]

Looking at the lower part of Figure 2 shows a similar pattern but on a higher level. Remember that this is the effect for being in regular or self-employment. Effects already start at a much higher level, around 80 percentage points, and remain higher throughout the whole observation period. In $t+28$ we have a positive effect for males which lies at 36.8 percentage points for participants in SUS and 21.4 percentage points for participants in BA. Even more extreme differences can be found for women, where the effect for SUS lies at 44.6 percentage points and 39.6 percentage points for BA. This is a strong indication that both programmes are not only effective in avoiding unemployment but that they also give individuals much higher chances of remaining employed (either in paid or self-employment). The strong differences in both outcome variables can be explained by the fact that outcome variable A only treats registered unemployment as a failure. When individuals retreat from the labour market—and this might be especially relevant for women—they are not counted as a failure. Hence, the second outcome variable, only treating individuals as a success if they are in employment, has more explanatory power.

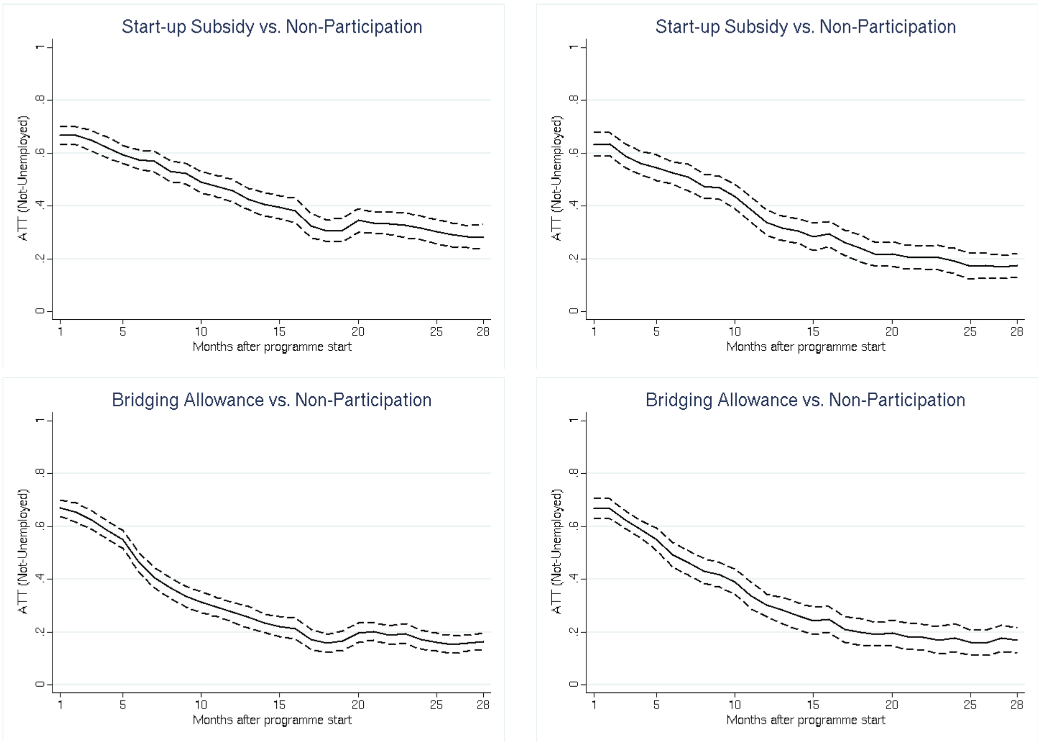
Cumulated Effects: Table 5 contains the cumulative effects over time, i.e., the cumulative monthly effects over the observation period. For the outcome variable 'not unemployed' this shows the difference in months spent in unemployment between participants and nonparticipants. It can be seen that male participants in SUS spend roughly 12.2 months less in unemployment than non-

[17]The dip in the effects, especially for men, between months 16 and 20, is caused by a change in the interview information. Individuals were interviewed twice, in 2005 and 2006. Months 16 to 20 might involve a time overlap between the first and second interview and might be prone to recall errors. Hence, information for these months should be interpreted with care. For the overall interpretation, especially when moving towards the end of the observation period, this should not pose any problems.

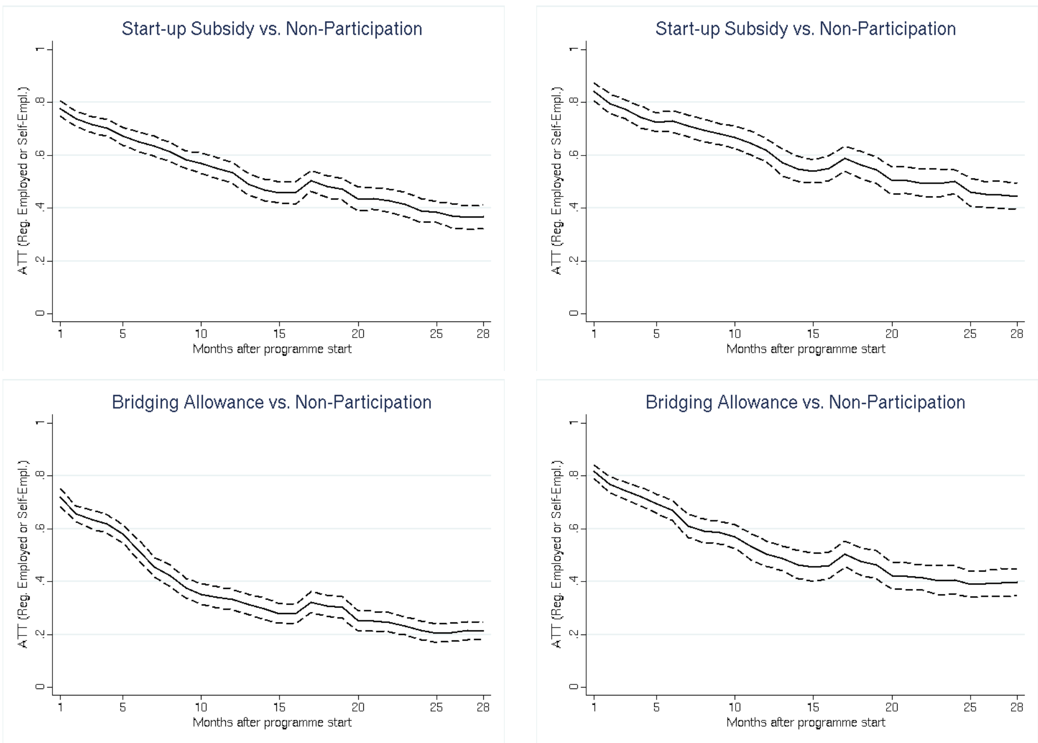
Figure 2: Treatment Effects over Time

Men Women

Outcome Variable A: Not Unemployed



Outcome variable B: Employed or Self-Employed



Note: Estimations are based on kernel matching as described in Section 5.4. Bootstrapped standard errors are based on 200 replications.

participants. For female participants in SUS the effect is approximately 9.7 months. The cumulative effect for participants in BA is slightly lower, at 8.6 months for men and 9.1 months for women. We have already discussed that the effects for the outcome variable 'self-employment or paid employment' are even higher, which is also reflected by the cumulative effects of around 14.7 (16.9) months for men (women) in SUS and 10.2 (14.8) months for men (women) in BA.

Table 5: Cumulated Effects - Matching and Conditional DiD

	Start-Up Subsidy				Bridging Allowance			
	Men		Women		Men		Women	
Outcome	Effect	s.e	Effect	s.e	Effect	s.e	Effect	s.e
Outcome Variable A: Not Unemployed (in months)								
Matching	12.19	(0.418)	9.72	(0.493)	8.55	(0.362)	9.13	(0.491)
DiD-1	11.82	(0.792)	9.26	(0.768)	8.46	(0.403)	6.79	(0.742)
DiD-2	11.97	(0.503)	9.80	(0.611)	8.63	(0.372)	7.96	(0.521)
DiD-3	12.03	(0.488)	9.18	(0.631)	8.39	(0.358)	7.96	(0.520)
Outcome Variable B: Employed or Self-Employed (in months)								
Matching	14.66	(0.474)	16.87	(0.496)	10.17	(0.382)	14.76	(0.505)
DiD-1	14.61	(0.944)	15.85	(1.332)	9.83	(0.930)	6.75	(1.347)
DiD-2	14.50	(0.636)	16.63	(0.764)	10.17	(0.599)	11.67	(0.917)
DiD-3	14.77	(0.791)	16.09	(0.906)	9.83	(0.683)	9.83	(1.012)

Note: Matching estimates are based on kernel matching as discussed in Section 5.4. Standard errors (in parentheses) are based on 200 bootstrap replications.
Reference level for DiD 1: Total month not spend in unemployment (outcome variable A) and spend in regular employment (outcome variable B) between 1997 and 2002.
Reference level for DiD 2: Same as DiD-1, but for the time period 2000-2002.
Reference level for DiD 3: Same as DiD-1, but for the time period 1997-1999.

As outlined in Section 4.3 we also tested the sensitivity of our results with respect to time-invariant unobserved heterogeneity by using a conditional difference-in-differences approach. Before using such an approach, one has to determine the reference level for the before/after difference. We choose three different time periods for the comparison. In the first approach we use the time period from 1997 to 2002, that is, the six-year employment history before entering the programme. For the first outcome variable, we sum the months not spent in unemployment, whereas for the second, we sum the months spent in paid employment. Additionally, we restrict the reference period to the latest three years, that is, the time period 2000-2002, and the earliest three years, that is, the time period 1997-1999.

Looking at the table, we see that the results are remarkably stable. For example, the effect on outcome variable B for men in SUS was 14.66 months with the matching approach and varies between 14.50 and 14.77 months with the CDID approaches. For women in SUS and men in BA the variation is slightly higher, but still negligible. This shows that additionally controlling for possible unobserved differences between participants and nonparticipants did not add much information for our estimates. This can be seen as evidence of the validity of the CIA in our context. Results are less favourable when looking at the smallest group under observation, that is, women in BA. Here the matching estimates are 9.13 months (not unemployed) and 14.76 months (regular or self-employed) respectively. The CDID results, however, vary from 6.8 to 8.0 in the first and 6.8 and 11.7 in the second case. This indicates that unobservable differences between the group of female participants in BA and nonparticipants remain even after matching. Given the fact that the CDID estimates are smaller than the matching estimates, one could argue that there are unobserved factors that drive not only the participation probability but also labour market outcomes. Hence, we have to treat these effects with caution.

Effects on the Personal Income: After having established that participants in both programmes are more likely to be employed and less likely to be unemployed than nonparticipants, we now investigate whether participants also earn more money. The questionnaire from January/February 2006 contained several questions related to individuals' personal income which allow us to generate two income-related outcome variables. The most relevant one is monthly income from self-employment or paid employment. This is the labour income that we are mainly interested in and that will be the focus of the analysis. However, since it is often argued that differences between (low) labour income and unemployment benefits are especially low in Germany, we will also look at the total personal income of individuals, that is, including support such as unemployment benefits.

Table 6: Effects on Monthly Income - Matching and Conditional DiD

	Start-up Subsidy				Bridging Allowance			
	Men		Women		Men		Women	
Outcome	Effect	s.e	Effect	s.e	Effect	s.e	Effect	s.e
Effect on Monthly Income from Self-Employment/Regular Employment (in €)								
Matching	596.27	(68.00)	298.96	(66.77)	770.69	(96.22)	975.80	(115.19)
DiD 1	601.44	(67.04)	295.57	(67.97)	768.72	(84.13)	738.01	(108.72)
DiD 2	586.95	(76.58)	289.80	(65.28)	742.37	(95.54)	429.69	(113.86)
DiD 3	586.76	(72.22)	316.32	(78.26)	770.97	(96.72)	510.58	(117.38)
Effect on Total Monthly Income (in €)								
Matching	465.99	(64.36)	237.46	(66.58)	639.23	(82.20)	950.87	(111.04)
DiD 1	471.17	(66.83)	234.07	(68.73)	637.26	(82.18)	713.08	(109.23)
DiD 2	456.67	(61.83)	228.29	(65.19)	610.90	(91.14)	404.76	(114.74)
DiD 3	456.48	(72.48)	254.81	(71.09)	639.51	(94.35)	485.65	(125.22)

Note: Matching estimates are based on kernel matching as discussed in Section 5.4. Standard errors (in parentheses) are based on 200 bootstrap replications.
Reference level for DiD 1: Unemployment Benefit before programme start.
Reference level for DiD 2: Average monthly income in 2002.
Reference level for DiD 3: Average monthly income from regular employment in 2002.

Table 6 contains the results for both outcome variables. Once again, we first present the results from matching estimates before presenting CDID results. For the DID procedure we use three reference levels: 1) The monthly unemployment benefits before the programme started, 2) the average monthly income in 2002 and 3) the average monthly income from regular employment in 2002. It is quite striking that all participants have significantly higher incomes than nonparticipants for both possible outcome variables. The upper half of Table 6 reveals that male participants in SUS earn around €600 per month more than their counterparts in the comparison group. Once again, the

CDID does not add much information to the matching estimates since all estimates range between €586 and €601. For female participants, the effect is much lower—between €290 and €316—but still significant. The effects for the participants in BA is even higher. Male participants earn about €770 more per month. For females, we once again have the problem that matching and CDID results differ significantly, making it hard to draw relevant policy conclusions.

Hence, we can conclude that participating in either of the two programmes has helped individuals to earn more money at the end of our observation period. This stays true even if we use the total personal income of individuals as an outcome variable, where we additionally take unemployment benefits and other government transfers into account.

6.2 Efficiency Analysis

So far, we have analysed the effectiveness of both programmes with respect to employment status and personal income. We have concluded that both programmes are effective, i.e., they increase both the probability of employment and income, and decrease the probability of unemployment. Whereas most of the evaluation studies of ALMP stop at that point, as mentioned above, we want to take the analysis a step further. Having established that participation is beneficial for participants, we now analyse whether the programme is beneficial for the provider, that is, the Federal Employment Agency. To do so, we conduct a basic four-step cost-benefit analysis, which we will explain briefly:

1. **Cumulated Effects:** We use the cumulative effects for the whole observation period and the outcome variable 'not unemployed' as a starting point. Let CE denote the cumulative effect, that is, the number of months less that a participant spends in unemployment than a nonparticipant.

2. **Average Savings per Participant in Months:** If we think about the money the FEA saved by placing individuals in these programmes, we have to take into account the remaining period of benefit entitlement (RBE). Clearly, if the cumulative effect is, for instance, six months for a certain group, but the RBE for the same group is only four months, the FEA only saves four unemployment months. Hence, we use the following decision rule to determine the average saving (AS) in unemployment months:

$$\begin{aligned} AS &= CE \text{ if } CE \leq RBE \\ &= RBE \text{ if } CE > RBE \end{aligned}$$

3. **Reduced Spending:** To put a monetary value on the savings of the FEA, we multiply AS by the average level of unemployment benefits in the group of treated individuals ($\overline{UB}$). Note that we add 70% on top of this value, since the FEA also covers social security contributions of individuals, which amount to approximately 70%.

4. **Monetary Efficiency:** Finally, to get the monetary efficiency we contrast AS with the direct costs of the programmes. For the SUS, these costs are fixed (€600/month in the first year, etc.) and depend only on the number of months an individual stays in self-employment. Hence,

they can be directly calculated from the number of months spent in self-employment. For BA they depend on the individual UB in the participants' group and are estimated accordingly.

Several things have to be noted about our approach. First of all, we only consider direct costs and benefits associated with the programmes. The direct costs only include programme costs that arise through payment of subsidies to SUS or BA participants. We do not consider any administrative costs that arise, e.g., through counseling services provided by the local employment office to the unemployed. If these costs are higher (lower) for participants than for nonparticipants, our approach over-(under)estimates the monetary efficiency. On the benefit side, we only consider the reductions in spending achieved through discontinuation of participants' unemployment benefits. It should be clear that unemployed individuals in Germany receive means-tested benefits after their entitlement to UB expires. These payments are usually not borne by the FEA but by other authorities. Hence, our approach underestimates the monetary efficiency. Finally, we also do not take into account the additional tax revenues of the entrepreneurs' businesses or the fact that some of these businesses might generate additional jobs. Clearly, these assumptions are necessary to facilitate our estimation of the direct costs and benefits, but a more thorough CBA should take these points into account. For the moment, we are willing to make these simplifying assumptions to get a first impression of the monetary efficiency of both programmes.

Table 7: Cost-Benefit Analysis - Results

Variable	Start-up Subsidy		Bridging Allowance	
	Men	Women	Men	Women
Cumulated Treatment Effect (in months)	12.19	9.72	8.55	9.13
Unemployment Benefits				
Remaining Time (in months)	4.60	4.94	7.24	6.74
Monthly Level (in €)	707.30	527.88	1,176.37	904.99
Direct Costs of the Treatment (in €)	11,285.58	11,591.54	11,955.71	9,161.89
Monetary Efficiency	-5,759.02	-7,155.53	2,524.19	1,203.75
	(94.30)	(77.71)	(107.10)	(254.40)

Note: Standard errors of the monetary efficiency (in parentheses) are based on 200 bootstrap replications.

Table 7 contains the results and more information on each of the four steps. The first line replicates the results from Table 5 showing the cumulative treatment effects in months. Rows two and three contain the remaining time of benefit entitlement (in months) and the monthly unemployment benefits (in euros). Two things should be noted immediately. First, participants in BA have on average much more time remaining in which they are entitled to unemployment benefits. Whereas this figure is below five months for men and women in SUS, male participants in BA have a remaining 7.2 months and female participants 6.7 months. Second, the level of UB is much higher for BA participants, reaching €1,171 per month for males and €900 per month for females. Participants in the SUS are entitled to unemployment benefits of €708 (men) and €530 (women). The final ingredient in the cost-benefit analysis are the direct costs of participation. Since the SUS subsidy does not depend on individual characteristics, costs are nearly the same for both genders, amounting to €11,300 for men and €11,600 for women. The difference can be explained by the fact that a slightly higher proportion of female participants stays in self-employment in our observation period. For the BA, on the other hand, costs depend directly on the individuals' unemployment benefits and are therefore higher for men (€11,900) than for women (€9,100).

Based on that, we get a clearly negative monetary efficiency for participants in SUS. Even though the cumulative effects are quite high for this group, the effect is dominated by the low remaining time of benefit entitlement and the relatively low level of benefits they would receive. For example, the direct costs for the FEA for female participants in SUS would have been only €4,500 (€530 plus 70% to cover social security per month for five months) had they not entered the programme. Compared with the direct costs of the programme, this amounts to a monetary efficiency of -€7,148. Since men in SUS have a higher benefit level, i.e., would incur higher costs on the FEA, the monetary efficiency is slightly better, but still negative at -€5,750. For participants in BA, however, we get a positive monetary efficiency of €2,490 for men and €1,178 for women.

These findings clearly show that BA support to unemployed people starting their own business has not only helped them to enhance their employment status and earn more income (when compared to nonparticipants), but has also saved the FEA money, decreasing its spending on unemployment benefits. For the SUS, our findings are not as encouraging. We have to keep in mind that the end of our observation period is 28 months after inception of the programmes. Hence, participants who continue in self-employment and do not earn more than €25,000 per year will receive further support of €240 for eight months. Clearly, this adds to the direct programme costs, but will not affect the savings of the FEA, resulting in even worse estimates for the monetary efficiency of SUS.

7 Conclusion

The aim of this paper has been to evaluate the effectiveness and efficiency of two active labour market programmes in Germany designed to encourage unemployed people to become entrepreneurs. These programmes have the potential not only to combat Germany's problem of persistently high unemployment, but also to increase its notoriously low self-employment rate. Our analysis is based on a dataset that combines administrative with survey data and allows us to follow the employment paths of individuals for up to 28 months after programmes have started. For the first programme under consideration—the bridging allowance—we observed participants for 22 months after the programme ended. However, participants in the second programme—the start-up subsidy—are in their third year of participation at the end of our observation period, and most likely will still receive further support (although at a reduced rate). Therefore, the results for SUS have to be treated as preliminary.

We have evaluated the effectiveness of both programmes relative to nonparticipation. To this end we used a kernel matching estimator and a conditional difference-and-differences estimator. Three outcome variables were of major interest. The first was 'not unemployed', corresponding to one of the main aims of the FEA. The second one combines the two possible labour market states 'in self-employment' and 'in paid employment' into one success criterion. The results indicate that both programmes are successful: at the end of our observation period, the unemployment rate of participants in BA was approximately 17 percentage points lower than that of nonparticipants, and for participants in SUS, around 18 percentage points lower for women and as much as 29 percentage points lower for men. Additionally, both the probability of being in self-employment and/or paid employment and the personal income are significantly higher for participants.

Based on the results of the effectiveness analysis we also conducted a basic cost-benefit analysis. Our results show that BA funding of individuals starting self-employment has not only helped them to enhance their employment status and earn more income (when compared to nonparticipants), but has also saved the FEA money by reducing its spending on unemployment benefits. For the SUS, the findings are not as encouraging, and result in a negative monetary efficiency.

Having said that, we can conclude that this is one of the first studies that allows inferences to be drawn about the effects of the start-up programmes that comprise part of Germany's ALMP. In contrast to other German ALMP programmes that have been evaluated recently (including job creation schemes and vocational training programmes), we find considerable positive effects for these two programmes. Hence, programmes aimed at turning the unemployed into entrepreneurs may be among the most promising for active labour market policy, both in Germany and elsewhere.

To allow more precise policy recommendations, further research is needed. First of all, the relative effects of both programmes should be estimated, which would allow their respective designs to be judged, as well as their suitability for different target groups. Additionally, it would be of interest to look at the development of the start-ups in terms of turnover and number of jobs directly created. Such an investigation would also enable a more extensive cost-benefit analysis taking not only the direct costs and benefits but also the indirect ones into account.

References

Abadie, A., and G. Imbens (2006): "On the Failure of the Bootstrap for Matching Estimators," Working Paper, Harvard University.

Biewen, M., B. Fitzenberger, A. Osikominu, and R. Völter (2006): "Employment Effects of Short and Medium Term Further Training Programs in Germany in the Early 2000s," Working Paper, Goethe-University Frankfurt.

Black, D., and J. Smith (2004): "How Robust is the Evidence on the Effects of the College Quality? Evidence from Matching," *Journal of Econometrics*, 121(1), 99–124.

Blanchflower, D. (2000): "Self-employment in OECD countries," *Labour Economics*, 7, 471–505.

Blien, U., F. Hirschenauer, M. Arendt, H. J. Braun, D.-M. Gunst, S. Kilcioglu, H. Kleinschmidt, M. Musati, H. Ross, D. Vollkommer, and J. Wein (2004): "Typisierung von Bezirken der Agenturen der Arbeit," *Zeitschrift für Arbeitsmarktforschung*, 37(2), 146–175.

Blundell, R., L. Dearden, and B. Sianesi (2005): "Evaluating the Impact of Education on Earnings in the UK: Models, Methods and Results from the NCDS," *Journal of the Royal Statistical Society, Series A*, 168(3), 473–512.

Bundesagentur für Arbeit (various issues): *Arbeitsmarkt*. Nürnberg.

Caliendo, M., R. Hujer, and S. Thomsen (2005): "The Employment Effects of Job Creation Schemes in Germany - A Microeconometric Evaluation," IZA Discussion Paper No. 1512, Bonn.

Caliendo, M., and S. Kopeinig (2005): "Some Practical Guidance for the Implementation of Propensity Score Matching," IZA Discussion Paper No. 1588, Bonn, forthcoming in: *Journal of Economic Surveys*.

Caliendo, M., and V. Steiner (2005): "Aktive Arbeitsmarktpolitik in Deutschland: Bestandsaufnahme und Bewertung der mikroökonomischen Evaluationsergebnisse," *Zeitschrift für Arbeitsmarktforschung / Journal for Labour Market Research*, 38(2-3), 396–418.

Caliendo, M., V. Steiner, and H. J. Baumgartner (2005): "Kapitel IV: Mikroökonometrische Analysen," Project Report to the BMWA – Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission, Wirksamkeit der Instrumente: Existenzgründungen (Modul 1e).

Cueto, B., and J. Mato (2006): "An analysis of self-employment subsidies with duration models," *Applied Economics*, 38, 23–32.

DiNardo, J., and J. Tobias (2001): "Nonparametric Density and Regression Estimation," *Journal of Economic Perspectives*, 15(4), 11–28.

European Commission (2005): *Labour Market Policy - Expenditure and Participants, Data 2003*. Office for Official Publications of the European Communities, Luxembourg.

Fredriksson, P., and P. Johansson (2004): "Dynamic Treatment Assignment - The Consequences for Evaluations Using Observational Data," Discussion Paper No. 1062, IZA.

Galdo, J. (2005): "Bandwidth Selection and the Estimation of Treatment Effects with Non-Experimental Data," Working paper, Syracuse University.

Heckman, J., H. Ichimura, J. Smith, and P. Todd (1998): "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66(5), 1017–1098.

Heckman, J., H. Ichimura, and P. Todd (1997): "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, 64(4), 605–654.

——— (1998): "Matching as an Econometric Evaluation Estimator," *Review of Economic Studies*, 65(2), 261–294.

Heckman, J., and R. Robb (1985a): "Alternative Methods for Evaluating the Impact of Interventions - An Overview," *Journal of Econometrics*, 30(1-2), 239–267.

——— (1985b): "Alternative Models for Evaluating the Impact of Interventions," in *Longitudinal Analysis of Labor Market Data*, ed. by J. Heckman, and B. Singer, pp. 156–245. Cambridge University Press.

Heckman, J., and J. Smith (1999): "The Pre-Program Earnings Dip and the Determinants of Participation in a Social Program: Implications for Simple Program Evaluation Strategies," *Economic Journal*, 109(457), 313–348.

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

Jones, M., S. Marron, and S. Sheather (1996): "A Brief Survey of Bandwidth Selection for Density Estimation," *Journal of the American Statistical Association*, 91(433), 401–407.

Koch, S., and F. Wiessner (2003): "Wer die Wahl hat, hat die Qual," *IAB Kurzbericht*, (2).

Lechner, M. (1999): "Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany After Unification," *Journal of Business Economic Statistics*, 17(1), 74–90.

Lechner, M., R. Miquel, and C. Wunsch (2005): "The Curse and Blessing of Training the Unemployed in a Changing Economy: The Case of East Germany After Unification," Discussion Paper No. 1684, IZA, Bonn.

Lechner, M., and C. Wunsch (2006): "Active Labour Market Policy in East Germany: Waiting for the Economy to Take Off," Discusssion Paper No. 2363, IZA, Bonn.

Leuven, E., and B. Sianesi (2003): "PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing," Software, http://ideas.repec.org/c/boc/bocode/s432001.html.

Meager, M., P. Bates, and M. Cowling (2003): "An Evaluation of Business Start-Up Support for Young People," *National Institute Economic Review*, 186, 59–72.

Meager, N. (1996): "From Unemployement to Self-employement: Labour Market Policies for Business Start-up," in *International Handbook of Labour Market Policy and Evaluation*, ed. by G. Schmidt, J. O'Reilly, and K. Schömann, pp. 489–519. Edward Elgar.

Pagan, A., and A. Ullah (1999): *Nonparametric Econometrics*. Cambridge University Press, Cambridge.

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

Pfeiffer, F., and F. Reize (2000): "Business Start-Ups by the Unemployed - an Econometric Analysis Based on Firm Data," *Labour Economics*, 7, 629–663.

Reize, F. (2004): *Leaving Unemploment for Self-Employment, An Empirical Study*. ZEW Economic Studies Vol. 25, Mannheim.

Rosenbaum, P., and D. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70(1), 41–50.

——— (1985): "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score," *The American Statistican*, 39(1), 33–38.

Roy, A. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), 135–145.

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

SIANESI, B. (2004): "An Evaluation of the Swedish System of Active Labour Market Programmes in the 1990s," *The Review of Economics and Statistics*, 86(1), 133–155.

SILVERMAN, B. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London.

SMITH, J., AND P. TODD (2005): "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?," *Journal of Econometrics*, 125(1-2), 305–353.

STONE, M. (1974): "Cross-Validatory Choice and Assessment of Statistical Predictions," *Journal of the Royal Statistical Society*, 36(2), 111–147.

WUNDERLICH, G. (2004): "Die Hartz-Gesetze I-IV," in *Hartz-Gesetze - Methodische Ansätze zu einer Evaluierung*, ed. by T. Hagen, and A. Spermann, pp. 15–31. Nomos-Verlag.

Appendix

Table A.1: Selected Descriptive Statistics[a]

	Males			Females		
	ExGZ	UEG	NT	ExGZ	UEG	NT
Number of observations	811	1207	1448	704	378	848
Socio-demographic characteristics						
Age category						
18-24	0.084	0.035	0.068	0.037	0.032	0.051
25-29	0.155	0.100	0.115	0.094	0.079	0.074
30-34	0.181	0.147	0.146	0.165	0.164	0.165
35-39	0.158	0.234	0.209	0.229	0.265	0.242
40-44	0.169	0.205	0.186	0.212	0.209	0.199
45-49	0.112	0.148	0.148	0.141	0.116	0.152
50-64	0.141	0.131	0.129	0.124	0.135	0.117
Family status						
Married	0.452	0.631	0.543	0.582	0.432	0.547
Children						
No children	0.731	0.613	0.675	0.479	0.701	0.538
One child	0.136	0.153	0.136	0.253	0.180	0.213
Two or more children	0.133	0.234	0.189	0.268	0.119	0.249
Health restrictions						
Yes	0.089	0.040	0.063	0.044	0.034	0.050
Nationality						
Non-German	0.338	0.286	0.277	0.295	0.241	0.276
Desired working time						
Full time	0.979	0.993	0.985	0.550	0.833	0.621
Qualification variables						
School degree						
No degree	0.036	0.015	0.021	0.007	0.005	0.011
Lower secondary schooling	0.439	0.309	0.367	0.303	0.159	0.241
Middle secondary schooling	0.237	0.239	0.231	0.335	0.278	0.320
Specialised upper sec. schooling	0.111	0.175	0.142	0.084	0.161	0.110
Upper secondary schooling	0.178	0.263	0.239	0.271	0.397	0.320
Job qualification						
Tertiary education	0.117	0.235	0.196	0.168	0.333	0.218
Technical college education	0.064	0.118	0.089	0.033	0.040	0.037
Skilled workers	0.552	0.503	0.537	0.601	0.545	0.591
Unskilled workers	0.267	0.143	0.178	0.199	0.082	0.153
Occupational group (in previous profession)						
Agriculture	0.039	0.013	0.015	0.011	0.005	0.009
Manufacturing	0.340	0.251	0.289	0.081	0.048	0.061
Technical	0.042	0.152	0.102	0.053	0.058	0.046
Services	0.502	0.548	0.517	0.700	0.828	0.723
Others	0.076	0.036	0.078	0.155	0.061	0.160
Labour market history						
Duration of last unemployment						
$<$ 3 months	0.300	0.321	0.236	0.341	0.325	0.241
3 months - $<$ 6 months	0.207	0.239	0.265	0.156	0.206	0.262
6 months - $<$ 1 year	0.284	0.314	0.326	0.344	0.352	0.316
1 year - $<$ 2 years	0.152	0.112	0.144	0.125	0.103	0.144
$\geq$ 2 years	0.058	0.014	0.029	0.034	0.013	0.038
Work experiences (Yes)	0.827	0.870	0.856	0.857	0.836	0.869
Number of placement propositions	5.610	3.792	5.326	3.683	3.394	4.302
	(9.210)	(7.185)	(7.565)	(7.015)	(6.230)	(7.150)
Current daily unemployment transfer	23.329	38.815	31.919	17.252	29.758	21.531
	(10.990)	(14.973)	(14.027)	(8.967)	(13.164)	(11.450)
Remaining time of benefit entitlement	4.716	7.314	6.318	5.018	6.828	5.571
	(5.550)	(6.238)	(6.338)	(5.879)	(6.074)	(5.993)

Continued on next page.

Table A.1 continued.

	Males			Females		
	ExGZ	UEG	NT	ExGZ	UEG	NT
Labour market history (ctd.)						
Employment status before unemployment						
Employment	0.596	0.774	0.736	0.570	0.730	0.678
Self-employment	0.056	0.025	0.039	0.050	0.034	0.029
School/Non-employed	0.110	0.069	0.074	0.077	0.069	0.055
Unemployability	0.080	0.046	0.052	0.075	0.040	0.054
Other but once employed	0.140	0.075	0.085	0.217	0.124	0.171
Others	0.019	0.011	0.014	0.011	0.003	0.012
Employment and earnings history						
Months in regular employment						
1996 (H2)	2.841	3.642	3.012	2.722	3.325	2.486
	(2.685)	(2.613)	(2.690)	(2.695)	(2.692)	(2.711)
1997	5.890	7.786	6.568	5.790	7.212	5.325
	(5.300)	(5.211)	(5.427)	(5.393)	(5.377)	(5.510)
1998	6.199	8.214	6.870	5.783	7.622	5.471
	(5.322)	(5.047)	(5.372)	(5.395)	(5.122)	(5.513)
1999	6.793	8.826	7.393	6.270	8.108	5.741
	(5.153)	(4.700)	(5.139)	(5.329)	(4.970)	(5.371)
2000	7.215	8.826	7.744	6.643	8.169	6.167
	(5.063)	(4.655)	(5.078)	(5.293)	(4.896)	(5.446)
2001	7.287	9.167	7.811	6.915	8.876	6.923
	(4.949)	(4.412)	(5.014)	(5.163)	(4.528)	(5.412)
2002	5.518	7.788	6.694	6.020	7.648	6.354
	(4.929)	(4.659)	(5.032)	(5.042)	(4.730)	(5.153)
2003 (H1)	1.063	1.451	1.294	1.101	1.431	1.269
	(1.986)	(2.157)	(2.060)	(2.032)	(2.236)	(2.070)
Months in unemployment						
1996 (H2)	0.298	0.079	0.198	0.126	0.063	0.199
	(0.960)	(1.175)	(0.612)	(0.994)	(0.765)	(0.551)
1997	0.695	0.914	0.292	0.616	0.419	0.259
	(2.430)	(2.703)	(1.608)	(2.386)	(1.956)	(1.479)
1998	0.853	1.116	0.340	0.804	0.570	0.357
	(2.546)	(2.895)	(1.592)	(2.601)	(2.153)	(1.506)
1999	1.140	1.453	0.507	1.024	0.980	0.561
	(2.816)	(3.210)	(1.850)	(2.802)	(2.764)	(1.967)
2000	1.515	1.829	0.781	1.265	1.145	0.796
	(3.193)	(3.414)	(2.240)	(3.002)	(2.787)	(2.122)
2001	1.731	2.383	1.001	1.465	1.695	0.783
	(3.288)	(3.718)	(2.472)	(3.161)	(3.254)	(2.045)
2002	3.075	4.051	2.428	3.058	3.172	2.548
	(3.980)	(4.158)	(3.401)	(4.142)	(3.878)	(3.459)
2003 (H1)	3.756	3.925	3.627	3.724	3.766	3.680
	(2.341)	(2.356)	(2.356)	(2.381)	(2.425)	(2.405)
Daily earnings from regular employment						
1996 (H2)	39.544	31.597	55.922	25.076	25.802	41.863
	(43.600)	(36.968)	(49.005)	(33.521)	(32.215)	(43.895)
1997	31.012	57.313	40.813	24.933	43.812	25.231
	(35.994)	(49.075)	(42.962)	(30.604)	(43.832)	(33.355)
1998	33.383	62.127	43.409	24.487	45.866	25.961
	(36.529)	(49.265)	(43.859)	(30.290)	(43.380)	(34.305)
1999	35.779	66.244	47.097	25.946	51.397	27.808
	(35.955)	(48.067)	(43.705)	(29.382)	(45.119)	(33.989)
2000	37.669	69.300	50.893	26.406	52.506	30.173
	(35.712)	(48.789)	(44.720)	(29.227)	(45.473)	(35.220)
2001	37.813	74.694	52.489	27.262	58.877	34.598
	(35.078)	(48.084)	(44.971)	(27.836)	(45.191)	(37.141)
2002	27.389	64.072	46.019	22.248	50.123	30.749
	(29.691)	(47.769)	(43.849)	(25.131)	(42.686)	(34.266)
2003 (H1)	9.423	24.179	17.408	7.708	19.087	12.697
	(20.121)	(40.108)	(31.525)	(16.055)	(35.700)	(24.900)
Regional labour market context						
Strategy clusters						
IIa	0.011	0.035	0.033	0.006	0.032	0.012
IIb	0.159	0.153	0.147	0.152	0.183	0.175
IIIa	0.127	0.069	0.095	0.108	0.082	0.097
IIIb	0.080	0.090	0.094	0.094	0.066	0.072
IIIc	0.222	0.226	0.223	0.200	0.164	0.213
IV	0.118	0.151	0.129	0.122	0.241	0.153
Va	0.036	0.039	0.037	0.037	0.032	0.048
Vb	0.168	0.152	0.176	0.175	0.140	0.151
Vc	0.079	0.085	0.066	0.107	0.061	0.079

[a] Standard deviations, where applicable, are in parenthesis.

Table A.2: Propensity Score Estimation Results - Coefficients[a]

	SUS vs. Non-Participation		BA vs. Non-Participation	
	Men	*Women*	*Men*	*Women*
	b/se	*b/se*	*b/se*	*b/se*
Socio-demographic characteristics				
Age category				
25-29	0.617 ∗ ∗	1.030 ∗ ∗	0.298	0.526
	(0.235)	(0.363)	(0.246)	(0.455)
30-34	0.871 ∗ ∗	0.834∗	0.274	0.419
	(0.245)	(0.352)	(0.248)	(0.439)
35-39	0.481+	0.678+	0.252	0.486
	(0.248)	(0.349)	(0.244)	(0.434)
40-44	0.669 ∗ ∗	0.861∗	0.105	0.427
	(0.254)	(0.351)	(0.253)	(0.442)
45-49	0.652∗	0.824∗	0.165	0.425
	(0.273)	(0.360)	(0.265)	(0.454)
50-64	1.223 ∗ ∗	1.369 ∗ ∗	0.356	0.803+
	(0.282)	(0.382)	(0.287)	(0.479)
Family status: Married	−0.028	−0.001	0.137	−0.113
(Ref.: Not married)	(0.139)	(0.134)	(0.120)	(0.161)
Children (Ref.: No children)				
One child	0.183	0.097	−0.133	0.272
	(0.176)	(0.175)	(0.144)	(0.224)
Two or more children	0.000	−0.027	−0.257+	0.036
	(0.177)	(0.188)	(0.135)	(0.257)
With health restrictions	−0.084	−0.116	−0.072	0.043
	(0.196)	(0.289)	(0.212)	(0.377)
Nationality: German	0.016	−0.074	0.180+	−0.127
	(0.114)	(0.137)	(0.103)	(0.175)
Desired working time: Full-time	−0.209	−0.023	−0.085	0.682 ∗ ∗
	(0.407)	(0.152)	(0.466)	(0.216)
Qualification variables				
School degree				
Lower secondary schooling	−0.064	0.903	0.370	0.165
	(0.309)	(0.597)	(0.347)	(0.847)
Middle secondary schooling	−0.024	0.864	0.443	0.536
	(0.323)	(0.599)	(0.357)	(0.847)
Specialised upper sec. schooling	−0.036	0.764	0.411	0.589
	(0.343)	(0.621)	(0.368)	(0.864)
Upper secondary schooling	−0.160	1.019+	0.380	0.323
	(0.343)	(0.608)	(0.368)	(0.856)
Occupational group (in previous profession)				
Idiwberuf1	0.517	−0.232	0.229	−0.376
	(0.322)	(0.636)	(0.378)	(0.943)
Idiwberuf4	−0.566∗	0.491	0.235	−0.090
	(0.249)	(0.350)	(0.168)	(0.458)
Idiwberuf5	−0.113	−0.101	−0.047	0.084
	(0.120)	(0.229)	(0.112)	(0.333)
Idiwberuf6	−0.476∗	−0.408	−0.715 ∗ ∗	−0.535
	(0.234)	(0.311)	(0.249)	(0.472)
Job Qualification				
Idiwquali0	−0.089	−0.075	−0.300	0.379
	(0.221)	(0.245)	(0.187)	(0.324)
Idiwquali1	−0.116	−0.151	−0.199	0.256
	(0.233)	(0.352)	(0.188)	(0.437)
Idiwquali2	−0.122	−0.098	−0.160	0.190
	(0.133)	(0.166)	(0.133)	(0.260)
Labour market history				
Duration of last unemployment				
3 months - $<$ 6 months	−0.353∗	−0.907 ∗ ∗	−0.406 ∗ ∗	−0.748 ∗ ∗
	(0.147)	(0.173)	(0.122)	(0.206)
6 months - $<$ 1 year	−0.436 ∗ ∗	−0.450 ∗ ∗	−0.459 ∗ ∗	−0.284
	(0.140)	(0.159)	(0.121)	(0.196)
$\geq$ 1 year	−0.517 ∗ ∗	−0.696 ∗ ∗	−0.629 ∗ ∗	−1.140 ∗ ∗
	(0.192)	(0.237)	(0.192)	(0.324)
With work experiences	−0.129	−0.340+	−0.169	−0.585 ∗ ∗
	(0.149)	(0.183)	(0.135)	(0.219)
Number of placement propositions	−0.004	−0.015	−0.015∗	−0.019
	(0.006)	(0.009)	(0.007)	(0.013)
Unemployment benefits	−0.046 ∗ ∗	−0.032 ∗ ∗	0.022 ∗ ∗	0.035 ∗ ∗
	(0.007)	(0.009)	(0.005)	(0.009)
Remaining benefit entitlement	−0.041 ∗ ∗	−0.065 ∗ ∗	−0.056 ∗ ∗	−0.050∗
	(0.013)	(0.016)	(0.011)	(0.020)

Continued on next page.

Table A.2 continued.

	SUS vs. Non-Participation		BA vs. Non-Participation	
	Men	*Women*	*Men*	*Women*
	b/se	*b/se*	*b/se*	*b/se*
Months in unemployment				
1999	0.000	−0.001	−0.030	−0.004
	(0.023)	(0.028)	(0.027)	(0.041)
2000	−0.040	−0.047	−0.047+	−0.013
	(0.024)	(0.029)	(0.025)	(0.041)
2001	0.017	0.065∗	−0.035	−0.071+
	(0.022)	(0.026)	(0.023)	(0.041)
2002	0.012	−0.024	−0.001	0.051
	(0.022)	(0.027)	(0.022)	(0.036)
Months in regular employment				
1999	−0.004	0.025	−0.004	0.021
	(0.022)	(0.023)	(0.020)	(0.029)
2000	0.002	0.010	−0.007	0.008
	(0.026)	(0.025)	(0.025)	(0.034)
2001	0.010	0.045+	−0.021	0.026
	(0.027)	(0.025)	(0.027)	(0.036)
2002	0.051+	0.027	−0.006	0.000
	(0.026)	(0.024)	(0.025)	(0.035)
Daily income from regular employment				
1999	0.002	0.002	0.005∗	0.008+
	(0.004)	(0.004)	(0.003)	(0.004)
2000	−0.003	−0.000	−0.005	−0.004
	(0.004)	(0.005)	(0.003)	(0.005)
2001	0.005	−0.004	0.008∗	−0.001
	(0.004)	(0.005)	(0.003)	(0.005)
2002	−0.012 ∗∗	−0.006	0.002	0.002
	(0.004)	(0.005)	(0.003)	(0.005)
Employment status before unemployment				
Self-employment	0.432+	0.607+	−0.248	−0.089
	(0.248)	(0.327)	(0.268)	(0.420)
School/Non-employed	0.417∗	0.547∗	0.449∗	0.693∗
	(0.190)	(0.258)	(0.189)	(0.333)
Unemployability	0.341	0.422+	0.327	−0.128
	(0.212)	(0.255)	(0.215)	(0.371)
Other but once employed	0.686 ∗∗	0.807 ∗∗	0.548 ∗∗	0.479+
	(0.183)	(0.217)	(0.185)	(0.286)
Others	1.033∗	0.198	0.937∗	−0.487
	(0.458)	(0.565)	(0.473)	(1.155)
Regional labour market context - Strategy clusters				
IIb	1.372 ∗∗	1.069	−0.088	−1.413 ∗∗
	(0.431)	(0.657)	(0.265)	(0.522)
IIIa	1.308 ∗∗	1.020	−0.456	−1.312∗
	(0.436)	(0.664)	(0.286)	(0.542)
IIIb	1.087∗	1.196+	0.012	−1.310∗
	(0.444)	(0.670)	(0.280)	(0.562)
IIIc	1.156 ∗∗	0.888	−0.214	−1.646 ∗∗
	(0.425)	(0.652)	(0.258)	(0.522)
IV	1.433 ∗∗	0.968	−0.103	−1.225∗
	(0.437)	(0.660)	(0.268)	(0.519)
Va	1.189∗	0.576	0.184	−1.669 ∗∗
	(0.486)	(0.700)	(0.327)	(0.615)
Vb	1.272 ∗∗	1.085+	−0.433	−1.284∗
	(0.430)	(0.657)	(0.264)	(0.521)
Vc	1.398 ∗∗	1.365∗	0.085	−1.568 ∗∗
	(0.453)	(0.671)	(0.288)	(0.568)
Constant	−0.393	−1.537	−0.500	−1.224
	(0.733)	(1.013)	(0.706)	(1.154)
Log-likelihood	−1196.329	−885.819	−1546.651	−596.322
Hit-Rate	40.133	48.135	47.317	39.896

Table A.3: Cross-Validation for the Bandwidth Selection

Start-Up Subsidy				Bridging Allowance			
Men		Women		Men		Women	
h	RMSE	h	RMSE	h	RMSE	h	RMSE
0.00558	0.45009	*0.04673*	*0.38910*	*0.09087*	*0.42773*	1.04953	0.36659
0.01558	0.45110	0.05673	0.38956	0.10087	0.42788	1.05953	0.36659
0.02558	0.45016	0.06673	0.38955	0.11087	0.42795	1.06953	0.36658
0.03558	0.44909	0.07673	0.38962	0.12087	0.42803	1.07953	0.36658
0.04558	0.44878	0.08673	0.38963	0.13087	0.42810	1.08953	0.36658
0.05558	*0.44870*	0.09673	0.38963	0.14087	0.42823	1.09953	0.36658
0.06558	0.44887	0.10673	0.38972	0.15087	0.42837	1.10953	0.36658
0.07558	0.44916	0.11673	0.38976	0.16087	0.42851	1.11953	0.36658
0.07558	0.44916	0.12673	0.38973	0.17087	0.42866	1.12953	0.36658
0.07558	0.44916	0.13673	0.38970	0.18087	0.42882	1.13953	0.36658
0.07558	0.44916	0.14673	0.38963	0.19087	0.42897	*1.14953*	*0.36657*

Note: We implement leave-one out cross-validation in a five step procedure (see, e.g., (Galdo, 2005)):

1. Define a bandwidth search grid. Here, we use $l_{bw} + 0.05 \times g$ for $g = 0, 1, 2, ..., 20$, where $l_{bw} = max[min[|P_{0i} - P_{0-i}|, |P_{0i} - P_{0+i}|]]$ is a lower bound defined by the propensity score values of comparison group members in the support region.
2. Starting with the lowest bandwidth and using only the comparison sample, estimate the counterfactual outcome of each comparison unit using kernel matching on the remaining $N_0 - 1$ observations. Find the weighted MISE for that particular bandwidth.
3. Repeat step 2 for each of the remaining bandwidth values. Find the particular bandwidth h^+ that minimizes the weighted MISE across all estimations.
4. Refine the bandwidth h^+ by defining a $+/-0.05$ neighborhood around h^+ and select a new search grid.
5. Repeat steps 2 and 3 and select the bandwidth that yields the minimum weighted $MISE$ among all estimations.

Table A.4: Sensitivity of the Results with Respect to Different Matching Algorithms - Outcome Variable: Cumulated Effects “Not unemployed”

	Start-up Subsidy						Bridging Allowance					
	Men			Women			Men			Women		
	Effect	s.e.	OS	Effect	s.e.	OS	Effect	s.e.	OS	Effect	s.e.	OS
Kernel Matching[1]												
epan nocommon	12.19	(0.418)	0	9.72	(0.493)	1	8.55	(0.362)	0	9.13	(0.491)	0
normal nocommon	12.09	(0.453)	0	9.82	(0.497)	0	8.79	(0.352)	0	9.16	(0.437)	0
epan common	12.19	(0.425)	1	9.83	(0.520)	7	8.55	(0.346)	4	9.13	(0.411)	1
normal common	12.09	(0.414)	1	9.92	(0.494)	7	8.78	(0.370)	4	9.15	(0.461)	1
epan common trim(10)	12.27	(0.436)	76	10.02	(0.472)	66	8.65	(0.356)	112	9.01	(0.464)	36
normal common trim(10)	12.14	(0.428)	76	10.17	(0.462)	66	8.91	(0.330)	112	9.02	(0.521)	36
epan common trim(5)	12.19	(0.413)	38	10.10	(0.475)	33	8.59	(0.380)	56	9.07	(0.462)	18
normal common trim(5)	12.10	(0.407)	38	10.13	(0.480)	33	8.85	(0.333)	56	9.09	(0.476)	18
Nearest-Neighbour Matching[2]												
withrep nocommon	12.08	(0.630)	0	8.79	(0.753)	0	8.38	(0.543)	0	8.13	(0.854)	0
norep nocommon	12.13	(0.452)	0	10.36	(0.443)	0	9.43	(0.355)	0	8.79	(0.607)	0
withrep common	12.10	(0.645)	1	8.93	(0.697)	7	8.32	(0.485)	4	8.15	(0.885)	1
norep common	12.11	(0.441)	1	10.33	(0.429)	7	9.40	(0.376)	4	8.79	(0.552)	1
Caliper 0.01	12.06	(0.612)	5	8.91	(0.782)	7	8.37	(0.458)	1	8.04	(0.892)	6
Caliper 0.02	12.08	(0.642)	0	8.84	(0.764)	3	8.37	(0.501)	1	8.13	(0.834)	0

Note: Standard errors (in parentheses) are based on 200 bootstrap replications. OS (off support) indicates the number of treated individuals discarded due to missing common support.

[1] Kernel matching algorithms are implemented with two kernel functions (Normal/Gaussian and Epanechnikov), with and without common support (common and nocommon) and partly with the additional imposition of a trimming level of 5% and 10%.

[2] Nearest-Neighbour algorithms are implemented with and without common support (common and nocommon), with and without replacement (withrep and norep). Additionally caliper matching with two caliper levels (0.01 and 0.02) is implemented.

Table A.5: Sensitivity of the Results with Respect to Different Matching Algorithms - Outcome Variable: Cumulated Effects “Self-employed or regular employed”

	Start-up Subsidy						Bridging Allowance					
	Men			Women			Men			Women		
	Effect	s.e.	OS	Effect	s.e.	OS	Effect	s.e.	OS	Effect	s.e.	OS
Kernel Matching[1]												
epan	14.66	(0.420)	0	16.87	(0.482)	1	10.17	(0.389)	0	14.76	(0.506)	0
normal	14.50	(0.464)	0	16.82	(0.489)	0	10.50	(0.358)	0	14.83	(0.505)	0
epan common	14.64	(0.407)	1	17.00	(0.459)	7	10.17	(0.348)	4	14.75	(0.470)	1
normal common	14.48	(0.408)	1	16.89	(0.448)	7	10.50	(0.371)	4	14.82	(0.507)	1
epan common trim(10)	14.30	(0.421)	76	16.68	(0.466)	66	10.39	(0.380)	112	14.70	(0.480)	36
normal common trim(10)	14.23	(0.418)	76	16.65	(0.522)	66	10.74	(0.360)	112	14.75	(0.522)	36
epan common trim(5)	14.36	(0.465)	38	16.93	(0.506)	33	10.29	(0.409)	56	14.73	(0.489)	18
normal common trim(5)	14.30	(0.375)	38	16.82	(0.515)	33	10.62	(0.357)	56	14.79	(0.502)	18
Nearest-Neighbour Matching[2]												
withrep nocommon	15.08	(0.627)	0	16.24	(0.767)	0	10.19	(0.568)	0	12.57	(0.989)	0
norep nocommon	14.38	(0.473)	0	15.84	(0.484)	0	11.46	(0.363)	0	12.70	(0.626)	0
withrep common	15.06	(0.576)	1	16.49	(0.659)	7	10.13	(0.499)	4	12.57	(0.949)	1
norep common	14.36	(0.449)	1	15.87	(0.532)	7	11.43	(0.408)	4	12.69	(0.624)	1
Caliper 0.01	15.00	(0.620)	5	16.38	(0.739)	7	10.17	(0.509)	1	12.58	(0.943)	6
Caliper 0.02	15.08	(0.601)	0	16.34	(0.736)	3	10.17	(0.525)	1	12.57	(0.983)	0

Note: Standard errors (in parentheses) are based on 200 bootstrap replications. OS (off support) indicates the number of treated individuals discarded due to missing common support.

[1] Kernel matching algorithms are implemented with two kernel functions (Normal/Gaussian and Epanechnikov), with and without common support (common and nocommon) and partly with the additional imposition of a trimming level of 5% and 10%.

[2] Nearest-Neighbour algorithms are implemented with and without common support (common and nocommon), with and without replacement (withrep and norep). Additionally caliper matching with two caliper levels (0.01 and 0.02) is implemented.