CEPA DP No. 14

NOVEMBER 2019

Do Start-Up Subsidies for the Unemployed Affect Participants' Well-Being? A Rigorous Look at (Un-)Intended Consequences of Labor Market Policies

Marco Caliendo
Stefan Tübbicke

Center for Economic Policy Analysis

CEPA Discussion Papers

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

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

ISSN (online) 2628-653X

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

Published online at the Institutional Repository of the University of Potsdam
https://doi.org/10.25932/publishup-43746

CEPA Discussion Paper No. 14 NOVEMBER 2019

Do Start-Up Subsidies for the Unemployed Affect Participants' Well-Being? A Rigorous Look at (Un-)Intended Consequences of Labor Market Policies*

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

Stefan Tübbicke
University of Potsdam

ABSTRACT

We estimate the long-term effects of start-up subsidies (SUS) for the unemployed on subjective outcome indicators of well-being, as measured by the participants' satisfaction in different domains. This extends previous analyses of the current German SUS program ("Gründungszuschuss") that focused on objective outcomes – such as employment and income – and allows us to make a more complete judgment about the overall effects of SUS at the individual level. This is especially important because subsidizing the transition into self-employment may have unintended adverse effects on participants' well-being due to its risky nature and lower social security protection, especially in the long run. Having access to linked administrative-survey data providing us with rich information on pre-treatment characteristics, we base our analysis on the conditional independence assumption and use propensity score matching to estimate causal effects within the potential outcomes framework. We find long-term positive effects on job satisfaction but negative effects on individuals' satisfaction with their social security situation. Further findings suggest that the negative effect on satisfaction with social security may be driven by negative effects on unemployment and retirement insurance coverage. Our heterogeneity analysis reveals substantial variation in effects across gender, age groups and skill levels. The sensitivity analyses show that these findings are highly robust.

Keywords: Start-Up Subsidies, Propensity Score Matching, Counterfactual Analysis, Well-Being
JEL Codes: C14, L26, H43, I31, J68

Corresponding author:
Marco Caliendo
University of Potsdam
Chair of Empirical Economics
August-Bebel-Str. 89
14482 Potsdam
Germany
E-Mail: caliendo@uni-potsdam.de
Tel: +49 331 977 3225

* The authors thank Lutz Bellmann for helpful comments and the Institute for Employment Research (IAB) for cooperation and institutional support within the research project no. 1755. This paper was prepared for the conference "Rigorous Impact Evaluation in Europe" (Torino, May 20-22, 2018) in honor of Prof. Alberto Martini. The authors thank Prof. Martini for his lifelong dedication to evaluation questions and especially his drive to communicate the need for evaluations (and the associated problems) to policy makers.

1 Introduction

Unemployment is known to have substantial and detrimental impacts on individuals' well-being that may last beyond the actual spell of unemployment through scarring effects (see, e.g. Clark *et al.*, 2001; McKee-Ryan *et al.*, 2005; Winkelmann and Winkelmann, 1998). Active labor market policies (ALMP) can be seen as one way to shelter individuals from these negative effects on well-being by increasing the likelihood of re-employment. To date, the ALMP literature has mostly focused on *objective* outcome indicators of success such as earnings and employment for evaluating the success of programs (see, e.g. Card *et al.*, 2017, for a meta-analysis). However, the literature also recently started to investigate the effects of ALMPs on *subjective* outcomes such as life satisfaction or self-assessed health. Examples include studies on the effects of job creation programs in Germany (Crost, 2016; Wulfgramm, 2011), a UK training program (Andersen, 2008) or job search programs in the US (Vinokur *et al.*, 2000; Vuori and Silvonen, 2005).[1] Expanding the set of outcomes to include measures of *subjective well-being* may provide useful information for analyzing effect heterogeneity and improving policy design. For example, it may be the case that groups of individuals display similar impacts in terms of objective outcomes but the effects on subjective outcomes may be very different, which allows further improving targeting of participants. In extreme cases, it may also be the case that effects on objective and subjective outcomes diverge such that inference based on objective measures only would provide a skewed picture of program effects on participants' overall welfare.

While the latter case seems relatively unlikely in the case of traditional ALMPs such as training, start-up subsidies (SUS) – which are a particular kind of ALMP that aim to help jobseekers to escape unemployment by granting them temporary transfers to take up self-employment and set up a business – are at higher risk of misleading inference based on objective outcomes alone. This is because the overall effect of unemployed individuals transitioning into self-employment on subjective well-being is ambiguous. On the one hand, there is a relatively large body of literature showing the positive effects of self-employment in the general population on job satisfaction (e.g. see Benz and Frey, 2008; Hurst and Pugsley, 2011, 2016). However, self-employment is often

[1] Korpi (1997) and Sage (2015) focus on effects of ALMPs without differentiating between different types of programs. For overview articles and meta-analyses, see Coutts *et al.* (2014), McKee-Ryan *et al.* (2005), and Puig-Barrachina *et al.* (2019).

also associated with higher earnings risk and lower social security coverage compared to regular employment (see European Commission, 2015, for European evidence), transferring more risk to the individual. This may be especially relevant among unemployed individuals who start their business from a position of relative scarcity (see Caliendo *et al.*, 2019, for a general discussion), facing multiple disadvantages such as capital constraints (Meager, 1996; Perry, 2006). Together with mixed evidence of health effects of self-employment (e.g. see Blanchflower, 2004; Nikolova, 2019), this may result in unintended negative effects of SUS on subjective well-being.

Despite these theoretical concerns, little is known about the effects of SUS on subjective well-being.[2] In this paper, we narrow this research gap by extending previous analyses on *objective outcomes* of the current German SUS program ("Gründungszuschuss", dubbed New Start-Up Subsidy, Bellmann *et al.*, 2018; Caliendo and Tübbicke, 2019) and estimate the long-run effects of participation on *subjective well-being* along several dimensions. Analyzing the German program provides a useful benchmark as many SUS programs in Europe have a relatively similar institutional setup (see O'Leary, 1999; Caliendo *et al.*, 2016, for details). We base our analysis on combined administrative and survey data, giving us access to a rich set of control variables to perform the estimation under the selection-on-observables assumption using propensity score matching techniques (Rosenbaum and Rubin, 1983). We find positive long-run effects on satisfaction with the overall job situation but negative long-run effects on satisfaction with social security. A supplementary analysis suggests that this effect may be driven by negative effects on social security coverage. Moreover, our analysis of effect heterogeneity displays substantial differences in program impacts across gender, age categories and skill levels. Sensitivity analyses suggest that these findings are highly robust.

The remainder of this paper is structured as follows. Section 2 provides details of the New Start-Up Subsidy program in Germany, offers some theoretical considerations on the effects of SUS on participants and describes the data used and displays some descriptive statistics. Section

[2] Examples of evaluation studies providing evidence on the effects of SUS on objective outcomes include Tokila (2009) for Finland, Duhautois *et al.* (2015) for France, Caliendo and Künn (2011) and Wolff *et al.* (2016) for Germany, O'Leary (1999) for Hungary and Poland, Perry (2006) for New Zealand, Rodríguez-Planas and Jacob (2010) for Romania and Behrenz *et al.* (2016) for Sweden. To our knowledge, Rose (2019) is the only study to investigate the effects of SUS on subjective well-being. However, the study is concerned with the immediate impact of SUS participation, providing no evidence on longer-run effects. Although the effects of entrepreneurship training are not the focus of this paper, we would also like to mention the paper by Fairlie *et al.* (2015) who do not find any significant long-run effects on labor market outcomes or subjective well-being in the US.

3 elaborates on the identification as well as estimation approach and details the results of our empirical analysis. Section 4 concludes.

2 The Program, Theory and Data

2.1 The New Start-Up Subsidy and Selection into the Program

The New Start-Up Subsidy has been in place in Germany since the end of 2011. In 2013, about 30,000 out of roughly 2.5 million unemployed individuals joined the program annually, according to official statistics from the Federal Employment Agency (FEA). In order to be eligible for the program, unemployed individuals have to be recipients of unemployment benefits (UB) I with at least 150 remaining days of entitlement when applying.[3] To apply for the program, individuals need to take several steps. First, they need to write up their business plan and have its sustainability approved by some external experts, e.g. from the local chamber of commerce. Second, using these documents, the individual can apply for the subsidy at the local employment agency before caseworkers make the final decision on whether the subsidy is awarded to the applicant.[4] In their decision, the caseworker is supposed to take the re-employment probability of the applicant into account, i.e. access to the program should only be granted if a re-integration of the individual into the labor market is unlikely without support.[5] Indeed, qualitative interviews by Bernhard and Grüttner (2015) suggest that caseworkers most often reject applicants if they find a sufficient number of job vacancies for them in the local labor market. In addition, the law requires that program choice is supposed to reflect the individuals' abilities and hence some entrepreneurial affinity is required. However, Bernhard and Grüttner (2015) also show that rejections due to a lack of quality in the presented business plan are relatively rare and hence the individual's re-employment probability appears to be the most important confounding factor.

Successful applicants receive a monthly transfer equivalent to their unemployment benefits

[3]As of 2008, individuals can receive UB I for a maximum of 24 months, depending on the length of their previous employment spells. Once benefit eligibility is exhausted or if individuals never qualify for UB I benefits, individuals are eligible for UB II (welfare) benefits. For UB II recipients, which make up roughly 70% of all unemployed individuals in 2012 and largely comprise long-term unemployed individuals with sparse employment histories, another SUS program is available, called "Einstiegsgeld" (see Wolff and Nivorozhkin, 2012, for an evaluation).

[4]Overall, about 30% of SUS applications were rejected by the local caseworker in 2012.

[5]This is called the "placement priority", i.e. ALMP shall only be used if they are deemed necessary for the person to find a job; otherwise, placements into the regular labor market are to be prioritized.

plus a fixed transfer of €300 for six months after entering the program.[6] A second and optional benefit phase – which only pays the fixed transfer for an additional nine months – can be granted if the business is still running and further support is needed. On average, participants received a total subsidy of around €10,350 in 2012.[7] The fixed transfer is explicitly paid to cover social insurance contributions (health care and unemployment insurance). While having health insurance is mandatory for everyone in Germany, public unemployment and to some degree also retirement insurance coverage is not obligatory for the self-employed.[8] However, individuals may sign up for both types of insurance. While access to the public retirement insurance system is unrestricted, joining the public unemployment insurance system voluntarily is only possible within three months after the start-up and only if individuals had been in regular employment for at least 12 months during the last two years. Self-employed individuals who do not opt for unemployment or retirement insurance retain their entitlement to both types of insurance based on their previous contributions. On the other hand, individuals in regular employment are automatically enrolled in all types of social insurance systems. For unemployed persons, the FEA covers the cost of health and retirement insurance while no contributions to the unemployment insurance are made.

2.2 Theoretical Considerations on Program Effects and Channels

SUS programs aim to re-integrate unemployed individuals into the labor market via the route of self-employment. Theoretical justification for this type of program is given by the existence of multiple entry barriers into self-employment for unemployed individuals. First, SUS reduce potential capital constraints that are possibly due to lower personal financial means or discrimination by the capital market (Meager, 1996; Perry, 2006). Second, unemployed individuals face a disadvantage regarding (start-up-specific) human and social capital as well as labor market experience (Pfeiffer and Reize, 2000). By providing start-ups out of unemployment with a secure minimum income for a limited duration, unemployed individuals are expected to (partially)

[6]For example, an individual with monthly unemployment benefits to the amount of €1,000 would receive a total subsidy of €7,800 over the course of the first six months.

[7]To put this into perspective, the mean investment at businesses foundation among individuals who started their business from employment is about €44,000 (Caliendo *et al.*, 2015b) and the mean monthly disposable household income between 2010 and 2013 was about €1,900 in Germany (Krause *et al.*, 2017).

[8]Certain subgroups of self-employed individuals are obliged to join the public retirement insurance scheme, e.g. freelancing teachers, although this does not concern the majority of the self-employed population though.

make up for these disadvantages (Caliendo *et al.*, 2015b). In addition, SUS may also remedy a lack of awareness regarding self-employment as a viable employment alternative among the unemployed (Storey, 2003). By ameliorating these constraints faced by the unemployed, their labor market prospects are expected to improve through SUS participation.

Theoretical considerations regarding the effects on subjective measures such as overall life satisfaction are less clear-cut. This is because – in contrast to other more traditional types of ALMPS – participation in SUS is likely to induce higher rates of self-employment relative to non-participation, even if there were no differences in overall employment rates between the two states. On the one hand, this may have positive effects on subjective well-being, as self-employment in the general population is associated with non-pecuniary benefits such as greater job-related freedom (Benz and Frey, 2008; Hurst and Pugsley, 2011, 2016). For example, this is supported by Lange (2012), who finds positive effects of self-employment on job satisfaction. On the other hand, self-employment is inherently more risky than regular employment in terms of future earnings and thus may take a toll on individuals' well-being. This may be especially relevant among unemployed individuals who start their business from a position of relative scarcity, facing multiple disadvantages described above. Moreover, there is mixed evidence on the association between health and self-employment. While Nikolova (2019) finds positive health effects of self-employment, Blanchflower (2004) reports increased rates of stress, exhaustion from work, a loss of sleep due to worry and feelings of pressure among the self-employed. Overall, this may potentially lead to unintended negative effects of SUS on subjective well-being, even if the program raises the employment prospects of participants.

2.3 Data

In order to evaluate the SUS program, we use a random sample of previously-unemployed participants who joined the program between February and June 2012, representing about 17% of all entries into the program in the respective time period. Our comparison group comprises individuals who were unemployed for at least one day, were eligible for the program (i.e. had at least 150 remaining days of UB I entitlement) but did not apply for it in this period. Both samples were drawn from the Integrated Labor Market Biographies (IEB) of the FEA. The IEB

– containing all individuals who have ever been employed subject to social security contributions or registered as unemployed – covers the employment history of individuals and provides information on socio-demographics, previous earnings, human capital, ALMP history and regional information. The extensive register data is enriched with informative survey data collected via two computer-assisted telephone interviews around 20 and 40 months after entering the program. In order to reduce survey costs, non-participants to be interviewed were selected via a pre-matching strategy to avoid interviewing individuals with very dissimilar observed characteristics compared to actual participants. Non-participants are assigned a hypothetical entry month into the program based on the month in which they were observed in unemployment and thus drawn as a comparison individual from the IEB (see Bellmann *et al.*, 2018; Caliendo and Tübbicke, 2019, and Appendix B for more details on the data). The survey includes information on the intergenerational transmission of education, labor force attachment and self-employment as well as personality traits such as the "Big 5", risk preferences or locus of control, which have proven important in the context of entrepreneurial decision-making (see Caliendo *et al.*, 2015b, 2016, for a detailed discussion).[9] Outcome information is also gathered through the survey and allows tracking individuals for 40 months in the panel sample. An analysis of panel attrition reveals non-selective attrition patterns with respect to our main outcomes of interest (see Appendix B for details). The final dataset contains 1,248 participants and 1,204 non-participants.

As in Caliendo and Tübbicke (2019), we use information on individuals' labor market status and net monthly earnings as objective outcome measures, but our main outcomes are measures of subjective well-being given by the individuals' self-reported satisfaction with life, their health, income, their job and social security situation. These items are measured on a seven-point Likert scale from 1 "completely dissatisfied" to 7 "very satisfied".[10] To shed some light on potential

[9]The five-factor model – most-often referred to as the "Big 5" – is probably the most well-known personality construct. Its five dimensions are conscientiousness, extraversion, agreeableness, neuroticism, and openness. Other measures of personality traits include locus of control, risk attitudes, patience, impulsiveness and general self-efficacy. All personality traits and preferences are measured 20 months after start-up. The Big 5, locus of control and general self-efficacy are measured on a Likert scale from 1 to 5. All other traits and preferences are measured on an eleven-point Likert scale from 0 to 10. In all cases, the lowest values refers to not having this trait/preference at all and the highest value refers to fully agreeing with having this trait/prefence (see Table 1 for more details).

[10]The survey question is as follows: "Now it's about your satisfaction. Please answer with a value from 1 "completely dissatisfied" to 7 "very satisfied". With the values in between you can grade your satisfaction. a) How satisfied are you with your life at the moment? b) How satisfied are you with your health at the moment? c) How satisfied are you currently with your income or earnings? d) Regardless of your income, how satisfied are you – generally speaking – with your employment status or your job situation? and e) How satisfied are you with your social security or social benefits?"

mechanisms, we also use the individuals' unemployment and retirement insurance contributions as well as a subjective assessment of the sufficiency of retirement plans as additional outcomes.

2.4 Descriptives

Table 1 presents descriptives on some of the pre-treatment characteristics of our sample of participants and non-participants. A full overview is provided in Table A.1 in the Appendix. Table 2 provides outcome descriptives for all of our measures of success of the SUS program.

[Insert Table 1 and Table 2 about here]

Table 1 reveals statistically significant differences in terms of several characteristics between our treatment group and the group of comparison individuals. For example, it can be seen that on average participants are slightly younger, less likely to be female and generally better educated than our non-participants. In addition, we find that our sample of treated individuals have a more favorable long- and short-term employment history. While participants had spent on average about 10% of the last 10 years in unemployment, our comparison group was unemployed for 17% of that time.[11] In addition, average daily earnings prior to unemployment were higher among participants. About 5% of participants were also already self-employed before the start of their unemployment spell, whereas this was only true for 1.2% of the pool of comparison individuals. With respect to intergenerational transmission, participants are more likely to have parents who have been self-employed, which is named as one of the key drivers for becoming self-employed in the entrepreneurship literature (Lindquist *et al.*, 2016). Moreover, participants and comparison individuals significantly differ with respect to a variety of personality traits and preferences. For instance, participants are on average more conscientious, more extraverted, less neurotic and more open to new experiences. They also possess a higher willingness to take risks and are more strongly convinced that much of their life outcomes depend on their own actions, i.e. they have a more internal locus of control (Rotter, 1966).

While the characteristics of the businesses created through the SUS scheme are not the focus of the paper, Table A.2 in the Appendix provides some auxiliary information in this regard. Put

[11] This may seem contradictive to the "placement priority" principle described in Section 2.1. However, the priority principle implies that successful applicants should display worse labor market histories compared to rejected applicants and not necessarily when compared to the stock of unemployed individuals.

simply, subsidized businesses are mainly started with prior industry-specific experience from regular employment, they invest on average around €19,000 – often financed entirely from own equity – and they are most commonly active in the service industry, followed by retail or wholesale and construction. In terms of business outcomes, subsidized businesses show high survival rates and substantial earnings. Slightly more than one-third of businesses create jobs, although innovation activity is limited.[12]

Outcomes of interest Table 2 presents descriptive statistics for our outcomes of interest. The results show that participants have more favorable objective labor market outcomes 40 months after entering the program. They are not only more likely to be self-employed but they are also much more likely to be employed in general (i.e. in either self- or regular employment). In addition, they have larger net monthly earned incomes. Regarding our main outcome measures of subjective well-being, we can state that they score significantly higher on the Likert scale for most variables, i.e. they are more satisfied with their life, health, income and job situation in general. However, participants show lower levels of average satisfaction with their social security protection. It can also be seen that participants are less likely to contribute to the public unemployment insurance system and they are also less likely to make contributions to a retirement plan. These tendencies hold for outcomes both measured after 20 and after 40 months. However, the differences in outcome means of participants and non-participants shrink over time.

[Insert Figure 1 about here]

As comparing means of ordinal variables can be misleading, we also provide descriptive statistics on the entire distribution of subjective well-being variables in Figure 1. The graphical analysis reveals that the distribution of participants' outcomes has more probability mass at the upper end of the Likert scale for life satisfaction, satisfaction with health and income as well as job satisfaction. However, the opposite is true regarding satisfaction with social security, i.e. non-participants are more likely to score high on the scale compared to participants. In our causal analysis, we will also estimate effects on the probability of scoring above the midpoint of

[12]For a more detailed comparison of subsidized businesses and regular start-ups, see Caliendo *et al.* (2015b, 2019).

the Likert scale. For example, 85% of participants are satisfied with life in general (score above the midpoint of the Likert scale) after 40 months while the same is only true for about 77% of non-participants. At the same time 55% of non-participants but only 48% of participants are satisfied with their social security situation (for more details, see Table 2).

3 Estimation Strategy and Empirical Analysis

This section first elaborates on our identification approach and describes the implementation of the propensity score matching strategy. Second, our main main empirical analysis regarding long-term effects of SUS and their heterogeneity is presented. Finally, the sensitivity of our results is assessed.

3.1 Parameter of Interest and Identification Strategy

In order to estimate the causal effects of the New Start-Up Subsidy program on labor market outcomes and subjective well-being for actual participants, we base our analysis on the potential outcomes framework usually attributed to Roy (1951) and Rubin (1974). The parameter that we want to estimate is the average treatment effect on the treated (ATT), defined as

$$\tau_{ATT} = E[Y^1 \mid D = 1] - E[Y^0 \mid D = 1], \tag{1}$$

where D is the treatment indicator, taking on the value of one if the person received SUS and zero otherwise. Y^1 corresponds to the outcome in the treated state and Y^0 is the potential outcome in the untreated state. Unfortunately, the counterfactual outcome Y^0 is not observed for participants. This *fundamental evaluation problem* implies that it is necessary to estimate the second expectation in equation (1) from data on non-participants. However, simple mean comparisons between participants and non-participants are inconsistent due to selection bias. This bias may result from differences in observed characteristics X and/or unobserved characteristics U.

Propensity score matching (PSM) methods as pioneered by Rosenbaum and Rubin (1983) eliminate bias due to observed characteristics by re-weighting comparison individuals such that characteristics are balanced across samples. For PSM methods to give a consistent estimate of the ATT, the vector of observed characteristics X needs to be sufficiently rich to satisfy the unconfoundedness assumption, also often called the conditional independence assumption (CIA)

$$Y^0 \perp\!\!\!\perp D \mid P(X), \tag{2}$$

where $P(X) = Pr(D = 1 \mid X)$ is the propensity score (see Lechner, 2001, for details). As Lechner and Wunsch (2013) note, this implies that the propensity score specification must include all characteristics that simultaneously influence the outcome of interest and the treatment probability. In our application, we should therefore include all covariates related to take-up of the SUS program, subjective well-being, labor market outcomes outcomes, and social insurance uptake. As previously noted, one of the main determinants of receiving SUS is the re-employment probability of participants, which is likely to be strongly correlated with their socio-demographics, educational attainment, parental background and their labor market history. Arguably, these factors are also highly-relevant determinants of our outcome measures. In addition, some measure of entrepreneurial affinity should be included in the estimation procedure. As entrepreneurial affinity is unobservable, we include several variables with a strong link to self-employment into the propensity score specification. These include parental self-employment, regional start-up activity and a very informative set of personality traits and preferences.[13] This is likely to substantially attenuate potential bias because personality traits and preferences have been shown to have a sizable and statistically significant impact not only on the likelihood of becoming self-employed but also on subjective well-being, insurance take-up as well as labor market outcomes.[14] Despite the rich data that we have, failure of the CIA poses a potential threat to the validity of our estimates and thus we will assess their sensitivity with respect to unobserved confounders.

[13]The personality traits and prefences in our dataset were surveyed 20 months after entry into the program. Including them as control variables assumes that they are unaffected by treatment. As Cobb-Clark and Schurer (2012) and Cobb-Clark and Schurer (2013) provide some evidence in favor of the stability of the Big 5 and locus of control over time, we also estimate effects based on a reduced spefication including only these personality traits, i.e. excluding the readiness to take risks, patience, impulsiveness and general self-efficacy from the estimation of the propensity score. We only find negligible differences to our baseline apporach.

[14]Caliendo *et al.* (2016) extensively discuss the role of personality traits for evaluating SUS, while other examples of labor market effects of personality traits are manifold: Uysal and Pohlmeier (2011) show that personality traits affect unemployment duration, Heckman *et al.* (2006) provide evidence that the effect of personality traits on wages are about as strong as for cognitive abilities, Caliendo *et al.* (2015a) show that unemployed with an internal locus of control search more intensively for jobs and set higher reservation wages, while Caliendo *et al.* (2014) show that personality traits have a significant impact on the probability of becoming and staying self-employed. Bucciol and Zarri (2017) and Cobb-Clark *et al.* (2016) find statistically significant relationships between personality traits and investment decisions. Regarding the relationship between personality traits and subjective well-being, DeNeve and Cooper (1998) provide an extensive meta-analysis and find a strong link.

In addition, the overlap assumption $P(X) < 1 \; \forall \; X$ is necessary to find suitable counterparts in the comparison group for every treated observation. If this is not the case, only the *sub-sample* average treatment effect on the treated is point-identified. However, Lechner (2008) develops worst-case bounds in the spirit of Manski (1990), which help to assess whether or not the true ATT may be substantially different from the estimates for the sub-sample. Finally, it is necessary to rule out spill-over effects (Stable Unit Treatment Value Assumption, or SUTVA for short). Since the current German SUS program is relatively small – with only around 30,000 annual entries compared to roughly 350,000 entries into predecessor programs in 2004 – this seems to be a reasonable assumption.

3.2 Implementing the Matching Procedure

The implementation of PSM requires estimating the propensity score and the imposition of common support before matching (see Caliendo and Kopeinig, 2008, for an overview). Subsequently, based on the chosen matching algorithm, the matching has to be performed and finally it is necessary to assess the resulting balancing quality. If the matching quality is not satisfactory, the propensity score specification needs to be re-examined, common support adjusted or the matching algorithm changed until the matched sample can be regarded as balanced.

Estimation of the Propensity Score and the Imposition of Common Support As the first step of the matching procedure, we estimate the propensity score using a logit regression including information on the individuals' labor market history, their socio-demographics, human capital acquisition, intergenerational transmission of education, labor force attachment and self-employment, as well as usually-unobserved personality traits and regional controls for macroeconomic conditions and self-employment activity. Overall, we use 91 control variables, including some interaction terms that have been added in an iterative manner to improve resulting balancing quality.[15] However, the quality of the control variables is more important than the sheer quantity of the variables included in the propensity score estimation. We have already

[15]There are minor differences in our empirical strategy and hence estimation results regarding labor market outcomes when compared to Caliendo and Tübbicke (2019). We include personality traits in our main specification as we believe that they are more important in the context of assessing the effects on subjective well-being, while in Caliendo and Tübbicke (2019) they were only included in the sensitivtiy anaylsis (resulting only in negligible differences). Furthermore, the specification regarding regional economic indicators is slightly more parsimonious and a logit regression is used as it yields slightly better balancing quality in the next step.

argued above which variables are likely to influence selection into the subsidy and outcome variables simultaneously and argued that especially the availability of personality traits is important to our application. The rich nature of our data is paramount for the CIA to provide a reasonable assumption. Details on the exact specification employed can be found in Table A.3 in the Appendix. Generally, we aim to make the specification relatively flexible by using categorical dummies on continuous variables to better balance higher moments of the confounders. Subsequently, the estimated coefficients are used to obtain predicted values of the propensity score. The distribution of these predicted values can be found in Figure 2.

[Insert Figure 2 about here]

As expected due to the covariate imbalance described in Section 3.1, scores are skewed towards one for the treated and towards zero for comparison individuals. As stressed by Heckman *et al.* (1998), comparing individuals off common covariate support is a major source of selection bias. To avoid this, we impose common support by dropping treated individuals from the analysis if they lie outside the range of propensity scores among non-participants as described by Dehejia and Wahba (2002). This procedure was found to most strongly improve the mean squared error of matching estimators in a recent simulation study by Lechner and Strittmatter (2019).

Matching on the Propensity Score Treated individuals are then matched with comparison individuals using Epanechnikov kernel matching to create a sample balanced in observed characteristics.[16] Since PSM does not match on characteristics directly, it is necessary to judge the appropriateness of the propensity score specification against the resulting matching quality (Smith and Todd, 2005). Once a sufficiently balanced synthetic sample is created, mean differences in outcomes between participants and *re-weighed* comparison individuals serve as the estimate of the ATT

$$\hat{\tau}_{ATT} = \frac{1}{N_1}\sum_{i=1}^{N_1} Y_i - \frac{1}{N_0}\sum_{j=1}^{N_0} \hat{w}_j Y_j, \tag{3}$$

[16] To choose the kernel bandwidth, we performed a grid search over possible values and chose the bandwidth that maximizes post-matching balance in terms of the pseudo R^2 in the re-estimation of the propensity score after matching. This is the case for $h = 0.15$.

where $\hat{w}_j$ are estimated balancing weights and N_1 and N_0 are the number of treated and untreated observations in the sample, respectively. Statistical inference is then performed using re-sampling methods. Specifically, we obtain p-values through bootstrapping the t-statistic with 999 replications based on asymptotic variance approximations (MacKinnon, 2006; Bodory *et al.*, 2016).[17]

Matching Quality Table 3 compares several commonly-used balance indicators before and after matching. In general, one can see that matching quality dramatically increases. The number of statistically different means at the 10% level decreases from 61 to just three variables.[18] Similarly, mean standardized bias decreases from 15.3% to 2.4%, which is below the 3-5% threshold suggested by Caliendo and Kopeinig (2008). Re-estimating the propensity score on the matched sample gives a pseudo-R^2 of 1.8% and a corresponding p-value of essentially one. Following this approach, covariates overall no longer possess any predictive power with respect to treatment after performing the matching procedure. Balancing indicators based on Rubin (2001) measure standardized mean difference in the linear index of the propensity score (the so-called Rubin's B) and the variance ratio of the propensity score (Rubin's R). Ideally, these should be close to zero or one, respectively. We can see that Rubin's B substantially decreases through the matching procedure, while Rubin's R remains relatively close to one. Overall, the balancing quality is strongly improved by the kernel matching approach, which allows us to proceed with our causal analysis.[19]

[Insert Table 3 about here]

3.3 Main Results and Discussion

Table 4 gives our main estimates of the average treatment effects on the treated of the New Start-Up Subsidy program for unemployed individuals.

[Insert Table 4 about here]

[17] The basic idea is to use an asymptotic approximation to the standard error to compute the t-statistic for every bootstrap replication. In a final step, the distribution of t-statistics is then compared to the t-statistic in the original sample to compute a p-value. This is likely to improve upon the standard bootstrap, i.e. bootstrapping the ATT directly (MacKinnon, 2006; Bodory *et al.*, 2016).

[18] For covariate means after matching, see Table A.1 in the Appendix.

[19] As part of our sensitivity analysis later on, we will show that bias due to residual imbalance does not play any major role in our case (see Section 3.5).

The PSM estimates suggest significant effects on individuals' probability of being in either self- or regular employment as well as net monthly earned income. The program led to a 20.8 percentage point increase in the likelihood of being self- or regular employed and an increase of 910€ a month in terms of net earned income 40 months after entering the SUS program.[20]

In terms of subjective well-being, our baseline estimates suggest that the program has a short-run impact on general life satisfaction which fades over time, i.e. 40 months after start of the program, the difference in life satisfaction between treated and matched controls is close to zero and not statistically significant. In order to gain a deeper understanding of the effects on overall well-being, we turn to analyzing effects with respect to certain sub-domains of subjective well-being. Estimates are given in terms of both Likert points as well as the percent of a standard deviation. First, we can state that there are positive short- and long-term effects of SUS participation on individuals' satisfaction with their job situation. The treatment increases job satisfaction by about 0.28 points or 17.3% after 40 months. Point estimates for the effect on satisfaction with income are positive but insignificant at any common significance level. The point estimates on satisfaction with one's health are negative and growing over time in magnitude but are also statistically insignificant. Finally, we can state that participation in the SUS program has a large negative effect on individuals' satisfaction with their social security situation of about 0.51 (0.57) Likert points in the short (long) run, equal to a decrease of (32%) 35.6% of a standard deviation.

[Insert Figure 3 about here]

As previously noted, mean comparisons of ordinal variables may be misleading. Hence, we also estimate effects on the probability of scoring above the midpoint of the Likert scale (i.e. five or above) for the subjective well-being outcomes after 40 months. The results of this analysis can be found in Figure 3. The results support the zero average effects obtained by mean comparison for life satisfaction and satisfaction with health and income. Regarding job satisfaction, SUS increase the probability of scoring high on the Likert scale by about 6.5 percentage points. Moreover, SUS participation clearly reduces satisfaction with social security as it reduces the

[20] The respective estimated counterfactual means are equal to 73% and € 1,304. These results extend the analysis of Bellmann *et al.* (2018) and confirm the findings of Caliendo and Tübbicke (2019).

likelihood of scoring above the midpoint of the scale by about seventeen percentage points. Thus, the evidence seems to suggest that participants – who mostly remain self-employed even after 40 months – show increased longer-run worries about their social security situation. As this additional analysis supports the mean comparisons, we refrain from presenting results in such detail for the remainder of the paper, rather focusing only on mean effects.

To provide some evidence on which channels drive this last result, we also estimate the effect of the SUS program on some auxiliary objective (social) insurance outcomes. More specifically, we estimate effects on the likelihood of contributing to the unemployment insurance system and retirement insurance plans. Our findings in Table 4 suggest that SUS participation substantially reduces the likelihood of contributing to the public unemployment insurance scheme by over twenty percentage points.[21]This may reflect that participants willingly choose investments in their business over joining the public unemployment insurance scheme. Alternatively, the effect may be driven by a lack of information regarding the three-month time restriction to enter the insurance scheme for the self-employed. In any case, this implies that participants are significantly more vulnerable to economic downturns in the future. Regarding retirement insurance, the program is estimated to reduce the probability of making contributions to some sort of retirement plan by about 4.8 percentage points after 40 months. In addition, participants are almost six percentage points more likely to view their retirement investment as insufficient at the same point in time. Taken together, this seems to suggest that the program leads to reduced investments in retirement plans among participants, which could potentially increase the risk of old-age poverty. Unfortunately, our data does not allow us to make more concrete statements here as we do not know contribution amounts (or accumulated retirement benefits for that matter).[22]

Overall, our new results imply a slightly more mixed assessment of program effects compared to previous studies. On the one hand, SUS improve the employment prospects, income as well as job satisfaction of participants. On the other hand, our results show negative effects on par-

[21]While we do not present results conditional on employment, it is noteworthy that these effects are even larger.

[22]In general, self-employed (compared to the regular employed) are more likely to invest in private retirement insurance plans characterized by different contribution rates depending on the chosen plan. We do not have information on these amounts and also do not know how much individual retirement insurance beneift entitlements have been accumulated at this stage. Hence, we need to interpret the findings with caution.

ticipants' satisfaction with their social security situation as well as their unemployment benefits (and retirement) insurance contributions. Hence, our analysis provides some evidence on unintended negative effects of SUS participation, suggesting that the program may potentially be improved by altering its institutional design. First, participants may need to be provided with more information on the legal constraints regarding their uptake of unemployment insurance to avoid being locked out of the system. Second, this may be combined with incentivizing individuals to increase their investments into unemployment and retirement insurance, e.g. through targeted (or higher) support in the second benefit period. Taken together, these measures could reduce participants' concerns regarding their social security situation.

3.4 Effect Heterogeneity

Knowledge about effect heterogeneity plays a crucial role in improving policy design and targeting individuals to be selected into treatment. Previous analyses of SUS programs have shown that effects on objective outcomes indeed appear to be highly heterogeneous (see Caliendo and Künn, 2011; Bellmann *et al.*, 2018, for example). Hence, effect heterogeneity is also likely in terms of subjective outcomes. In order to analyze this, we split our sample by gender, age (above or below 45 years) as well as skills (being low- or high-skilled, where the latter are individuals with a craftsmanship or a university degree).

The previously-described estimation steps are then repeated on these sub-samples. Table 5 shows the results from this analysis after 40 months, including the number of treated and untreated individuals in each sub-sample as well as aggregate measures of covariate balance. While balancing quality tends to be worse than in the full sample, the MSB remains below the 5% threshold suggested by Caliendo and Kopeinig (2008) and the covariates remain statistically insignificant in the re-estimation of the propensity score in the matched sample.

[Insert Table 5 about here]

Starting with effects by gender, we find minimal differences in terms of effects on employment outcomes. However, men do tend to gain much more from participating in the SUS scheme than women in terms of earnings. In line with this finding, men display positive and statistically significant effects on satisfaction with their income, while the effects for women are statistically insignificant and negative. Similarly, men display stronger effects on job satisfaction than women. This may be due to the fact that men are more strongly affected by unemployment and therefore may also gain more from re-employment compared to women (Meer, 2014). Regarding the effects on satisfaction with social security, women show more detrimental effects. Estimating effects by age groups shows that older individuals profit much more from SUS participation than younger individuals in terms of objective outcomes. In line with this, older individuals also display stronger effects on job satisfaction than younger participants. Another interesting finding in this comparison is that younger participants tend to reduce their investment in unemployment insurance substantially more than older individuals. However, older participants show more adverse effects on their assessment of the sufficiency of their retirement plans. Splitting the sample by skills, we find larger employment effects for high-skilled individuals. Moreover, our results suggest marginally significant negative effects on participants' health satisfaction among low-skilled individuals. Finally, high-skilled participants display larger negative effects on retirement insurance contributions.

3.5 Sensitivity Analysis

For a rigorous impact evaluation, it is necessary to critically assess the applicability of identifying assumptions as the causal interpretation of our estimates crucially depends on them. In addition, the results of PSM or weighting should be analyzed for their robustness, given that these estimators require several steps of implementation with a moderate to relatively large number of discretionary choices, depending on the algorithm applied. Thus, in this Section we analyze the sensitivity of the results with respect to the matching or weighting approach chosen as well as deviations from the underlying identifying assumptions necessary for the matching approach to deliver consistent estimates.

Choice of Estimator As a first assessment of the robustness of our estimates, we test whether the kernel matching approach that we used in the previous section plays a crucial role for our results. Table 6 gives estimates for alternative estimation methods for selected outcomes.[23] First, we perform an alternative bandwidth selection for the kernel matching based on leave-one-

[23]Results for all outcomes can be found in Table A.4 in the Appendix.

out cross-validation (Frölich, 2005; Galdo *et al.*, 2008). The resulting estimates are practically identical to our baseline results. Moreover, we augment our main approach by a post-matching regression to control for residual imbalance in the matched sample and find that point estimates and p-values are largely unchanged compared to our baseline approach. Second, we employ inverse probability weighting with weights scaled to unity as well as radius matching with bias adjustment as suggested by Lechner *et al.* (2011). We chose these alternative estimators as Huber *et al.* (2013) and Busso *et al.* (2014) show that they tend to perform well in finite samples. Both of these approaches tend to deliver estimates that are even somewhat larger in magnitude than our baseline estimates. Given that some point estimates turn insignificant when using IPW despite larger absolute values of coefficients, this seems to point towards lower precision of estimates potentially due to the sensitivity of the IPW estimator with respect to the estimation of the propensity score (see Waernbaum, 2012, for details). Overall, our results are relatively stable with respect to tuning parameters as well as the choice of matching or weighting estimator.

[Insert Table 6 about here]

Unconfoundedness In order to assess the sensitivity of our estimates with respect to the unconfoundedness assumption, we follow Rosenbaum (2002) and use a bounding approach (see DiPrete and Gangl, 2004; Ichino *et al.*, 2008, for other similar applications). Assume that the treatment probability is given by $Pr(D = 1 \mid X) = \Lambda(x_i'\beta + \gamma u_i)$, where $\Lambda(\cdot)$ is the logistic cumulative distribution function. Further assuming that the unobserved confounder u_i is a dummy variable, the relative odds of two observationally-identical individuals i and j receiving treatment can be written as

$$\frac{1}{e^{\gamma}} \leq \frac{Pr(D=1 \mid X_i)/Pr(D=0 \mid X_i)}{Pr(D=1 \mid X_j)/Pr(D=0 \mid X_j)} \leq e^{\gamma}, \tag{4}$$

where $\Gamma := e^{\gamma}$ is a measure of departure from the unconfoundedness assumption. The Rosenbaum bounds approach is a worst-case scenario as it assumes that the unobserved covariate u_i is close to being a perfect predictor of the outcome. Based on this assumption, the bounding approach allows us to establish whether our estimated treatment effects would still be significantly different from zero for a hypothetical value of Γ. While this strategy does not tell us whether our results are actually inconsistent due to the failure of the CIA, it provides us with valuable information on how sensitive the results are to unobserved confounders that are very predictive of the outcome. In this bounding analysis, we will assume that we over-estimated the true effect for significantly positive effects and that we under-estimated the true effect for negative effects. Hence, we assume that the true effects are closer to zero. Table 7 gives the critical values of Γ, i.e. the value for which our estimated effects would turn insignificant. In general, we can state that the long-term labor market effects are very robust to "hidden bias" as an unobserved confounder must increase the odds of receiving treatment by a factor of at least 2.58 to render our inference with respect to positive effects on earnings insignificant. The critical Γ for the effect on the probability of being in self- or regular employment is even larger. Turning to the robustness of our results on subjective well-being, we can state that these are more sensitive to deviations from the unconfoundedness assumption. Effects on job satisfaction and satisfaction with social security after 40 months turn insignificant at a value of Γ around 1.2 and 1.8, respectively. With respect to social security outcomes, we can see that the negative effects of SUS participation on the probability of not having unemployment or retirement insurance are similarly robust as our inference regarding positive labor market effects. The estimated effect regarding individuals deeming their retirement insurance insufficient becomes insignificant at a value of $\Gamma = 1.25$. Thus, we can state that for the least robust conclusions to be overturned, an unobserved confounder would need to be 1.22 times as predictive of treatment as all covariates in our propensity score specification combined. Effects regarding labor market outcomes, satisfaction with social security, uptake of unemployment insurance and retirement insurance would require an even stronger role of unobserved confounders. This seems to be relatively unlikely given the detailed information that we have at our disposal through the administrative and survey data. Hence, the main conclusions drawn in our analysis are highly robust to deviations from the unconfoundedness assumption.

[Insert Table 7 about here]

Common Support In Section 3.2, we described how the analysis is restricted to treated individuals who are in the region of common support. We implemented this by discarding treated observations that have estimated propensity score values outside the range of scores of comparison individuals, since no comparable untreated individual can be found for those treated individuals. Table 3 shows that this restriction leads to the exclusion of 73 participants from our main estimation procedure, which corresponds to roughly 5.8% of treated individuals in our sample. If effects are heterogenous as indicated by our analysis, the true average treatment effect on the treated may well be different from our estimates even if the CIA holds. In order to assess the sensitivity of the estimates with respect to this issue, Lechner (2008) developed worst-case bounds as given by

$$\tau_{ATT}^{Low} = E[Y^1 - Y^0 \mid D=1, S=1] \cdot \pi_1 + \{E[Y^1 \mid D=1, S=0] \; - \overline{Y}^0\} \cdot (1-\pi_1), \tag{5}$$

$$\tau_{ATT}^{High} = E[Y^1 - Y^0 \mid D=1, S=1] \cdot \pi_1 + \{E[Y^1 \mid D=1, S=0] \; - \underline{Y}^0\} \cdot (1-\pi_1), \tag{6}$$

where S is an indicator for being on common support and π_1 is the probability of a treated person being on support. The values of $\overline{Y}^0$ and $\underline{Y}^0$ are given by the minimum and maximum of the support of Y^0 and hence the difference in curly braces gives the minimum and maximum possible average effect for the subgroup of treated units off support. Replacing the terms in (5) and (6) with their empirical counterparts from our sample, we obtain the bounds and p-values given by the last four columns of Table 7. It is important to keep in mind that these are worst-case bounds, i.e. they give the range of estimates that are consistent with the data, but they do not make any statement on the likelihood of those bounds being reached in reality. Hence, the results have to be interpreted with the necessary caution if they suggest that inference may be overturned.

Our findings show that the estimates of labor market effects are completely robust to the support problem and the lower bounds still indicate large and highly significant effects on employment and income. Similarly, the estimates on the probability of not contributing to the unemployment insurance still yield a significant lower bound. The results on the other measures are slightly more nuanced. Upper bounds on the effects regarding satisfaction with life in general as well as social security are marginally insignificant at the 10% level. The bounding analysis also reveals that negative effects on satisfaction with respect to health and positive effects on satisfaction with income cannot be ruled out under extreme assumptions about the missing counterfactual estimates. The data would also be consistent with null effects on the uptake of retirement insurance contributions and its sufficiency under these assumptions. Thus, the main conclusions from Section 3.3 are predominantly supported by this bounding analysis.

4 Conclusion

Using non-experimental counterfactual evaluation techniques, this paper estimates the long-term effects of participation in the German New Start-Up Subsidy program on individuals' *subjective* well-being. Combining this with results on *objective* labor market outcomes, this allows us to analyze the effects of the program using a more thorough welfare definition than previously-existing evaluations of start-up subsidy programs. Using a broader welfare measure is especially important in this context as subsidizing unemployed individuals into self-employment exposes them to more risk compared with regular employment. This problem is exacerbated by the fact that social security protection is lower for self-employed individuals and hence the program may have unintended negative effects on participants' well-being in the long run. Our results based on PSM suggest that the program has relatively large positive and statistically significant effects on participants' employment prospects, income and satisfaction regarding the individuals' job situation. On the other hand, we find sizable, robust and statistically significant negative effects on individuals' satisfaction with their social insurance situation. Supplementary analyses suggests that these effect may be driven by reduced investment in unemployment insurance – making them more vulnerable to economic downturns in the future – and retirement insurance – potentially increasing the risk of old-age poverty (even though our data does not allow us to make concrete statements here). Thus, the program's overall assessment is slightly less optimistic when taking these unintended negative effects into account. In our heterogeneity analysis, we find substantial variation in effects across gender, age groups and skill levels. Our sensitivity analyses suggest that our main results are highly robust to deviations from the identifying assumptions and alternative choices regarding the implementation of the estimation strategy.

These findings underscore the relevance of using more *subjective* indicators of success to complement the analysis with respect to *objective* economic outcomes to improve policy design. Regarding policy conclusions, one lesson from our results is that it may be advisable to make more information on legal constraints regarding unemployment insurance for the self-employed available to SUS participants to avoid them being locked out of the system. Second, this may be combined with incentivizing individuals to increase their investments into unemployment and retirement insurance, e.g. through targeted (or higher) support in the second benefit period.

Looking forward, it would be desirable to validate the non-experimental results obtained so far with experimental evidence. This would provide researchers with the opportunity to vary the design of the program and thus ascertain whether the proposed solutions to participants' dissatisfaction with their social security situation can be ameliorated in this way. Furthermore, start-up subsidy programs are in need of macroeconometric evaluations since they may provide further macroeconomic benefits (or losses) that cannot be assessed using the microeconometric approach chosen here.

References

Andersen, S. H. (2008). The short- and long-term effects of government training on subjective well-being. *European Sociological Review*, **24** (4), 451–462.

Behrenz, L., Delander, L. and Månsson, J. (2016). Is starting a business a sustainable way out of unemployment? Treatment effects of the Swedish start-up subsidy. *Journal of Labor Research*, **37** (4), 389–411.

Bellmann, L., Caliendo, M. and Tübbicke, S. (2018). The post-reform effectiveness of the new german start-up subsidy for the unemployed. *LABOUR*, **32** (3), 293–319.

Benz, M. and Frey, B. S. (2008). Being independent is a great thing: Subjective evaluations of self-employment and hierarchy. *Economica*, **75** (298), 362–383.

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

Blanchflower, D. G. (2004). *Self-Employment: More may not be better*. Working Paper 10286, National Bureau of Economic Research.

Bodory, H., Camponovo, L., Huber, M. and Lechner, M. (2016). *The finite sample performance of inference methods for propensity score matching and weighting estimators*. Discussion Paper 9706, Institute for the Study of Labor (IZA), Bonn. Forthcoming in *Journal of Business & Economic Statistics*.

Bucciol, A. and Zarri, L. (2017). Do personality traits influence investors' portfolios? *Journal of Behavioral and Experimental Economics*, **68**, 1–12.

Busso, M., DiNardo, J. and McCrary, J. (2014). New evidence on the finite sample properties of propensity score reweighting and matching estimators. *The Review of Economics and Statistics*, **96** (5), 885–897.

Caliendo, M., Cobb-Clark, D. and Uhlendorff, A. (2015a). Locus of control and job search strategies. *The Review of Economics and Statistics*, **97** (1), 88–103.

—, Fossen, F. and Kritikos, A. (2014). Personality characteristics and the decisions to become and stay self-employed. *Small Business Economics*, **42** (4), 787–814.

—, Hogenacker, J., Künn, S. and Wiessner, F. (2015b). Subsidized start-ups out of unemployment: a comparison to regular business start-ups. *Small Business Economics*, **45** (1), 165–190.

— and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, **22** (1), 31–72.

— and Künn, S. (2011). Start-up subsidies for the unemployed: Long-term evidence and effect heterogeneity. *Journal of Public Economics*, **95** (3-4), 311–331.

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

—, — and — (2019). *Catching up or lagging behind? The long-term business and innovation potential of subsidized start-ups out of unemployment*. Discussion Paper 12690, Institute for the Study of Labor (IZA), Bonn.

— and Tübbicke, S. (2019). *New evidence on long-term effects of start-up subsidies: matching estimates and their robustness*. Discussion Paper 12261, Institute for the Study of Labor (IZA), Bonn. Forthcoming in *Empirical Economics*.

Card, D., Kluve, J. and Weber, A. (2017). What works? A meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, **16** (3), 894–931.

Clark, A., Georgellis, Y. and Sanfey, P. (2001). Scarring: The psychological impact of past unemployment. *Economica*, **68**, 221–41.

Cobb-Clark, D. A., Kassenboehmer, S. C. and Sinning, M. G. (2016). Locus of control and savings. *Journal of Banking & Finance*, **73**, 113 – 130.

— and Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, **115** (1), 11 – 15.

— and — (2013). Two economists' musings on the stability of locus of control. *The Economic Journal*, **123** (570), 358–400.

Coutts, A. P., Stuckler, D. and Cann, D. J. (2014). The health and well-being effects of active labor market programs. In *Wellbeing*, American Cancer Society, pp. 1–18.

Crost, B. (2016). Can workfare programs offset the negative effect of unemployment on subjective well-being? *Economics Letters*, **140**, 42 – 47.

Dehejia, R. H. and Wahba, S. (2002). Propensity score matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, **84** (1), 151–161.

DeNeve, K. and Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin*, **124** (2), 197–229.

DiPrete, T. and Gangl, M. (2004). Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological Methodology*, **34**, 271–310.

Duhautois, R., Redor, D. and Desiage, L. (2015). Long term effect of public subsidies on start-up survival and economic performance: An empirical study with French data. *Revue d'Économie Industrielle*, **149** (1), 11–41.

European Commission (2015). *Recent changes in self-employment and entrepreneurship across the EU*. Research note 6, Directorate-General for Employment, Social Affairs and Inclusion.

Fairlie, R. W., Karlan, D. and Zinman, J. (2015). Behind the gate experiment: Evidence on effects of and rationales for subsidized entrepreneurship training. *American Economic Journal: Economic Policy*, **7** (2), 125–61.

Frölich, M. (2005). Matching estimators and optimal bandwidth choice. *Statistics and Computing*, **15** (3), 197–215.

Galdo, J., Smith, J. A. and Black, D. (2008). Bandwidth selection and the estimation of treatment effects with unbalanced data. *Annales d'Économie et de Statistique*, **91/92**, 189–216.

Heckman, J., Ichimura, H., Smith, J. and Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica*, **66** (5), 1017–1098.

Heckman, J. J., Stixrud, J. and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, **24**, 411–482.

Huber, M., Lechner, M. and Steinmayr, A. (2014). Radius matching on the propensity score with bias adjustment: Tuning parameters and finite sample behavior. *Empirical Economics*, **49** (1), 1–13.

—, — and Wunsch, C. (2013). The performance of estimators based on the propensity score. *Journal of Econometrics*, **175** (1), 1–21.

Hurst, E. and Pugsley, B. (2011). What do small businesses do? *Brookings Papers on Economic Activity*, **42** (2 (Fall)), 73–142.

Hurst, E. G. and Pugsley, B. W. (2016). Wealth, tastes, and entrepreneurial choice. In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, University of Chicago Press, pp. 111–151.

Ichino, A., Mealli, F. and Nannicini, T. (2008). From temporary help jobs to permanent employment: what can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics*, **23** (3), 305–327.

Korpi, T. (1997). Is utility related to employment status? Employment, unemployment, labor market policies and subjective well-being among Swedish youth. *Labour Economics*, **4**, 125–147.

Krause, P., Franz, C. and Fratzscher, M. (2017). Income Groups and Types of Employment in Germany since 1995. *DIW Economic Bulletin*, **7** (27), 267–278.

Lange, T. (2012). Job satisfaction and self-employment: Autonomy or personality? *Small Business Economics*, **38** (2), 165–177.

Lechner, M. (2001). Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In *Econometric Evaluation of Labour Market Policies*, Heidelberg: Physica-Verlag HD, pp. 43–58.

— (2008). A note on the common support problem in applied evaluation studies. *Annales d'Economie et de Statistique*, **91/92**, 217–235.

—, Miquel, R. and Wunsch, C. (2011). Long-run effects of public sector sponsored training in west germany. *Journal of the European Economic Association*, **9** (4), 742–784.

— and Strittmatter, A. (2019). Practical procedures to deal with common support problems in matching estimation. *Econometric Reviews*, **38** (2), 193–207.

— and Wunsch, C. (2013). Sensitivity of matching-based program evaluations to the availability of control variables. *Labour Economics*, **21**, 111–121.

Lindquist, M., Sol, J., van Praag, M. and Vladasel, T. (2016). *On the Origins of Entrepreneurship: Evidence from Sibling Correlations*. Discussion Paper 11562, C.E.P.R., London.

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

Manski, C. (1990). Nonparametric bounds on treatment effects. *American Economic Review, Papers and Proceedings*, **80**, 319–323.

McKee-Ryan, F., Song, Z., Wanberg, C. and Kinicki, A. (2005). Psychological and physical well-being during unemployment: A meta-analytic study. *The Journal of applied psychology*, **90**, 53–76.

Meager, N. (1996). From unemployement to self-employement: Labour market policies for business start-up. In G. Schmidt, J. O'Reilly and K. Schömann (eds.), *International Handbook of Labour Market Policy and Evaluation*, Edward Elgar, pp. 489–519.

Meer, P. (2014). Gender, unemployment and subjective well-being: Why being unemployed is worse for men than for women. *Social Indicators Research*, **115**.

Nikolova, M. (2019). Switching to self-employment can be good for your health. *Journal of Business Venturing*, **34** (4), 664 – 691.

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

Perry, G. (2006). *Are business start-up subsidies effective for the unemployed: evaluation of enterprise allowance*. Working paper, Auckland University of Technology.

Pfeiffer, F. and Reize, F. (2000). Business start-ups by the unemployed - an econometric analysis based on firm data. *Labour Economics*, **7** (5), 629–663.

Puig-Barrachina, V., Giró, P., Artazcoz, L., Bartoll, X., Cortés-Franch, I., Fernández, A., González-Marín, P. and Borrell, C. (2019). The impact of Active Labour Market Policies on health outcomes: A Scoping review. *European Journal of Public Health*.

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

Rose, D. (2019). The impact of active labour market policies on the well-being of the unemployed. *Journal of European Social Policy*, **29** (3), 396–410.

Rosenbaum, P. and Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, **70** (1), 41–55.

Rosenbaum, P. R. (2002). *Observational Studies*. New York: Springer, 2nd edn.

Rotter, J. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, **80** (1), 1–28.

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

Rubin, D. (1974). Estimating causal effects of treatments in randomised and nonrandomised studies. *Journal of Educational Psychology*, **66** (5), 688–701.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, **2** (3), 169–188.

Sage, D. (2015). Do active labour market policies promote the well-being, health and social capital of the unemployed? Evidence from the uk. *Social Indicators Research*, **124** (2), 319–337.

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.

Smith, J. and Todd, P. (2005). Does matching overcome lalonde's critique of nonexperimental estimators? *Journal of Econometrics*, **125** (1-2), 305–353.

Storey, D. (2003). Entrepreneurship, small and medium sized enterprises and public policies. In Z. Acs and D. Audretsch (eds.), *Handbook of Entrepreneurship Research*, *18*, Kluwer Academic Publishers, pp. 473 – 511.

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

Uysal, S. D. and Pohlmeier, W. (2011). Unemployment duration and personality. *Journal of Economic Psychology*, **32** (6), 980–992.

Vinokur, A., Schul, Y., Vuori, J. and Price, R. (2000). Two years after a job loss: Long-term impact of the jobs program on reemployment and mental health. *Journal of occupational health psychology*, **5**, 32–47.

Vuori, J. and Silvonen, J. (2005). The benefits of a preventive job search program on re-employment and mental health at 2-year follow-up. *Journal of Occupational and Organizational Psychology*, **78** (1), 43–52.

Waernbaum, I. (2012). Model misspecification and robustness in causal inference: comparing matching with doubly robust estimation. *Statistics in Medicine*, **31** (15), 1572–1581.

Winkelmann, R. and Winkelmann, L. (1998). Why are the unemployed so unhappy? Evidence from panel data. *Economica*, **65**, 1–15.

Wolff, J. and Nivorozhkin, A. (2012). Start me up: The effectiveness of a self-employment programme for needy unemployed people in Germany. *Journal of Small Business & Entrepreneurship*, **25** (4), 499–518.

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

Wulfgramm, M. (2011). Can activating labour market policy offset the detrimental life satisfaction effect of unemployment? *Socio-Economic Review*, **9**, 477–501.

Tables and Figures

Table 1: Selected Descriptives

	Part.	Non-Part.	p-value
Number of observations	1,248	1,204	
Socio-demographics			
Age in years	42.77	43.94	0.002
Share female	0.425	0.509	0.000
Share with university degree	0.411	0.267	0.000
Labor market history			
Fraction of last 10 years in unemployment	0.101	0.170	0.000
Daily last income (Euro)	81.39	61.60	0.000
Share in dependent employment before unemployment	0.674	0.516	0.000
Share in self-employment before unemployment	0.054	0.012	0.000
Intergenerational transmission			
Share with father/mother having been self-employed	0.349	0.252	0.000
Big five personality traits			
Conscientiousness	4.383	4.106	0.000
Extraversion	3.692	3.297	0.000
Agreeableness	3.401	3.368	0.300
Neuroticism	2.365	2.677	0.000
Openness	3.723	3.255	0.000
Readiness to take risks	5.777	5.316	0.000
Locus of control	4.260	3.785	0.000

Note: Reported are sample means, unless indicated otherwise. p-values are based on t-tests of equal means. The "big 5" and locus of control are measured on a five-point Likert scale from 1 "does not apply at all" to 5 "applies completely". Readiness to take risks is measured on a 11-point Likert scale from 0 "not at all willing to take risks" to 10 "very willing to take risks". A complete overview of covariates used in the analysis is provided by Table A.1 in the Appendix.

Table 2: Outcome Descriptives

	Outcomes after 20 months			Outcomes after 40 months		
	Part.	Non-Part.	p-value	Part.	Non-Part.	p-value
Number of observations	1,248	1,204		1,248	1,204	
Labor market outcomes						
Share in self- or regular employed	0.958	0.615	0.000	0.938	0.676	0.000
Net monthly earned income (Euro)	1,901	886.6	0.000	2,264	1,046	0.000
Subjective well-being						
Satisfaction with life in general						
Likert points	5.677	5.021	0.000	5.620	5.208	0.000
Share above midpoint	0.863	0.713	0.000	0.845	0.769	0.000
Satisfaction with health status						
Likert points	5.626	5.058	0.000	5.497	5.113	0.000
Share above midpoint	0.812	0.701	0.000	0.807	0.727	0.000
Satisfaction with income						
Likert points	4.580	3.872	0.000	4.591	4.208	0.000
Share above midpoint	0.585	0.427	0.000	0.586	0.497	0.000
Satisfaction with job situation						
Likert points	5.529	4.378	0.000	5.366	4.740	0.000
Share above midpoint	0.823	0.551	0.000	0.785	0.631	0.000
Satisfaction with social security						
Likert points	4.192	4.249	0.382	4.229	4.445	0.001
Share above midpoint	0.466	0.493	0.185	0.477	0.548	0.001
Social security						
Share not contributing to UI	0.534	0.371	0.000	0.603	0.336	0.000
Share not contributing to a retirement plan	0.083	0.016	0.000	0.059	0.021	0.000
Share deeming retirement plan insufficient	0.615	0.628	0.523	0.583	0.594	0.529

Note: Reported are sample means, unless indicated otherwise. p-values are based on t-tests of equal means. The number of observations for the earnings and satisfaction variables is slightly lower due to item non-response. Satisfaction outcomes are measured on a Likert scale from 1 "completely dissatisfied" to 7 "very satisfied". In addition to the mean of satisfaction variables, the table also displays figures on the share of individuals scoring above the midpoint of the Likert scale.

Table 3: Balancing Quality

	Before Matching	After Matching
Number of variables with significant differences in means[a]		
at 1%-level	46	0
at 5%-level	57	1
at 10%-level	61	3
Number of variables with absolute standardized bias[b]		
0% to less than 1%	11	28
1% to less than 3%	5	32
3% to less than 5%	11	21
5% to less than 10%	18	10
more than 10%	46	0
Mean absolute standardized bias in %	15.29	2.43
(Re-)Estimation of the propensity score[c]		
Pseudo-R^2	0.309	0.018
p-value of joint-significance test	0.000	0.996
Other measures		
Rubin's B[d]	142.7	31.9
Rubin's R[e]	0.704	1.389
Number of variables	91	91
Number of participants off support		73

Note: Different indicators are shown for covariate balancing before and after Epanechnikov-kernel matching using a bandwidth of 0.15.

[a]: Number of variables with statistically different means is based on a t-test of equality of means.

[b]: The standardized absolute bias of a variable is the difference in means between treatment and comparison group as a percentage of the square-root of the mean of pre-matched variances of both groups.

[c]: Following Sianesi (2004) Pseudo-R^2 and p-value of joint significance from a logit estimation on the unmatched and the matched sample are also calculated.

[d]: Rubin's B is the standardized mean difference of the linear index of the propensity score of the treatment and comparison group.

[e]: Rubin's R is the variance ratio of the propensity score index of the treatment and comparison sample.

Table 4: Main Estimates

	Outcomes after 20 months		Outcomes after 40 months	
	ATT	p-value	ATT	p-value
Labor market outcomes				
Self- or regular employed	0.266	0.000	0.208	0.000
Net monthly earned income (Euro)	740.9	0.000	910.0	0.000
Subjective well-being				
Satisfaction with life in general				
Likert points	0.152	0.022	0.010	0.865
in %	11.0%		0.8%	
Satisfaction with health status				
Likert points	-0.003	0.960	-0.083	0.169
in %	-0.2%		-5.8%	
Satisfaction with income				
Likert points	0.162	0.116	0.037	0.713
in %	9.5%		2.3%	
Satisfaction with job situation				
Likert points	0.662	0.000	0.278	0.009
in %	37.6%		17.3%	
Satisfaction with social security				
Likert points	-0.512	0.000	-0.567	0.000
in %	-32.2%		-35.6%	
Social security				
Not contributing to UI	0.217	0.000	0.288	0.000
Not contributing to a retirement plan	0.065	0.000	0.048	0.000
Retirement plan deemed insufficient	0.073	0.010	0.057	0.084

Note: The ATT estimates are based on kernel matching using an Epanechnikov kernel, a bandwidth of $h = 0.15$ (chosen by grid-search to maximize post-matching balance in terms of pseudo R^2 in the re-estimated propensity score regression), and common support imposition via min/max criterion using a logit regression. The estimated effects on satisfaction variables are given both in terms of Likert points and in percent of a standard deviation. p-values are estimated by bootstrapping the t-statistic using 999 replications (MacKinnon, 2006; Huber *et al.*, 2014).

Table 5: Effect Heterogeneity

	Male		Female		Age $<$ 45 years		Age $\geq$ 45 years		Low Skilled		High Skilled	
	ATT	p-value	ATT	p-value	ATT	p-value	ATT	p-value	ATT	p-value	ATT	p-value
Labor market outcomes												
Self- or regular employed	0.221	0.000	0.223	0.000	0.153	0.000	0.300	0.000	0.178	0.000	0.247	0.000
Net monthly income (Euro)	1,193	0.000	640	0.000	766	0.000	1,210	0.000	864	0.000	897	0.000
Subjective well-being												
Satisfaction with life in general	0.097	0.240	0.030	0.740	0.110	0.155	0.050	0.694	-0.034	0.757	0.139	0.223
Satisfaction with health status	-0.044	0.634	-0.126	0.258	-0.119	0.202	-0.014	0.902	-0.182	0.099	0.104	0.409
Satisfaction with income	0.270	0.041	-0.067	0.623	0.030	0.781	0.040	0.846	0.087	0.570	0.146	0.409
Satisfaction with job situation	0.363	0.006	0.249	0.117	0.270	0.021	0.348	0.062	0.215	0.127	0.439	0.032
Satisfaction with social security	-0.390	0.008	-0.697	0.000	-0.532	0.000	-0.603	0.000	-0.489	0.000	-0.533	0.001
Social security												
Not contributing to UI	0.270	0.000	0.305	0.000	0.410	0.000	0.143	0.004	0.281	0.000	0.290	0.000
Not contributing to a retirement plan	0.048	0.000	0.044	0.027	0.052	0.000	0.050	0.000	0.027	0.055	0.064	0.000
Retirement plan deemed insufficient	0.085	0.040	0.037	0.398	0.019	0.679	0.132	0.003	0.031	0.487	0.052	0.255
Number of observations												
Participants	717		531		691		557		704		544	
Non-participants	591		613		626		578		432		772	
Balancing quality after matching												
MSB	3.536		3.920		3.510		4.956		3.818		4.326	
Pseudo-R^2	0.042		0.035		0.032		0.067		0.049		0.049	
p-value	0.727		0.999		0.986		0.115		0.207		0.903	

Note: The heterogeneity analysis is based on binary splits of the estimation sample. ATT estimates after 40 months are obtained kernel matching using an Epanechnikov kernel, a bandwidth of $h = 0.15$ (chosen by grid-search to maximize post-matching balance), and common support imposition via min/max criterion using a logit regression. p-values are estimated by bootstrapping the t-statistic using 999 replications (MacKinnon, 2006; Huber *et al.*, 2014). While low skilled refers to having a vocational training degree or less, high skilled refers to having a craftsmanship degree or university degree.

Table 6: Sensitivity with Respect to Matching Approach for Selected Outcomes

	Self- or regular employed		Satisfaction with income		Satisfaction with job situation		Currently not contributing to UI	
	ATT	p-value	ATT	p-value	ATT	p-value	ATT	p-value
Baseline estimates	0.208	0.000	0.037	0.713	0.278	0.009	0.288	0.000
Kernel matching with optimal bandwidth	0.210	0.000	0.042	0.636	0.290	0.004	0.286	0.000
Kernel matching with bias adjustment	0.206	0.000	0.031	0.839	0.262	0.075	0.294	0.000
Inverse probability weighting	0.224	0.005	0.196	0.550	0.353	0.232	0.282	0.000
Radius matching with bias adjustment	0.211	0.000	0.079	0.393	0.301	0.027	0.307	0.000

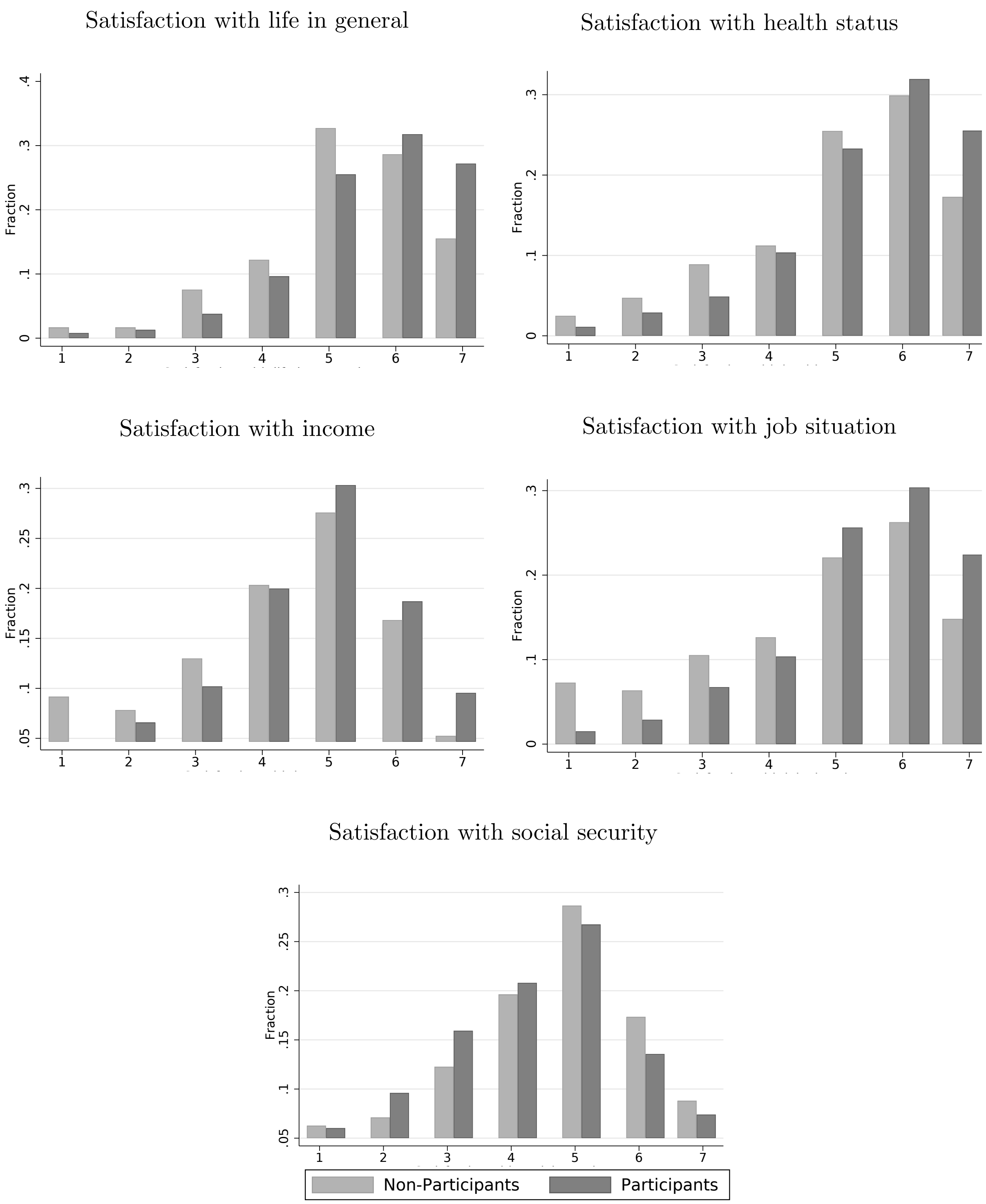
Note: The table reports ATT estimates after 40 months and corresponding p-values for different matching or weighting approaches. We compare our baseline estimates (from Table 4) to kernel matching estimates with an optimal bandwidth as chosen by leave-one-out cross-validation (Frölich, 2005; Galdo *et al.*, 2008). The table also shows results based on kernel matching with post-matching regression, inverse probability weighting (IPW) with weights scaled to unity and bias adjusted radius matching with the radius being equal to 300% of the largest distance in terms of the propensity score when using nearest neighbor matching as suggested by Huber *et al.* (2014). For information on the other outcomes, please see Table A.4 in the Appendix.

Table 7: Sensitivity with Respect to Identifying Assumptions

	Baseline Estimates		Rosenbaum Bounds	Lechner Bounds			
	ATT	p-value	Γ^*	ATT^{low}	p-value	ATT^{high}	p-value
Labor market outcomes							
Self- or regular employed	0.208	0.000	5.030	0.190	0.000	0.249	0.000
Net monthly income (Euro)	910.0	0.000	2.580	767.5	0.000	1,039.6	0.000
Subjective well-being							
Satisfaction with life in general	0.010	0.865	–	-0.070	0.300	0.282	0.120
Satisfaction with health status	-0.083	0.169	–	-0.177	0.002	0.174	0.309
Satisfaction with income	0.037	0.713	–	-0.120	0.331	0.265	0.053
Satisfaction with job situation	0.278	0.009	1.220	0.167	0.122	0.553	0.001
Satisfaction with social security	-0.567	0.000	1.780	-0.718	0.000	-0.284	0.080
Social security							
Not contributing to UI	0.288	0.000	3.220	0.254	0.000	0.313	0.000
Not contributing to a retirement plan	0.048	0.000	4.690	-0.011	0.811	0.047	0.000
Retirement plan deemed insufficient	0.057	0.084	1.250	0.034	0.278	0.092	0.003

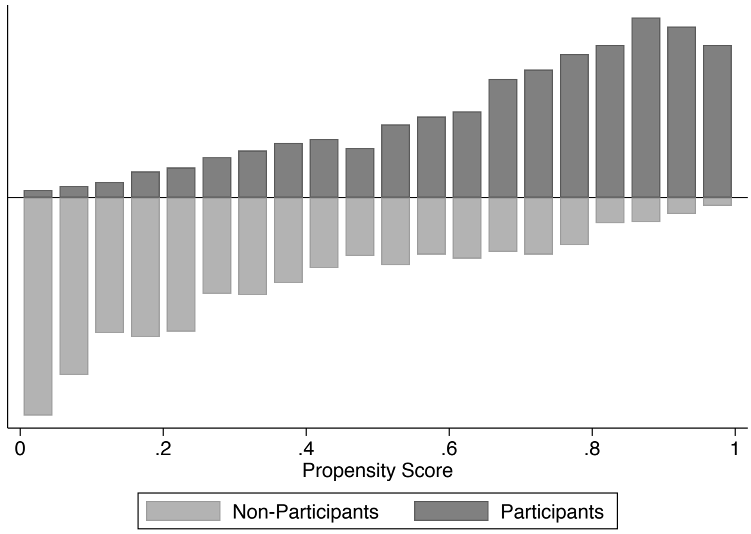
Note: Reported are the results for our assessment of the identifying assumptions for estimates after 40 months. For comparison, column one and two contain our baseline ATT estimates and p-values from Table 4. Next, the table shows the critical values for the Rosenbaum bounds on deviations from the unconfoundedness assumption. The remainder of the table gives results from the Lechner bounding approach regarding deviations from the overlap assumption.

Figure 1: Subjective Well-Being Distributions



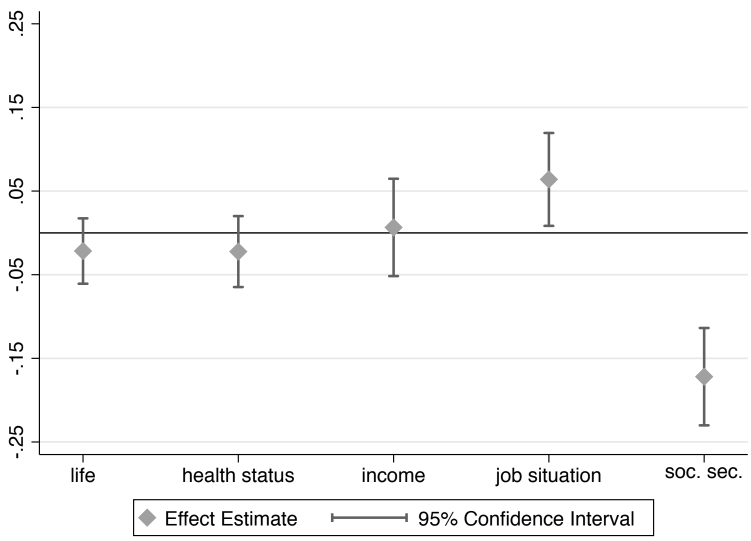
Note: This graph shows the distribution of subjective well-being of participants and non-participants after 40 months as measured by the individuals' self-reported satisfaction with life, their health, income, job situation in general and their social security situation. These items are measured on a seven-point Likert scale from 1 "completely dissatisfied" to 7 "very satisfied'.

Figure 2: Propensity Score Distribution



Note: This graph shows the distribution of estimated propensity scores for the treated and comparison group. The propensity score was estimated based on a logit regression with 91 variables in total, including information on socio-demographics, human capital, labor market history, intergenerational transmission, local macroeconomic conditions, personality traits and some interaction terms chosen to maximize post-matching balance. For the exact specification and estimated coefficients, see Table A.3.

Figure 3: Effect Estimates on Binary Measures of Subjective Well-Being



Note: This graph shows effect estimates on the probability of scoring above the midpoint of the Likert scale for all measures of subjective well-being after 40 months. 95% confidence bands based on bootstrapping are shown.

A Supplementary Tables and Figures

Table A.1: Descriptive Statistics for Control Variables

	Before Matching			After Matching		
	Part.	Non-Part.	p-value	Part.	Non-Part.	p-value
N	1,248	1,204		1,207	1,204	
Socio-demographics						
Age						
(ref.: less than 25 years)						
25 to less than 35 years	0.206	0.192	0.383	0.205	0.208	0.890
35 to less than 45 years	0.336	0.322	0.478	0.340	0.341	0.938
45 to less than 56 years	0.195	0.191	0.817	0.192	0.176	0.310
56 years and older	0.252	0.289	0.037	0.252	0.263	0.523
Female	0.425	0.509	0.000	0.434	0.440	0.767
Not German citizen	0.038	0.017	0.002	0.031	0.043	0.102
Health restrictions	0.038	0.058	0.017	0.039	0.033	0.448
Married	0.599	0.581	0.388	0.600	0.592	0.696
Number of children						
(ref.: no children)						
one child	0.218	0.255	0.031	0.220	0.207	0.462
two children and above	0.212	0.230	0.269	0.216	0.222	0.754
Children under 10 present	0.243	0.218	0.139	0.246	0.250	0.809
Single parent	0.054	0.049	0.600	0.054	0.048	0.509
Human capital						
Highest schooling degree						
(ref.: no schooling degree)						
lower secondary school	0.107	0.198	0.000	0.109	0.107	0.856
middle secondary school	0.278	0.376	0.000	0.283	0.265	0.346
upper secondary school (specialized)	0.170	0.128	0.004	0.170	0.177	0.620
upper secondary school (general)	0.433	0.282	0.000	0.425	0.436	0.580
Professional education						
(ref.: other/no training)						
vocational training	0.371	0.576	0.000	0.383	0.353	0.124
professional/vocational academy	0.153	0.090	0.000	0.149	0.143	0.666
college/university degree	0.411	0.267	0.000	0.407	0.431	0.221
Professional qualification						
(ref.: unskilled workers)						
skilled workers	0.258	0.381	0.000	0.265	0.265	0.979
skilled workers with technical college	0.035	0.031	0.531	0.035	0.035	0.990
top management	0.141	0.109	0.016	0.142	0.166	0.104
Labor market history						
Fraction of time in unemployment in the last 10 years						
(ref.: less than 10%)						
10 to less than 20 %	0.175	0.227	0.001	0.181	0.208	0.087
20 to less than 40 %	0.118	0.199	0.000	0.122	0.118	0.787
40 to less than 60 %	0.027	0.073	0.000	0.028	0.028	0.978
more than 60 %	0.007	0.035	0.000	0.007	0.008	0.808
Five years before (hypothetical) entry						
mean months employed	8.841	8.626	0.266	8.841	8.867	0.893
mean months in labor market program	0.326	0.534	0.002	0.337	0.281	0.356
Four years before (hypothetical) entry						
mean months employed	9.252	8.921	0.069	9.250	9.160	0.618
mean months in labor market program	0.368	0.518	0.030	0.380	0.365	0.815

(Table continued on next page)

(Table A.1 continued)

	Before Matching			After Matching		
	Part.	Non-Part.	*p*-value	Part.	Non-Part.	*p*-value
Three years before (hypothetical) entry						
mean months employed	9.825	9.158	0.000	9.793	9.636	0.344
mean months in labor market program	0.261	0.500	0.000	0.270	0.359	0.115
Two years before (hypothetical) entry						
mean months employed	10.360	9.452	0.000	10.318	9.957	0.013
mean months in labor market program	0.191	0.403	0.000	0.197	0.266	0.162
One year before (hypothetical) entry						
mean months employed	7.773	6.699	0.000	7.721	7.385	0.027
mean months in labor market program	0.401	0.390	0.834	0.400	0.467	0.216
Employment status before entering unemployment (ref.: other)						
dependent employment	0.674	0.516	0.000	0.679	0.651	0.153
self-employment/family worker	0.054	0.012	0.000	0.046	0.050	0.613
school/apprenticeship	0.017	0.020	0.567	0.017	0.018	0.736
disable to work/unemployable	0.033	0.135	0.000	0.034	0.050	0.047
Occupational group before entering unemployment (ref.: other)						
forestry, fishing, animal breeding	0.026	0.027	0.883	0.027	0.022	0.436
manufacturing	0.127	0.204	0.000	0.131	0.133	0.854
technical profession	0.080	0.056	0.020	0.075	0.076	0.947
services	0.765	0.711	0.002	0.766	0.768	0.920
Mean daily income from last employment (Euro)	81.393	61.597	0.000	80.573	82.446	0.422
Duration of last unemployment spell (ref.: less than 1 month)						
1 to less than 3 months	0.388	0.383	0.802	0.394	0.377	0.415
3 to less than 6 months	0.280	0.319	0.034	0.282	0.283	0.953
6 to less than 12 months	0.192	0.224	0.051	0.194	0.209	0.355
12 to less than 24 months	0.027	0.038	0.127	0.028	0.036	0.296
24 months and above	0.009	0.008	0.891	0.009	0.010	0.855
Monthly unemployment benefit (ref.: less than 300 Euros)						
300 to less than 600 Euros	0.127	0.196	0.000	0.129	0.130	0.939
600 to less than 900 Euros	0.201	0.321	0.000	0.204	0.204	0.996
900 to less than 1200 Euros	0.184	0.207	0.160	0.190	0.192	0.885
1200 to less than 1500 Euros	0.145	0.118	0.047	0.146	0.148	0.874
1500 Euros and above	0.253	0.115	0.000	0.244	0.258	0.433
Intergenerational information						
Father and/or mother was born abroad	0.151	0.177	0.079	0.144	0.179	0.019
Father and/or mother is/was self-employed	0.349	0.252	0.000	0.345	0.355	0.590
Father employed when respondent 15 years old	0.913	0.855	0.000	0.911	0.904	0.560
Professional education of father (ref.: other / no training)						
vocational training	0.349	0.428	0.000	0.352	0.351	0.956
professional/vocational academy	0.256	0.196	0.000	0.251	0.221	0.078
technical college/university degree	0.288	0.192	0.000	0.287	0.301	0.427

(Table continued on next page)

(Table A.1 continued)

	Before Matching			After Matching		
	Part.	Non-Part.	p-value	Part.	Non-Part.	p-value
Local macroeconomic conditions						
mean local unemployment rate in %	7.905	7.927	0.853	7.914	7.918	0.975
mean vacancies to unemployed rate in %	16.922	17.235	0.439	16.887	17.105	0.580
mean GDP per capita in 2011(1,000 Euro)	31.263	30.669	0.088	31.267	31.351	0.813
mean local start-up rate out of unemployment	0.058	0.058	0.932	0.058	0.059	0.144
mean local self-employment rate	0.112	0.114	0.002	0.112	0.112	0.494
East Germany	0.333	0.374	0.033	0.334	0.337	0.887
Personality traits and preferences						
Big five						
mean conscientiousness	4.383	4.106	0.000	4.374	4.342	0.256
mean extraversion	3.692	3.297	0.000	3.681	3.692	0.764
mean agreeableness	3.401	3.368	0.300	3.399	3.392	0.845
mean neuroticism	2.365	2.677	0.000	2.371	2.373	0.948
mean openness	3.723	3.255	0.000	3.695	3.674	0.604
mean locus of control	4.260	3.785	0.000	4.249	4.223	0.251
mean readiness to take risks	5.777	5.316	0.000	5.752	5.778	0.756
mean patience	6.370	6.380	0.911	6.374	6.347	0.762
mean impulsiveness	5.388	5.409	0.809	5.389	5.525	0.121
mean general self-efficacy	4.391	4.062	0.000	4.381	4.328	0.012

Note: Reported are sample shares, unless indicated otherwise. Statistics are presented for the raw data as well as after kernel matching. p-values are based on t-tests of equal means.

Table A.2: Business-related Descriptives

Business-related characteristics	
Industry-specific knowledge before start-up	
None	0.154
From regular employment	0.571
From self-employment	0.196
From other sources	0.080
Finances	
Mean start-up capital (Euro)	18,743
Capital is entirely equity	0.602
Sector	
Retail or wholesale	0.110
Construction	0.092
Manufacturing	0.031
Logistics	0.016
Services	0.514
Agriculture	0.014
Other	0.223
Business outcomes after 40 months	
Unconditional outcomes	
Survival share in self-employment	0.805
Outcomes conditional on survival	
Mean net monthly earnings from self-employment (Euro)	2,526
Firms with employees	0.362
Mean number of full-time equivalent employees	0.808
Firms with a patent application	0.017
Firms with a corporate ID protection application	0.058

Note: Reported are sample shares, unless indicated otherwise. To measure full-time equivalent employment by the subsidized start-ups, the number of full-time, part-time and marginally employed individuals are aggregated using weights of one, 0.5 and 0.25, respectively.

Table A.3: Logit Estimation of the Propensity Score

	Coefficient
Socio-demographics	
Age	
(ref.: less than 25 years)	
25 to less than 35 years	-1.128*
35 to less than 45 years	-0.948
45 to less than 56 years	-0.726
56 years and older	-0.780
Female	-0.624***
Not German citizen	1.325***
Health restrictions	0.566*
Married	-0.088
Number of children	
(ref.: no children)	
one child	-0.399**
two children and above	-0.269
Children under 10 present	0.273
Single parent	0.195
Human capital	
Highest schooling degree	
(ref.: no schooling degree)	
lower secondary school	0.755
middle secondary school	0.860*
upper secondary school (specialized)	0.958**
upper secondary school (general)	1.890***
Professional education	
(ref.: other/no training)	
vocational training	-0.306
professional/vocational academy	0.439
college/university degree	-0.205
Professional qualification	
(ref.: unskilled workers)	
skilled workers	-0.163
skilled workers with technical college education	-0.203
top management	-0.136
Labor market history	
Fraction of time in unemployment in the last 10 years	
(ref.: less than 10%)	
10 to less than 20 %	-0.212
20 to less than 40 %	-0.447**
40 to less than 60 %	-1.122***
more than 60 %	-1.551***
Five years before (hypothetical) entry	
months employed	-0.026
months in labor market program	-0.002
Four years before (hypothetical) entry	
months employed	0.015
months in labor market program	0.078*
Three years before (hypothetical) entry	
months employed	0.006
months in labor market program	-0.121**
Two years before (hypothetical) entry	
months employed	0.015
months in labor market program	0.063
One year before (hypothetical) entry	
months employed	0.027
months in labor market program	0.289***

(Table continued on next page)

(Table A.3 continued)

	Coefficient
Employment status before entering unemployment (ref.: other)	
dependent employment	0.583***
self-employment/family worker	1.869***
school/apprenticeship	0.038
disable to work/unemployable	-0.879***
Occupational group before entering unemployment (ref.: other)	
forestry, fishing, animal breeding	-1.159
manufacturing	-2.268
technical profession	-1.749
services	-1.719
Daily income from last employment (Euro)	0.001
Duration of last unemployment spell (ref.: less than 1 month)	
1 to less than 3 months	-0.857***
3 to less than 6 months	-0.823***
6 to less than 12 months	-0.902***
12 to less than 24 months	-0.739*
24 months and above	0.746
Monthly unemployment benefit (ref.: less than 300 Euros)	
300 to less than 600 Euros	-0.543**
600 to less than 900 Euros	-0.901***
900 to less than 1200 Euros	-0.586**
1200 to less than 1500 Euros	-0.342
1500 Euros and above	0.227
Intergenerational information	
Father and/or mother was born abroad	-0.229
Father and/or mother is/was self-employed	0.369***
Father employed when respondent 15 years old	0.181
Professional education of father (ref.: other / no training)	
vocational training	0.035
professional/vocational academy	0.363*
technical college/university degree	0.371*
Eastern Germany	0.020
Local macroeconomic conditions	
local unemployment rate in %	0.015
ratio of vacancies to unemployed	-0.003
GDP per capita in 2011	-0.051**
local start-up rate out of unemployment	-6.487*
local self-employment rate	-5.390
Personality traits	
Big five	
conscientiousness	0.398***
extraversion	0.245***
agreeableness	0.050
neuroticism	-0.079
openness	0.356***
Other traits	
readiness to take risks	-0.240**
locus of control	0.941***
patience	-0.075***
impulsiveness	-0.066**
general self-efficacy	0.293***

(Table continued on next page)

(Table A.3 continued)

	Coefficient
Interaction terms	
self-empl. bef. UE x	
months empl. two years bef.	0.154**
self-empl. bef. UE x	
unempl. for 10 to 20% in last 10 yrs.	-1.193
months in treat. 1 years bef. x	
months in treat. 2 years bef.	-0.132***
last UE-spell longer than 2 years x	
months in treat. 1 year bef.	0.717**
last UE-spell longer than 2 years x	
spent at least 60% in UE in last 10 years	-3.785*
empl. bef. UE x	
months in treat. 1 years bef.	-0.037
last UE-spell longer than 2 years x	
months in treat. 3 years bef.	0.609
last UE-spell longer than 2 years x	
months in treat. 4 years bef.	-1.254**
last UE-spell longer than 2 years x	
months in treat. 2 years bef.	0.334
months in treat. 2 years bef. x	
months in treat. 3 years bef.	0.019
lower sec. school x	
health issues	-1.471**
last daily inc. x	
highest sec. school	-0.008***
risk attitude x	
GDP per capita	0.008**
parents from abroad x	
fath. med. school	0.313
Const.	-2.941
Obs.	2452
Pseudo-R^2	0.31
log-Likelihood	-1173.01

Note: Reported are logit coefficients. ***/**/* denote significance at the 1/5/10 % level.

Table A.4: Sensitivity with Respect to Matching Approach

	Baseline Estimates		Kernel Mat. with optimal bandwidth		Kernel Mat. with Bias Adjustment		Inverse Probability Weighting		Radius Mat. with Bias Adjustment	
	ATT	p-value	ATT	p-value	ATT	p-value	ATT	p-value	ATT	p-value
Labor market outcomes										
Self- or regular employed	0.208	0.000	0.210	0.000	0.206	0.000	0.224	0.005	0.211	0.000
Net monthly income (Euro)	910	0.000	926	0.000	905	0.000	945	0.000	924	0.000
Subjective well-being										
Satisfaction with life in general	0.010	0.865	0.023	0.675	0.011	0.916	0.021	0.880	0.058	0.170
Satisfaction with health status	-0.083	0.169	-0.058	0.334	-0.079	0.451	-0.117	0.360	-0.094	0.243
Satisfaction with income	0.037	0.713	0.042	0.636	0.031	0.839	0.196	0.550	0.079	0.393
Satisfaction with job situation	0.278	0.009	0.290	0.004	0.262	0.075	0.353	0.232	0.301	0.027
Satisfaction with social security	-0.567	0.000	-0.550	0.000	-0.573	0.000	-0.591	0.007	-0.511	0.000
Social Security										
Not currently contributing to UI	0.288	0.000	0.286	0.000	0.294	0.000	0.282	0.000	0.307	0.000
Not currently contributing to a retirement plan	0.048	0.000	0.049	0.000	0.050	0.001	0.039	0.252	0.042	0.006
Retirement plan deemed insufficient	0.057	0.084	0.053	0.091	0.061	0.179	0.093	0.168	0.059	0.102

Note: The table reports ATT estimates and corresponding p-values after 40 months for different matching or weighting approaches. We compare our baseline estimates (from Table 4) to kernel matching estimates with an optimal bandwidth as chosen by leave-one-out cross-validation (Frölich, 2005; Galdo *et al.*, 2008). The table also shows results based on kernel matching with post-matching regression, inverse probability weighting (IPW) with weights scaled to unity and bias adjusted radius matching with the radius being equal to 300% of the largest distance in terms of the propensity score when using nearest neighbor matching as suggested by Huber *et al.* (2014).

B Detailed Data Description and Attrition Analysis

The dataset used for this analysis and also in Bellmann *et al.* (2018) and Caliendo and Tübbicke (2019) was produced in cooperation between the Institute for Employment Research (IAB) of the Federal Employment Agency and the Institute for Social Science Research (Infas). It combines administrative data from the Integrated Employment biographies (IEB) – which cover all individuals who have ever been employed subject to social security contributions or covered by welfare or the unemployment insurance – and two computer-assisted telephone surveys conducted about 20 and 40 months after participants entered the start-up subsidy program. The survey was conducted to obtain outcome measures and enrich the information on socio-demographics, human capital, regional characteristics as well as detailed employment history available through the administrative data with variables on personality traits, preferences and intergenerational transmission.

Sampling design and surveys In a first step, independent random samples of unemployment insurance benefits I recipients who were eligible to enter the start-up subsidy scheme and either applied for the program and received the subsidy or did not apply in month m between February and June 2012 were drawn from the IEB. After duplicates among the non-participants were dropped, they were assigned month m as their hypothetical month of entry as a reference point compared to the actual month of entry among participants. In a second step, participants and non-participants were matched on some basic socio-demographic characteristics (age, sex, region, duration of the previous unemployment spell, employment status at the end of 2011, number of children, qualification, nationality, disability and marital status) using propensity score nearest neighbor matching. Only non-participants who were matched to at least one participant were considered to be interviewed. Moreover, closer matching partners in terms of the propensity score were given a higher priority in terms of being contacted for the survey. The first surveys took place between October 2013 and May 2014, yielding 1,922 and 2,091 interviews. Associated response rates were approximately 60% and 40%, respectively. See also Bellmann *et al.* (2018) and Caliendo and Tübbicke (2019) for more information on the data.

Panel attrition About 89% of first-wave interviewees agreed to be contacted again for the second interview, which were held between July and November 2016. A total of 1,248 participants and 1,204 non-participants could be interviewed again. Hence, attrition rates between the first and the second interviews are roughly 35% and 45%. If this attrition is selective, the panel sample may no longer be representative of the underlying populations and therefore induce bias in the estimation of causal effects. To shed some light on attrition patterns, Table (B.1) shows sample means on basic socio-demographics and our outcome measures for the wave 1 sample as well as the panel sample from the second interview. This comparison shows only relatively small differences in terms of sample means, most of which are highly statistically insignificant. Moreover, comparing sample means across treatment groups for variables with statistically significant differences shows similar attrition patterns. Thus, the attrition analysis yields practically no empirical evidence of a systematic attrition pattern and hence we do not introduce attrition weights in our analysis.

Table B.1: Analysis of Panel Attrition

	Wave 1 Sample	Wave 2 Sample	p-value
Participants			
Number of obs.	1,922	1,248	
Age at (hypothetical) entry			
25 to less than 35 years	0.223	0.206	0.253
35 to less than 45 years	0.353	0.336	0.325
45 to less than 56 years	0.184	0.195	0.441
56 years and older	0.220	0.252	0.039
Female	0.435	0.425	0.579
Eastern Germany	0.338	0.333	0.771
Self- or regular employment	0.950	0.958	0.288
Mean net monthly earned income (Euro)	1,946	1,901	0.449
Mean satisfaction with life in general	5.646	5.677	0.487
Mean satisfaction with health status	5.604	5.626	0.658
Mean satisfaction with income	4.567	4.580	0.824
Mean satisfaction with job situation	5.484	5.529	0.379
Mean satisfaction with social security	4.271	4.192	0.172
Not contributing to UI	0.540	0.534	0.741
Not contributing to a retirement plan	0.095	0.083	0.244
Retirement plan deemed insufficient	0.612	0.615	0.865
Non-participants			
Number of obs.	2,091	1,204	
Age at (hypothetical) entry			
25 to less than 35 years	0.220	0.192	0.058
35 to less than 45 years	0.315	0.322	0.679
45 to less than 56 years	0.187	0.191	0.779
56 years and older	0.263	0.289	0.109
Female	0.500	0.509	0.621
Eastern Germany	0.367	0.374	0.690
Self- or regular employment	0.584	0.615	0.082
Mean net monthly earned income (Euro)	847	886	0.228
Mean satisfaction with life in general	5.039	5.021	0.733
Mean satisfaction with health status	5.022	5.058	0.549
Mean satisfaction with income	3.784	3.872	0.173
Mean satisfaction with job situation	4.369	4.378	0.897
Mean satisfaction with social security	4.262	4.249	0.825
Not currently contributing to UI	0.395	0.371	0.175
Not contributing to a retirement plan	0.019	0.016	0.533
Retirement plan deemed insufficient	0.652	0.628	0.168

Note: Reported are sample shares, unless indicated otherwise. Statistics from the first wave and the panel sample from the second wave are presented. p-values are based on t-tests of equal means.