The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers

www.cesifo.org/wp

Personality Traits and the Evaluation of Start-Up Subsidies

Marco Caliendo
Steffen Künn
Martin Weißenberger

CESIFO WORKING PAPER NO. 5733
CATEGORY 4: LABOUR MARKETS
JANUARY 2016

An electronic version of the paper may be downloaded
- *from the SSRN website: www.SSRN.com*
- *from the RePEc website: www.RePEc.org*
- *from the CESifo website: www.CESifo-group.org/wp*

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 5733*

Personality Traits and the Evaluation of Start-Up Subsidies

Abstract

Many countries support business start-ups to spur economic growth and reduce unemployment with different programmes. Evaluation studies of such programmes commonly rely on the conditional independence assumption (CIA), allowing a causal interpretation of the results only if all relevant variables affecting participation and success are accounted for. While the entrepreneurship literature has emphasised the important role of personality traits as predictors for start-up decisions and business success, these variables were neglected in evaluation studies so far due to data limitations. In this paper, we evaluate a new start-up subsidy for unemployed individuals in Germany using propensity score matching under the CIA. Having access to rich administrative-survey data allows us to incorporate usually unobserved personality measures in the evaluation and investigate their impact on the estimated effects. We find strong positive effects on labour market reintegration and earned income for the new programme. Most importantly, results including and excluding individuals' personalities do not differ significantly, implying that concerns about potential overestimation of programme effects in absence of personality measures might be less justified if the set of other control variables is rich enough.

JEL-codes: C140, L260, H430, J680.

Keywords: start-up subsidies, evaluation, self-employment, personality, treatment effects.

*Marco Caliendo**
University of Potsdam
Chair of Empirical Economics
August-Bebel-Str. 89
Germany – 14482 Potsdam
caliendo@uni-potsdam.de

Steffen Künn
Maastricht University
Maastricht / The Netherlands
s.kuenn@maastrichtuniversity.nl

Martin Weißenberger
University of Potsdam
Potsdam / Germany
martin.weissenberger@uni-potsdam.de

*corresponding author

December 7, 2015

The authors thank Alexander Kritikos, participants at the 2015 CESifo Venice Summer Institute, the 2014 Conference of the European Association of Labour Economists (EALE), the 2014 European Meeting of the Econometric Society (ESEM), the 2014 IZA European Summer School in Labour Economics, the 2014 Potsdam PhD Workshop in Empirical Economics, the 2013 meeting of the Verein für Socialpolitik, the 2013 IZA Workshop on Entrepreneurship Research, the 2013 Conference of the European Society for Population Economics (ESPE), and seminars at the University of Potsdam and the Berlin Network for Labour Market Research (BeNA) for valuable comments. We further thank the Institute of Employment Research (IAB) for cooperation and institutional support within the research project 1755.

1 Introduction

It has been shown that entrepreneurship induces economic growth and lowers unemployment by increasing competition and hence firm productivity, inducing innovation and new technologies and transmitting knowledge spillovers (see Koellinger and Thurik, 2012; Fritsch, 2008; Audretsch and Thurik, 2001; Storey, 1994, for theoretical concepts and empirical evidence). Therefore, many countries provide various support measures to remove existing barriers for nascent entrepreneurs with the goal of increasing the overall start-up rate in their economies. The support ranges from soft measures such as counselling, coaching, training, or technical advice to direct financial support such as subsidised loans, grants, or start-up subsidies for the unemployed. In order to understand whether these programmes indeed achieve their main goal – i.e. fostering successful start-ups – causal empirical evidence is needed. In the past, many evaluation studies have been conducted investigating the effectiveness of soft (e.g. Fairlie *et al.*, 2015; Rotger *et al.*, 2012; Wren and Storey, 2002, among others) as well as hard support measures (e.g. Desiage *et al.*, 2015; Caliendo and Künn, 2011; Tokila *et al.*, 2008; Rodriguez-Planas and Jacob, 2010, among others). Since experimental evidence is very limited, in particular for industrialised countries, most of the studies rely on a comparison of a treated and a non-treated group under the conditional independence assumption (CIA). This is a very strong assumption as it only allows a causal interpretation of the estimated outcome difference between the two groups if all information relevant for the decision to start a business and/or receive support as well as business development and labour market outcomes in general are observed by the researcher. Although the quantity and quality of data has significantly been improved in recent years, in particular due to the better availability of administrative data, there still exist substantial concerns about the justification of using the CIA in the context of the evaluation of start-up support programmes and hence the causal interpretation of treatment effects.

One of the reasons for this scepticism is based on the recent entrepreneurship literature that emphasises the key role of an individual's personality in affecting not only (i) the decision to become an entrepreneur but also (ii) the business development/success over time (for two meta-analytical surveys on this topic, see Rauch and Frese, 2007; Zhao *et al.*, 2010). In a similar vein, it has been shown that personality also affects other labour market outcomes – such

as wages (Heckman *et al.*, 2006), search intensity (Caliendo *et al.*, 2015a; McGee, 2015) and unemployment duration (Uysal and Pohlmeier, 2011) – which are relevant for the estimation of programme effects.

Hence, one major aim of our paper is to bring together these two strands of literature and investigate the role that individuals' personalities play for the estimation of causal programme effects under the CIA. This is especially relevant for our empirical analysis of a new start-up subsidy programme for unemployed individuals in Germany (*"Gründungszuschuss"*). The programme, which was introduced in August 2006 and has replaced two already existing programmes (*"Überbrückungsgeld"* and *"Existenzgründungszuschuss"*), financially supports start-ups from previously unemployed potential entrepreneurs for up to 15 months. While the former programmes have been evaluated extensively (using the aforementioned identifying assumption, see e.g. Caliendo and Künn, 2011, 2015), no evidence for the new programme exists so far. One important advantage for our purpose is the availability of very informative data. In addition to administrative records on programme participants using the start-up subsidy and a comparison group of other unemployed job seekers, we have access to extensive information collected in a survey. Besides information on individuals' family backgrounds and intergenerational transmissions, the survey added specific information on individuals' personalities such as the "big-five," locus of control, and risk preferences. Since this type of data is (usually) not observable in administrative sources, it is ideally suited for our research purpose.

We contribute to the existing literature in three important dimensions: (i) We provide the first empirical evidence on the short- and long-run effectiveness of the new subsidy programme and assess whether the high effectiveness of the former programmes can be confirmed. (ii) Most importantly, we examine the sensitivity of the treatment effects with respect to the inclusion of usually unobserved personality variables in the estimation procedure. Although we do not claim that personality is the only component that was unobserved in earlier studies and, thus, might have biased the results under the CIA, it can be argued, based on the evidence stemming from the entrepreneurship literature, that it plays a significant role. Therefore, this study is of high relevance to the literature as our results will contribute to the discussion of whether earlier results estimated without explicit consideration of individuals' personalities are reliable or not, and it will investigate the necessity to collect personality information in prospective evaluation studies.

And finally (iii), we analyse effect heterogeneity with respect to personality characteristics. While earlier studies revealed heterogeneity regarding education, age, gender, and the regional economic situation, the personality component has yet to be examined due to data limitations. This is especially interesting as the expected relationship between the effectiveness of the start-up subsidy and higher degrees of particular personality characteristics (e.g., is the programme more effective for risk-averse or risk-loving individuals) is ambiguous from a theoretical point of view.

Based on an extensive propensity score specification – including not only a standard set of control variables similar to other studies, such as socio-demographics, labour market history, regional characteristics, and intergenerational determinants of self-employment, but in addition also personality traits – we find positive employment and income effects for the new subsidy programme over the entire 40 month observation period after start-up. We further find that the inclusion of personality variables in addition to the standard set of control variables leads to only small and mostly insignificant changes in the treatment effects. This indicates that the large set of control variables in the estimation of the propensity score, even when not directly controlling for personality, already sufficiently captures individuals' personalities. The analysis on effect heterogeneity reveals that there is only limited interaction between programme effectiveness and personality.

The paper is organised as follows. In Section 2, we provide a summary of the economic rationale behind start-up subsidies for the unemployed, take a closer look at the entrepreneurship literature by discussing recent theoretical concepts and empirical findings about the importance of individuals' personalities on the start-up decision and business development, introduce the institutional setting in Germany, summarise the previous empirical findings and outline the research questions of our study. Section 3 presents the data and some descriptive results. Section 4 discusses the estimation strategy, the potential occurrence of a hidden bias, and the implementation of the propensity score matching approach. In Section 5, we present the main estimation results and robustness analyses before Section 6 concludes.

2 Start-Up Subsidies and the Role of Personality

2.1 The Economic Rationale behind Start-Up Subsidies for the Unemployed

The main justification for the existence of start-up subsidies for the unemployed is based on the assumption that unemployed nascent entrepreneurs face disadvantages compared to business founders out of non-unemployment. Such disadvantages might arise because of more severe credit constraints due to lower financial means or discrimination by credit markets (see Meager, 1996; Perry, 2006), a depreciation of their start-up specific human and social capital during unemployment (Pfeiffer and Reize, 2000; Niefert, 2010), a strong focus on dependent employment during job search resulting from imperfect information (Storey, 2003, refers to it as "lack-of-awareness") and finally a higher share of necessity start-ups due to missing employment alternatives. Caliendo *et al.* (2015c) provide descriptive evidence for the existence of such disadvantages by comparing subsidised start-ups (considering the same programme as under scrutiny here) with regular start-ups in Germany. The start-up subsidy aims at removing such barriers for the unemployed by providing financial assistance to compensate for these disadvantages. Moreover, in a recent study Bianchi and Bobba (2013) show that insurance (instead of credit) constraints are most binding for nascent entrepreneurs, i.e., they are hindered by the (financial) risk of failure. In this sense, the subsidy can be considered as insurance against the risk of low or no income during the start-up period, stimulating nascent entrepreneurs among the unemployed to start a business. However, the existence of the subsidy might also induce some negative effects such as adverse selection, moral hazard, and deadweight effects (see Caliendo *et al.*, 2015c, for a discussion and empirical evidence).

2.2 The Influence of Personality in Entrepreneurship Research

The entrepreneurship literature stresses the importance of personality with respect to business creation and performance. Entrepreneurs identify opportunities and create businesses to pursue them (Bygrave and Hofer, 1991). Setting up and successfully managing a business is inherently related to higher degrees of personal initiative and discretion, making risky decisions in uncertain environments, showing perseverance in the face of obstacles, setbacks, and stress, as well as setting up and maintaining relationships with investors, suppliers, and clients. A comparison of personality characteristics with the tasks required in an entrepreneurial context yields intuitive

indications about the direction of the relationship. One of the most widely examined personality constructs in this context is the five-factor model, commonly referred to as the "big five" with its dimensions conscientiousness, extraversion, agreeableness, neuroticism, and openness (Costa and McCrae, 1992; McCrae and Costa, 2008). On top of that, more specific personality constructs have gained attention, most importantly risk attitudes (Chell *et al.*, 1991), locus of control (Rotter, 1966), achievement orientation (McClelland, 1965), self-efficacy (Baum and Locke, 2004), or innovativeness (Heunks, 1998).[1] We restrict the discussion below to the personality characteristics available in our dataset, i.e., big five, locus of control, and readiness to take risks. Given that these items are highly correlated with others like achievement orientation or self-efficacy (Judge *et al.*, 2002), we are confident that we capture the most important variables, without making any claim that all usually unobservable personality-related factors are included.

Personality and Start-Up Comparing the entrepreneurial tasks with the attributes associated with each of the personality dimensions, the following intuitive predictions can be made (see, e.g., Zhao *et al.*, 2010; Caliendo *et al.*, 2014a, and column 1 in Table 1 for a summary): The decision to set up a business should be related positively to extraversion (indicating higher levels of ambition and optimism, seeking leadership roles) and openness (higher levels of creativity) and negatively to neuroticism (higher levels of self-confidence and self-esteem, less vulnerable to psychological stress in face of challenges). Conscientiousness (dedication, perseverance, efficiency) might only become more relevant once the business is set up, such that the influence on the start-up decision is ambiguous. The same is true for agreeableness, where both extremes of the factor – high values (trusting, altruistic, cooperative) and low values (self-centered, hard-bargaining, suspicious) – might have positive and negative effects on the entry decision. For internal locus of control a positive association is expected reflecting that individuals who believe their own actions determine their future outcomes are more likely to actively pursue new business opportunities (Rauch and Frese, 2007). Given that being self-employed is a more risky occupational choice, a positive relationship with readiness to take risks is assumed (Cramer *et al.*, 2002; Ekelund *et al.*, 2005).

[Insert Table 1 about here]

[1] This list is not intended to be exhaustive. For meta-analytical surveys on the topic, see, e.g., Stewart and Roth (2001), Rauch and Frese (2007), and Zhao *et al.* (2010).

The empirical evidence on these hypotheses is less clear-cut. While Zhao *et al.* (2010) conclude that entrepreneurial intentions are positively related to conscientiousness, extraversion, and openness, and negatively to neuroticism, Caliendo *et al.* (2014a) find significant positive impacts on actual start-up activity only for extraversion and openness. Further, a positive link between start-up activity and internal locus of control (Evans and Leigthon, 1989; Caliendo *et al.*, 2014a) as well as risk tolerance (Cramer *et al.*, 2002; Ekelund *et al.*, 2005; Skriabikova *et al.*, 2014) is supported by empirical results. When taking gender differences into account, findings for openness and risk hold for both men and women whereas the predictive power for start-up activity of extraversion and locus of control is confirmed only for men (Hansemark, 2003; Caliendo *et al.*, 2009, 2015b).

Personality and Business Success The influence of personality on professional behaviour and success is likely to be stronger for entrepreneurs compared to most other professions due to the characteristics of the entrepreneurial role itself (Brandstätter, 2011). The hypotheses regarding business survival can be derived with a similar intuitive reasoning as mentioned above for the start-up activity, but some noteworthy deviations occur (see, e.g., Ciavarella *et al.*, 2004; Zhao *et al.*, 2010; Caliendo *et al.*, 2014a, and column 2 in Table 1). The positive relations to conscientiousness (higher work motivation, dedication, perseverance, and efficiency), extraversion (higher levels of assertiveness, advantages in developing and maintaining social networks with investors, suppliers, and customers), and internal locus of control as well as the negative link to neuroticism (higher levels of stress-tolerance and self-security, less prone to anxiety and depression) are straightforward. With respect to agreeableness, higher levels, that imply more trusting and cooperative business relations with stakeholders, might be beneficial on the one hand (Ciavarella *et al.*, 2004), whereas on the other hand, too agreeable entrepreneurs might lack bargaining abilities and the required ruthlessness to survive (Zhao *et al.*, 2010). The expectations for openness are also ambiguous as higher levels (innovative thinking, creativity) might be less important once the business is set up (Baron and Markman, 2005). Finally, entrepreneurs are required to manage risk to preserve sustainability and avoid too risky investments that could lead to large losses, resulting in an inverse-u shaped relationship with business success (Chell *et al.*, 1991, Zhao *et al.*, 2010.)

The empirical evidence on business survival finds support for the positive impacts of conscientiousness (Ciavarella *et al.*, 2004) and internal locus of control (Rauch and Frese, 2000), the negative link with agreeableness (Caliendo *et al.*, 2014a), and suggests a negative association with openness (Ciavarella *et al.*, 2004). Finally, the assumed inverse u-shaped relation of risk tolerance on entrepreneurial survival also finds empirical support (Caliendo *et al.*, 2010a).

Personality and Programme Effectiveness In the heterogeneity analysis, we examine the interplay between personality and effectiveness of the subsidy programme to provide the first empirical evidence on which type of individual benefits most from participation. As we will show in the following, it is not obvious from a theoretical point of view if individuals with higher degrees in a particular personality variable are expected to benefit more or less from the start-up subsidy programme. To derive hypotheses about the direction of this interaction, we need to examine the net effect of two distinct relationships: First, the connections between personality and business survival, as elaborated above, yield indications in the case of participation. Second, the link between personality and exit from unemployment (column 3 in Table 1) reveals the direction of the impact of personality in the counterfactual situation of non-participation. The net effect of these two gives the expected interaction effect between personality and programme effectiveness (column 4). The literature on the influence of personality on job search behaviour and the transition out of unemployment is very scarce. Uysal and Pohlmeier (2011) hypothesise that exit from unemployment is positively connected to conscientiousness and openness, whereas they assume a negative relation to neuroticism. Further, they reason that, ex ante, the link to agreeableness is ambiguous and stress the context-dependency (e.g., job, sector) of these expectations. With respect to locus of control, McGee (2015) and Caliendo *et al.* (2015a) argue that a more internal locus of control predicts both a higher search intensity as well as higher reservation wages. As a consequence, the net effect on transitions from unemployment is ambiguous. Furthermore, it is assumed that higher risk-aversion is associated with lower reservation wages and thus shorter unemployment durations (Pissarides, 1974).

Empirically, Uysal and Pohlmeier (2011) find support for the positive link with conscientiousness and openness as well as for the negative influence of neuroticism on transitions from unemployment. The empirical results for internal locus of control are inconclusive, with McGee

(2015) finding no significant impact, whereas Caliendo *et al.* (2015a) report a positive net effect on the probability of leaving unemployment. The evidence on risk attitudes is also less clear cut: while more risk-loving individuals tend to have higher reservation wages (Pannenberg, 2010) and longer durations in unemployment (Feinberg, 1977), Diaz-Serrano and O'Neill (2004) find that they are less likely to be unemployed. The findings in Oberschachtsiek and Ullrich (2010) meanwhile suggest a non-linear pattern between risk aversion and unemployment duration.

2.3 Institutional Setting in Germany

The new start-up subsidy (*"Gründungszuschuss"*, SUS) offers unemployed job seekers financial support to start their own business and hence to escape unemployment. The programme was introduced in August 2006 and replaced two already existing start-up subsidy programmes, the bridging allowance (*"Überbrückungsgeld"*) and a former version of the start-up subsidy (*"Existenzgründungszuschuss"*) (see Caliendo and Kritikos, 2010, for a description). The SUS programme pays a subsidy for a maximum duration of 15 months after start-up which is split into two parts: (i) All sponsored individuals receive a monthly amount equivalent to the individual's last unemployment benefit and a lump sum of 300 euros to cover social security costs for nine months. (ii) Afterwards, individuals can apply for an optional second period (no legal claim) by proving sufficient business activity. Based on the caseworker's discretion, individuals received the lump sum payment for another six months.[2] Eligibility for the SUS programme required unemployed individuals to have a minimum entitlement to *unemployment benefit I*[3] of at least 90 days at the time of programme start. Moreover, individuals applying for the subsidy had to provide a business case and financing plan to the Employment Agency, which had to be evaluated by a competent external institution. Between 2007 and 2011 (afterwards the programme conditions changed), around 130,000 job seekers entered the start-up subsidy programme per year, resulting in public annual expenditures of about 1.6 billion euros (e.g., compared to 0.8 billion euros for vocational training). This illustrates that the programme is an integral part of the German Active Labour Market Policy (ALMP) and is expected to remove existing dis-

[2] 65.5% of the business founders received the subsidy for 15 months in our sample.

[3] In Germany, every individual who has been in employment subject to social security for at least one out of the last three years is eligible for unemployment benefit I. The amount of the benefit consists of 60% (67% with children) of the last net wage and is basically paid for a period of 12 months, with the exception of older individuals (see Caliendo and Hogenacker, 2012).

advantages faced by nascent entrepreneurs among the unemployed compared to the employed workforce.

2.4 Previous Evidence and Research Questions

While the effectiveness of this new programme has not been examined yet, evaluation studies on the two former programmes show very positive results in terms of employment and income (Caliendo and Künn, 2011, 2015, 2014) that are larger than those reported for traditional ALMP programmes such as training or wage subsidies. All of these studies are using a propensity score matching approach – the workhorse in this literature – and claim that the rich data at hand allows them to control for all relevant variables to make the CIA a reliable assumption. A similar picture arises from international evidence on start-up programmes for the unemployed. The identification of causal programme effects is most often based on the CIA, and the findings are predominately positive based on propensity score matching approaches (see, e.g., Desiage *et al.*, 2015, for France, O'Leary *et al.*, 1998, and O'Leary, 1999, for Hungary and Poland, Rodriguez-Planas and Jacob, 2010, and Rodriguez-Planas, 2010, for Romania, or Perry, 2006, for New Zealand).

The large positive results for Germany, in particular in comparison to other ALMP programmes, and other countries raise concerns about the justification of the CIA in this context, i.e., whether all relevant aspects are included in the vector of observable characteristics or whether the results are still affected by (at this time) unobserved factors. These concerns stem primarily from the growing entrepreneurship literature stressing the importance of personality as outlined above. However, due to data limitations, this aspect has not (directly) been included in existing evaluation studies estimating causal programme effects under the CIA, and hence, the previous very positive findings might be biased as important personality variables were missing.[4] Therefore, the central question in this paper is whether the inclusion of personality variables, in addition to other control variables as used in earlier studies, would change the estimation of treatment effects significantly. On top of that, we will further provide the first long-term

[4]For the evaluation of traditional ALMP like training and wage subsidy programmes under the CIA, Caliendo *et al.* (2014b) find no significant differences in treatment effects when including these measures in addition to a standard set of control variables. However, this evidence is not directly adoptable for the evaluation of business support programmes given that they are likely to be most prone to remaining selection bias due to unobserved personality variables because they involve a higher level of individual initiative, risky decisions, and uncertainty.

evidence on the effectiveness of the new SUS programme. This is interesting itself as the new programme combines elements of the two earlier programmes, leading to a different selection of participants where the average participant is more similar to the former bridging allowance than the former start-up subsidy (Caliendo *et al.*, 2012). The question thus arises as to whether the new programme is as successful as its two predecessors. Moreover, we will have a closer look at effect heterogeneity with respect to personality. While earlier evaluation studies have shown that start-up subsidy programmes are particularly effective for certain subgroups of the labour market, e.g., women, low-educated, or low-qualified individuals (see Caliendo and Künn, 2011, 2015), the question remains whether the effects also vary with personality characteristics.

3 Data and Descriptive Results

3.1 Estimation Sample

For the empirical analysis, a random sample of unemployed individuals who entered the subsidy programme in the first quarter of 2009 serves as our treatment group; a sample of other unemployed job seekers who did not join the programme during that period are the control group.[5] The data combines administrative information (Integrated Employment Biographies) provided by the Institute for Employment Research (IAB) with survey information collected in telephone interviews. The survey is constructed as a panel, where the same individuals were interviewed twice – in the last quarter of 2010 (21 months after business start) and in the fall of 2012 – such that we observe all respondents for 40 months after business start-up.

From the administrative data, we obtain detailed information on the time prior to participation in the start-up programme, including spells in employment and participation in ALMP programmes as well as wages and unemployment benefits. For the period after entry into the subsidy, we use the survey information to calculate labour market outcomes, as spells in self-employment are not recorded in the administrative data. Moreover, the survey allows us to observe characteristics usually not included in the administrative records such as parental self-employment. In addition, and central to our analysis, the questionnaire contains items measuring

[5]Non-participants were selected by a pre-matching procedure, i.e., those most similar to participants in key socio-demographic characteristics were selected. A fictitious entry month for the programme was attributed to each non-participant which corresponds to the actual entry month of the pre-matching partner in the participant group. We further note that having access to only one particular quarter of entrants might restrict the external validity of the results if the composition of subsidized business founders has changed over time.

various personality characteristics, e.g., for the big five (locus of control), respondents were given 10 (6) different statements about themselves and were asked how much they agreed with them on a seven-point Likert scale. Risk preferences were measured on a scale ranging from zero to ten, where higher values indicate a higher readiness to take risks. The item wordings and the construction of the variables are documented in Table B.1.2 in the Supplementary Appendix and are similar to other questionnaires such as the German Socio-Economic Panel (SOEP). It has to be noted, however, that the personality variables in our data were surveyed more parsimoniously than in other surveys, such that we need to be careful with the interpretation for some of the traits. The personality characteristics were surveyed during the second interview and thus recorded after the programme start. Following the literature, we assume in our analysis that personality variables are exogenous and thus not related to labour market events, i.e., unaffected by the entry into the start-up subsidy programme and subsequent success.[6]

[Insert Table 2 about here]

Table 2 shows the definition of our estimation sample. Initially, 2,306 participants and 2,307 non-participants were interviewed in the first wave. We use a 50% random subsample for which the information on personality and business characteristics was collected, and we further consider only individuals who participated in the second interview, gave their consent to link their survey information to the administrative data, and responded to all questions relevant for our analysis. Our final estimation sample consists of 589 participants (367 men and 222 women) and 699 non-participants (439 men and 260 women). A selectivity analysis at each step in Table 2 yields practically no empirical evidence for a systematic attrition pattern.[7] The gender composition in our treatment group of 62% men and 38% women is relatively similar to the general shares observed for business founders in Germany in 2009 (cf. Fritsch *et al.*, 2012). Since start-up decisions and actually founded businesses are very different across gender (Georgellis and Wall, 2005; Caliendo *et al.*, 2015b), we conduct our analysis separately for men and women.

[6]Personality variables are shown to be relatively stable over the adult life-cycle and not related to lifetime events in a meaningful way (Cobb-Clark and Schurer, 2012, 2013). Although we cannot explicitly test the exogeneity of personality variables in our setting, Hamilton *et al.* (2015) did not find any evidence for simultaneity or reverse causality of personality variables and self-employment status or earnings.

[7]Detailed results for the selectivity of attrition analysis are available in the Supplementary Appendix B.2.

3.2 Descriptive Results

Table 3 presents selected descriptive statistics with respect to basic individual characteristics at start-up (Panel A), personality characteristics (Panel B), and labour market outcomes 21 months (Panel C) and 40 months after business start (Panel D). Results are reported separately for male (columns 1 through 3) and female (column 4 to 6) participants and non-participants.

[Insert Table 3 about here]

Socio-Demographics and Labour Market History: Both male participants (column 1) and non-participants (column 2) are on average 41 years old, and more than 50% have completed upper secondary school. While basic socio-demographics are balanced between the two groups,[8] we find that participants are characterised by a higher labour market attachment in the past but do not substantially differ in benefit levels from their non-participant counterparts. For women (columns 4 and 5), the comparison of the two groups yields a similar picture. Between men and women, however, we find the usual differences. Women are less attached to the labour market, earn less and have stronger family commitments, irrespective of participation status.

Personality: Both male and female participants are characterised by stronger 'entrepreneurial personality' characteristics than non-participants (as expected in Table 1, column 1). For instance, participants are significantly more extraverted, more open to new experiences, and have a more internal locus of control. Also, male participants show a higher willingness to take risks, while female participants are significantly more confident (less neurotic). A comparison of men and women indicates meaningful differences in mean levels, again irrespective of participation status. For instance, women have higher values in the big five, while men show a higher readiness to take risks. This reinforces our decision to analyse men and women separately.

Labour Market Outcomes: The relatively high descriptive shares of self-employed participants in the short- (after 21 months) and long-run (after 40 months) indicate a persistent integration into self-employment of a striking majority of former subsidy recipients. Given a 15 month maximum duration of the subsidy, 77% of male and 69% of female participants are

[8] This is not surprising given the pre-matching procedure of participants and non-participants with respect to key socio-demographic characteristics mentioned above.

self-employed two years after the subsidy expired. A comparison of self-employment rates between participants and non-participants is not very informative, however, as all participants are self-employed at the start of the programme by definition, whereas non-participants might seek dependent employment instead. Thus, we focus on a joint employment outcome, i.e., self- or regular employment subject to social security contributions. For male participants, we find employment shares consistently higher than 90% while for females the numbers are slightly lower. Both experience an advantage in employment chances over non-participants of around 15 points each. In addition to employment, we also consider earned income. Former male participants earn 2,813 euros/month after 40 months, whereas the mean monthly income in the male comparison group amounts to 1,836 euros. Conditional on being employed, the difference between the two groups is less pronounced (3,003 euros versus 2,300 euros) but still meaningful.[9] For women, we again observe a similar pattern between participants and non-participants but on substantially lower absolute levels compared to men.

[Insert Table 4 about here]

Business Outcomes: We further consider business outcomes for those 77% male and 69% female participants who are self-employed after 40 months (see Table 4). First, focussing on the long-run outcomes after 40 months (column 2) shows that men are working on average 50 hours per week with an hourly income of 15 euros. While a majority of male self-employed remain solo-entrepreneurs, around 40% create, on average, 3.6 full-time equivalent jobs. Formal indications of innovation show relatively low rates. For female founders in our sample (column 4), we observe lower weekly working time, less innovation implemented, and a lower amount of job creation in terms of both the extensive and intensive margin. Second, a comparison of the business outcomes achieved after 20 and after 40 months reveals that businesses show improvements in growth, productivity, innovation, and job creation over time. This finding holds true for both genders.

[9]To set these figures in perspective, the German Federal Statistical Office (2012, p. 106) reports average monthly gross earnings of 2,976 Euros for a male full-time worker in dependent employment in Germany in the third quarter of 2012 (when income measures 40 months after start-up were surveyed). Applying a net-to-gross ratio of 70%, assuming a three-person household (married couple, one child) with one breadwinner and residence in West Germany (Federal Statistical Office, 2013, p. 10), this translates into net monthly earnings of 2,083 euros. Although the range of earned incomes among the group of employed male participants is pretty broad, the mean income level 40 months after start-up clearly exceeds this benchmark.

4 Estimation Strategy

4.1 Identification of Causal Treatment Effects under the CIA

The aim of the paper is to estimate the causal impact of participating in SUS on labour market outcomes and to investigate the sensitivity of the treatment effects to the inclusion of personality variables. Similar to the majority of evaluation studies in the past (see Section 2.4), we do this by applying a propensity score (PS) matching approach. While matching is easy to implement, the validity of its results hinges on the strong identifying conditional independence assumption (CIA).

To illustrate the idea behind PS matching, we use the well known potential outcome framework (Roy, 1951; Rubin, 1974). The two potential outcomes are denoted as Y^1 (in case of treatment) and Y^0 (in case of non-treatment). We focus on the usual parameter of interest in most evaluation studies, the average treatment effect on the treated (ATT):

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

where D is a binary treatment indicator. The fundamental evaluation problem arises because the last term on the right hand side of equation (1) is not observed. If participants and non-participants are selected groups in terms of (un)observed characteristics who would have different potential outcomes even in the absence of treatment, selection bias arises.

To correct for this selection bias, propensity score matching estimators rely on the unconfoundedness or conditional independence assumption (CIA), which states that conditional on observed characteristics (X), the counterfactual outcome is independent of treatment (Rosenbaum and Rubin, 1983). In addition, we also assume overlap: $Pr(D = 1 \mid P(X)) < 1$, for all X. The ATT is then identified as:

$$\Delta_{ATT}^{MAT} = E(Y^1 \mid P(X), D = 1) - E_X\Big[E(Y^0 \mid P(X), D = 0) \;\Big|\; D = 1\Big], \tag{2}$$

where the counterfactual situation can now be estimated from the mean outcomes of the matched control group, i.e., taking the outer expectation over the distribution of $P(X)$ in the treatment group.

The CIA is a very strong assumption which relies heavily on the availability of relevant data that allow the researcher to control for all relevant variables that simultaneously influence the

participation decision and the (untreated) outcome variable (Lechner and Wunsch, 2013). In previous evaluation studies of start-up subsidies for the unemployed, it has been argued that controlling for individual socio-demographic and qualification factors along with information on labour market history and parental self-employment makes it plausible that the CIA holds (see, e.g., Caliendo and Künn, 2011). This can be criticised because crucial information was missing, in particular, due to the recent findings in the entrepreneurship literature stressing the key role of personality characteristics affecting the decision to start a business and the subsequent performance (see Section 2.2).

4.2 The Risk of Hidden Bias

If the concerns are justified and the missing variables indeed have a significant impact on the selection into the programme and labour market outcomes, a *hidden bias* might arise to which the above defined Δ_{ATT}^{MAT} is not robust (see Rosenbaum, 2002; Caliendo *et al.*, 2014b, for an extensive discussion and recent application). To illustrate the underlying idea, we introduce a vector U in addition to the usually observed vector X and assume that the participation probability depends on both sets of variables. The participation probability can then be specified as:

$$P(D = 1 \mid X, U) = F(\beta X + \gamma U), \tag{3}$$

where γ is the effect of U on the participation decision. If $\gamma = 0$, the study is free of hidden bias and the participation decision is solely determined by X. However, if there is hidden bias, two individuals with the same observed covariates X have different chances of receiving treatment. The magnitude of the bias depends on γ and the correlation between X and U.

In contrast to earlier studies evaluating the effectiveness of start-up subsidies, we now have access to more informative data which allow us to observe the standard set of control variables used in earlier studies (X) and, in addition, usually unobserved characteristics (U) such as personality and risk preferences. Therefore, we can now model the selection process with and without personality variables (U) and compare the estimated treatment effects. Thus, we can examine the sensitivity of the effects of the start-up subsidy programme with respect to the availability of personality variables. Finally, we clearly emphasise that we do not claim that personality is the only component that was unobserved in earlier studies which might have

biased the results under the CIA. However, based on the evidence as presented in Section 2.2, it can be argued that it is likely to be a significant part of U.

4.3 Propensity Score Estimation and Matching Quality

The first step of our matching routine is to estimate the propensity score based on a probit model. The specification of the model is primarily guided by previous evaluation studies (see e.g. Caliendo and Künn, 2011, for an elaborate discussion on the choice of variables). It contains a rich set of detailed information on socio-demographics, intergenerational transmissions, regional labour markets, human capital, and labour market history including details on the unemployment spell preceding start-up (summarised as X). On top of these standard controls, we extend our model by usually unobserved personality measures of the big five, locus of control, and readiness to take risks (summarised as U). The results of the probit estimation in Table A.1 in the Appendix show that, for both men (column 3) and women (column 6), openness and locus of control have a positive and significant influence on the start-up decision, in line with our theoretical expectations (cf. Table 1, column 1). For the other big five factors as well as risk attitudes, we do not find any significant impact. At first glance, especially the insignificant result on risk might be surprising, but it is in line with previous empirical evidence that suggests that risk preferences do not play a role in start-up decisions for unemployed or inactive individuals (Caliendo *et al.*, 2009). Overall, the additional set of information on personality characteristics has a significant impact on the decision to participate in the start-up programme and improves the overall model fit, as indicated by the joint significance test for all personality variables. Among the conventional controls, primarily variables on household composition, parental self-employment, characteristics on labour market history, the current unemployment spell, as well as the regional cluster influence the selection into the programme.

[Insert Figure 1 and Table 5 about here]

Figure 1 plots the distributions of the estimated propensity scores for participants and non-participants, separately by gender. As expected, the distributions for men (Figure 1a) and women (Figure 1b) are both asymmetric between participants and non-participants and skewed towards the tails. Hence, participants have, on average, a higher probability to enter the programme.

Although we find individuals in each group along the entire distribution of the propensity score, there is only limited overlap between participants and non-participants in the tails. To ensure that we only compare individuals with similar values of the propensity score, we impose common support by excluding treated observations with a propensity score above (below) the maximum (minimum) value in the non-participant group.

In the second step of the matching routine, we implement an Epanechnikov kernel matching algorithm with optimal bandwidth choice based on leave-one-out cross-validation.[10] The estimator choice is based on results by Huber *et al.* (2013), who assess the finite sample properties of different matching estimators and find kernel matching with optimally chosen bandwidth parameters to perform very well in particular with small sample sizes as in our case. Based on this matching algorithm, Table 5 reports different indicators summarising the very good matching quality for both the male (column 1 and 2) and female sample (column 3 and 4) for an optimal bandwidth of 0.24.[11] While the characteristics of male (female) participants and non-participants differ significantly in 25 (12) out of 72 covariates in the unmatched sample, all significant differences disappear at the 5% (10%) level in the matched sample (Panel A). Matching also reduces the mean standardised bias (Rosenbaum and Rubin, 1985) from 10.0% (9.2%) before matching to 3.7% (3.6%) after matching (Panel B), confirming the good matching quality, which is usually characterised by values lower than 3-5% (Caliendo and Kopeinig, 2008). In line with these results, Panel C reports for both genders that the Pseudo-R^2 from a probit reestimation of the propensity score for the matched sample (Sianesi, 2004) sharply decreased compared to the unmatched case, whereas the p-value of joint significance test increases to 1. Thus, the included characteristics have no significant explanatory power for the selection into the subsidy programme after matching, implying a successful matching procedure.

[10] See Table B.1.3 in the Supplementary Appendix for details on the bandwidth choice. Given that the implementation of the matching estimator might affect our results, we test the sensitivity of the results with respect to the matching algorithm, definition of the estimation sample and common support in Section 5.4.

[11] Matching quality results for all optimal bandwidths are presented in Table B.1.4 in the Supplementary Appendix.

5 Estimation Results

5.1 The Effectiveness of the New Start-Up Subsidy in Germany

We start the discussion of our estimation results with an answer to our first research question. Using the extended specification including personality variables to estimate the propensity score, we find strong positive effects of participation in the subsidy programme on employment and income (see Table 6, column 3). It can be seen from Panel A (B) that 40 months after start-up, former male (female) participants face an 8%-point (10.5%-point) higher probability to be in self- or regular employment than matched non-participants. In addition to the static effect, the black solid lines in Figure 2 show the effects at each month during our observation window. The effects are positive throughout but decrease over time. The ATT drops from initially 60%-points to 20%-points after one year and becomes somewhat stable at 8-10%-points after 30 months. For females, the development of the effects is similar over time but on a slightly higher level.

[Insert Figure 2 and Table 6 about here]

If we cumulate all monthly effects over the entire observation period, we find that male (female) participants spent on average 7.8 (8.6) months more in self- or regular employment than matched non-participants (see Table 6, column 3 of Panel A and B). With respect to income, we also find statistically significant positive effects. For instance, 40 months after start-up, male (female) participants earn on average 740 (610) euros per month more than matched non-participants. Given the average working income of 2,800 (1,600) euros for participants (see Table 3, Panel D), this treatment effect is substantial but can be partly attributed to the significant gap in the employment probability between participants and matched non-participants.

The overall positive results for the new start-up subsidy confirm previous findings for its two predecessors in Germany (as reported by Caliendo and Künn, 2011; Caliendo *et al.*, 2010b), indicating that the programme is an effective tool in helping unemployed individuals reintegrate into the labour market. Compared to both previous schemes, though, the ATTs for the new subsidy with regard to labour market integration after 40 months are with 8-11%-points substantially smaller in magnitude (for the old programmes the ATTs were in the area of 20-30%-points). On the one hand, the effect differences might arise due to the institutional changes resulting in a different selection pattern of participants, or the different observation periods with different

economic conditions. The latter point is empirically supported by the (descriptive) shares of participants in self- or regular employment after 40 months. These shares are higher for the new subsidy programme (around 90%) compared to its two predecessors (closer to 80%). On the other hand, however, the lower ATTs for the new programme might also be explained by the additional consideration of individuals' personalities in the estimation process. Due to data limitations, this was not possible in evaluation studies on the two former programmes, and hence, the estimated effects might have been biased. We will take a closer look at this issue in the next section.

5.2 Influence of Personality on the Estimation of Programme Effects

5.2.1 Estimation Results

To investigate whether the inclusion of individuals' personalities indeed has a significant impact on the estimation of programme effects, we rerun the matching procedure (estimation of the propensity scores and ATTs), yet this time, we exclude the usually unobserved personality variables (U). We thus end up with a specification similar to those used in previous evaluation studies of start-up subsidy programmes in Germany that contains detailed information on socio-demographics, intergenerational transmissions, regional labour markets, human capital, and labour market history including the unemployment spell preceding start-up (X).

As shown by the dotted grey line in Figure 2a, the estimated ATTs for men based on this standard specification are very close to the ones with the extended specification over the whole observation period with a small deviation upward. For women (Figure 2b), the upward differences are slightly more pronounced but overall appear moderate in size as well. Full estimation results for the standard specification are presented in column 2 of Table 6 whereas column 5 reports the differences in ATTs between both specifications, where p-values are based on a bootstrapped robust Hausman test.[12] For instance, while we estimate a cumulated effect of 7.88 months more in self- or regular employment over the observation period for male participants compared to matched nonparticipants (column 3), a specification ignoring the personality variables yields an effect of 8.12 months (column 2). The difference of 0.36 months (column 5) implies an insignificant overestimation of 4% if we neglect personality measures. Overall, we find a relatively

[12]The robust bootstrapped Hausman test does not require one of the estimators to be fully efficient under the null hypothesis, see Cameron and Trivedi (2009, p. 443f.) for details.

consistent pattern. The evidence suggests that taking personality characteristics U into account corrects for a positive selection that remains even after having controlled for a large set of other important characteristics X. However, the differences between both point estimates are overall small to moderate in size, and we do not find any significant differences even on the 10% significance level. We will explore potential reasons for this finding in the next subsection.

5.2.2 Explaining the Weak Role of Personality

Why does the explicit inclusion of personality have no significant impact on the estimation of the programme effects? One possible explanation is that personality is already implicitly reflected to a large extent by other covariates which have been affected by personality themselves. Considering the strong role of personality in human capital decisions (e.g., Coleman and Deleire, 2003; Almlund *et al.*, 2011) and for (previous) labour market performance (e.g., Heckman *et al.*, 2006; Heineck and Anger, 2010; Judge and Bono, 2001), we split our set of standard covariates X into two groups: the first group is assumed to be unaffected by personality U, while the second group, denoted as W, contains covariates that are potentially affected by personality, in particular variables on human capital and labour market history. Thus, we expect that controlling for these potentially affected variables W in the standard specification already removes part of the selection bias due to personality measures without explicitly accounting for them. We therefore examine the role of W in more detail. For one, we rerun the propensity score matching procedure, this time considering an auxiliary specification that excludes those variables potentially affected by personality (X excluding W) and compare the results to our standard specification (X) and our extended specification ($X + U$). Figure 2a illustrates for men that the major part of the selection bias is removed when covariates on human capital and labour market history are added to the matching specification (standard specification, $P(X)$) compared to a specification where variables potentially affected by personality are excluded (auxiliary specification $P(X \setminus W)$). The bias accounted for by the additional explicit inclusion of usually unobserved personality measures U is very small in comparison (extended specification, $P(X + U)$).

[Insert Table 7 about here]

Moreover, we explore in Table 7 how the balancing of the personality variables is impacted by the inclusion of these different sets of covariates in the propensity score specifications, in

particular the special role of W. In the upper part of each Panel, we report balancing indicators of the separate covariate blocks that jointly constitute our extended matching specification. In addition, the lower parts contain balancing measures of the single personality variables where higher values of the reported standardised biases indicate a worse balancing. We compare the balancing for the unmatched sample (column 1) as well as for the matched sample based on the auxiliary $P(X \setminus W)$, standard $P(X)$, and extended specification $P(X + U)$.

As can be seen in the upper part of Panel A, the inclusion of variables potentially affected by personality W in the matching procedure improves the balancing of the personality variables for men from a mean standardised bias of 10.7 (column 2) to under 9.8 (column 3) without explicitly controlling for the personality measures. When we look at the evidence on the single personality variables, the impact of additionally controlling for W is much stronger for most of the personality measures. For instance, the standardised bias of locus of control changes from 27.7 in the unmatched sample only slightly to 25.2 for the auxiliary specification, while adding human capital and labor market history to the specification reduces it to 18.2 without explicitly accounting for personality. It has to be noted, however, that the balancing of openness actually worsens once W is included, which explains the only moderate impact of W on the mean standardised bias over all personality variables mentioned above. Still, in total, these results support the notion that covariates potentially affected by personality at least partly capture the usually unobserved personality characteristics U, resulting in a small and insignificant impact of personality on the estimated programme effects. For women, however, the evidence is less clear cut. The inclusion of W does not, on average, lead to a better balancing of personality, and the results for the single personality variables are rather mixed (Table 7, Panel B). This might be one explanation for why the differences in ATTs between the standard and extended specification are, on average, higher for women than for men (cf. Table 6 and Figure 2). A possible reason for this observation might be found in the lower labour market attachment of women, which results in personality being less captured by these variables compared to men.

5.3 Effect Heterogeneity with Respect to Personality Characteristics

In the final part of our analysis, we investigate effect heterogeneity to address the question of which participant "personality type" benefits most from the programme. Therefore, for each of

the big five factors – conscientiousness, extraversion, agreeableness, neuroticism, and openness – as well as locus of control and readiness to take risks, we construct a dummy indicator that reflects a high degree (i.e., higher than the median in the gender-specific participant group) in this characteristic. We then conduct post-matching weighted regressions of our outcome variables on a constant, the treatment indicator, these personality dummies, and the interactions between treatment indicator and personality dummies using weighted least squares and the matching weights obtained for the extended specification:

$$\begin{aligned} Y = \lambda_0 + \lambda_1 conscient.^{high} + \ldots + \lambda_7 risk^{high} \\ + \delta_0 D + \delta_1 (D \times conscient.^{high}) + \ldots + \delta_7 (D \times risk^{high}) + v. \end{aligned} \tag{4}$$

The coefficients $\delta_1 \ldots \delta_7$ of the interaction terms between the treatment dummy and each personality dummy then indicate the average difference in the ATT between individuals with a high and a low degree in this personality characteristic while holding all other personality variables constant. We choose this strategy for two reasons: (i) The alternative, where we would split the sample into subsamples based on the personality dummies and conduct the full matching procedure on these subsamples, would lead to very small sample sizes, resulting in unsatisfying common support, poorer matching quality, and results of limited validity. (ii) We observe moderately sized significant correlations between various personality variables. The joint inclusion of all personality variables in the post-matching regressions allows us to estimate the effect heterogeneity with regard to one particular personality variable while holding all other personality variables constant. Therefore, the differences in ATTs are not confounded by these correlations.

[Insert Figure 3 about here]

The results for the cumulated employment effect are graphically displayed in Figure 3. Overall, there is only limited interaction between the programme effects and personality. For men (upper bars), we find a significantly higher effectiveness for more open (compared to less open) and for less risk loving (compared to more risk loving) individuals. For the other big five factors and for locus of control, the differences in ATTs between males with a lower and a higher degree are limited in size and insignificant. These findings hold quite consistently for the other outcome variables as well (full estimation results are reported in Table A.2 in the Appendix). For women

(lower bars), the picture remains rather mixed as we find no significant interactions with the cumulated employment effect. Over all outcomes we only find two significances with respect to openness and readiness to take risk, yet in the reverse direction compared to men.

In total, the empirical evidence on the effect heterogeneity thus suggests only limited interaction between the effectiveness of the subsidy and the personality variables. Given that the theoretical expectations were entirely ambiguous, these results might express that the two opposing effects of personality – on the one hand, on labour market reintegration in the case of participation and, on the other hand, on exit from unemployment in the counterfactual case of non-participation – are generally similar in size. We also emphasise that significance levels might be improved with larger sample sizes; results should thus be interpreted with caution.

5.4 Robustness Analysis

The practical use of propensity score matching requires a series of choices in the implementation that can affect the estimation results. We therefore test the sensitivity of our main effects with respect to different issues. We impose three alternative common support restrictions, alter the choice of the matching algorithm and examine the robustness of the ATT if we account for time-invariant unobserved heterogeneity by estimating two conditional difference-in-differences approaches. In total, the results of the robustness checks are very similar to our main analysis (for details, see Section B.3 in the Supplementary Appendix). Moreover, for the comparison between the standard and the extended specification, we conduct a sensitivity analysis on a working age subsample (30 to 60 year olds) because for those individuals the evidence supporting no systematic changes in personality variables is strongest (Cobb-Clark and Schurer, 2013). We find a similar pattern with regard to the differences in ATTs between the two specifications. They are generally small to moderate in size and highly insignificant (see Table B.1.5 in the Supplementary Appendix for details), confirming the weak impact of personality variables on the estimated ATT. On top of that, a sensitivity analysis with respect to potential classical measurement error in the personality variables reveals that our results are very robust (for details, see Section B.4 in the Supplementary Appendix).

6 Conclusion

The recent entrepreneurship literature emphasises the key role of an individual's personality on the start-up decision and future business success. Whilst this finding is well established for many different countries and settings, it has not yet been transferred into the literature on the evaluation of start-up subsidies and other business support programmes mainly due to data limitations. In our paper, we provide the first evidence on the long-term effectiveness of a new start-up subsidy for unemployed individuals in Germany and incorporate the growing evidence on the important role of personality traits in our evaluation approach. The data at hand – a combination of administrative and survey data – gives us the unique opportunity to study the specific role of personality characteristics when evaluating start-up subsidies. This is of high relevance as it contributes to the ongoing debate about the reliability of earlier evaluation results which have been estimated under the CIA without taking personality into account.

To this end, we implement a propensity score matching approach and control not only for covariates that have been used in earlier studies – such as socio-demographics, human capital, and labour market history – but also for different dimensions of an individual's personality such as the big five, locus of control, and risk preferences. We find that the new subsidy programme has strong positive effects on employment probabilities and income for both men and women over the whole 40 month observation window. To answer our central question as to whether the inclusion of personality variables, in addition to the other control variables, changes the estimated treatment effects significantly, we rerun the analysis excluding the personality traits and compare results. We find only small and insignificant differences in the estimated treatment effects between the two specifications. One possible explanation is that personality is already implicitly reflected to a large extent by other covariates which have been affected by personality themselves. We find evidence supporting this notion, with particular emphasis on the important role of human capital attainment and labour market history. In this sense, our empirical evidence is in line with findings by Lechner and Wunsch (2013), who stress the importance of detailed employment histories for the validity of propensity score matching estimators in evaluating traditional ALMPs. Additionally, we complement similar results for the evaluation of traditional ALMP instruments like short-/long-term training and wage subsidies (Caliendo *et al.*, 2014b).

Furthermore, we also consider effect heterogeneity with respect to personality. From a theoretical point of view, there is no clear prediction of which "personality type" is likely to benefit most from programme participation. We provide the first evidence on this issue and find a limited interaction between the effectiveness of the subsidy and openness and risk attitudes for men. The overall evidence is rather mixed and does not allow us to draw major conclusions about effect heterogeneity with respect to personality; further research is needed, ideally with larger samples.

From a policy perspective, our results have the following two important implications: First, the new start-up subsidy programme is an effective tool to persistently reintegrate formerly unemployed individuals into the labour market, a similar finding to earlier versions of start-up subsidy programmes in Germany. While it should be noted that general equilibrium effects, such as substitution or crowding out, cannot be taken into account within the micro-setting of this study, the positive results on the individual level contribute to the overall promising evidence on the benefits of start-up subsidies for unemployed individuals. Second, existing concerns about the potential overestimation of programme effects in earlier evaluation studies of start-up subsidy programmes, because of missing information on individuals' personalities, might be less justified as long as the set of observed control variables is rich enough. It should be clear that our findings are restricted to the personality variables available in our data (and some of them are measured more parsimoniously than in other surveys), and we do not claim that these measures reflect all factors that were unobserved in earlier studies. Still, given the overwhelming evidence from the entrepreneurship literature stressing the important role of personality in start-up decisions and business success, we are confident that they represent a major component of what usually remains unobserved. However, it should be noted that we have to be cautious in generalising these findings to other programmes or institutional settings. The inclusion of personality measures might be of greater importance for certain subgroups that have lower labour market attachment (like women) or in situations with simply not rich enough data on human capital attainment or employment histories available yet (like younger individuals). Here, personality is probably insufficiently captured by other control variables, and thus an explicit inclusion of personality might be necessary.

References

Almlund, M., Duckworth, A., Heckman, J. and Kautz, T. (2011). Personality psychology and economics. In E. Hanushek, S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education*, vol. 4, Amsterdam: Elsevier, pp. 1–181.

Audretsch, D. B. and Thurik, R. (2001). *Linking Entrepreneurship to Growth*. OECD Science, Technology and Industry Working Papers 2001/2, OECD Publishing, Paris.

Baron, R. A. and Markman, G. D. (2005). Toward a process view of entrepreneurship: The changing relevance of individual-level variables across phases of new firm development. In M. A. Rahim, R. T. Golembiewski and K. D. Mackenzie (eds.), *Current topics in management*, vol. 9, New Brunswick, NJ: Transaction, pp. 45–64.

Battistin, E. and Chesher, A. (2014). Treatment effect estimation with covariate measurement error. *Journal of Econometrics*, **178** (2), 707–715.

Baum, J. R. and Locke, E. A. (2004). The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth. *Journal of Applied Psychology*, **89** (4), 587–598.

Bianchi, M. and Bobba, M. (2013). Liquidity, risk, and occupational choices. *Review of Economics Studies*, **80** (2), 491–511.

Brandstätter, H. (2011). Personality aspects of entrepreneurship: A look at five meta-analyses. *Personality and Individual Differences*, **51** (3), 222–230.

Bygrave, W. D. and Hofer, C. W. (1991). Theorizing about entrepreneurship. *Entrepreneurship: Theory & Practice*, **16** (2), 13–22.

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

—, Fossen, F. and Kritikos, A. (2009). Risk Attitudes of Nascent Entrepreneurs: New Evidence from an Experimentally-Validated Survey. *Small Business Economics*, **32** (2), 153–167.

—, — and — (2010a). The impact of risk attitudes on entrepreneurial survival. *Journal of Economic Behavior and Organization*, **76** (1), 45–63.

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

—, —, — and Wetter, M. (2015b). The gender gap in entrepreneurship: Not just a matter of personality. *CESifo Economic Studies*, **61** (1), 202–238.

— and Hogenacker, J. (2012). The German Labor Market after the Great Recession: Successful Reforms and Future Challenges. *IZA Journal of European Labor Studies*, **1:3**.

—, —, Künn, S. and Wiessner, F. (2012). Alte Idee, neues Programm: Der Gründungszuschuss als Nachfolger von Überbrückungsgeld und Ich-AG. *Journal for Labour Market Research*, **45(2)**, 99–123.

—, —, — and — (2015c). 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 Kritikos, A. (2010). Start-Ups by the Unemployed: Characteristics, Survival and Direct Employment Effects. *Small Business Economics*, **35(1)**, 71–92,

— 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 — (2014). Regional effect heterogeneity of start-up subsidies for the unemployed. *Regional Studies*, **48** (6), 1108–1134.

— and — (2015). Getting back into the labor market: The effects of start-up subsidies for unemployed females. *Journal of Population Economics*, **28** (4), 1005–1043.

—, — and Wiessner, F. (2010b). Die Nachhaltigkeit von geförderten Existenzgründungen aus Arbeitslosigkeit: Eine Bilanz nach fünf Jahren. *Journal for Labour Market Research / Zeitschrift für ArbeitsmarktForschung*, **42** (4), 269–291.

—, Mahlstedt, R. and Mitnik, O. A. (2014b). *Unobservable, but Unimportant? The Influence of Personality Traits (and Other Usually Unobserved Variables) for the Estimation of Treatment Effects*. IZA Discussion Paper No. 8337, IZA, Bonn.

Cameron, C. and Trivedi, P. K. (2009). *Microeconometrics Using Stata*. College Station, Texas: Stata Press.

Chell, E., Harworth, J. and Brearley, S. (1991). The search for entrepreneurial traits. In E. Chell, J. Harworth and S. Brearley (eds.), *The Entrepreneurial Personality: Concepts, Cases and Categories*, Routledge Small Business Series, London: Thomson Learning, pp. 29–53.

Ciavarella, M. A., Buchholtz, A. K., Riordan, C. M., Gatewood, R. D. and Stokes, G. S. (2004). The big five and venture survival: Is there a linkage? *Journal of Business Venturing*, **19** (4), 465–483.

Cobb-Clark, D. A. 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), F358–F400.

Coleman, M. and Deleire, T. (2003). An economic model of locus of control and the human capital investment decision. *Journal of Human Resources*, **38** (3), 701–721.

Costa, P. and McCrae, R. (1992). *Revised Neo Personality Inventory (NEO-PI-R) and NEO Five Factor Inventory (NEO-FFI)*. Psychological Assessment Resources (Odessa).

Cramer, J., Hartog, J., Jonker, N. and Van Praag, C. (2002). Low Risk Aversion Encourages the Choice for Entrepreneurship: An Empirical Test of a Truism. *Journal of Economic Behavior and Organization*, **48** (1), 29–36.

Crump, R., Hotz, V. J., Imbens, G. W. and Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, **96** (1), 187–199.

Dehne, M. and Schupp, J. (2007). *Persönlichkeitsmerkmale im Sozio-oekonomischen Panel (SOEP) - Konzept, Umsetzung und empirische Eigenschaften*. Research Notes 26, DIW Berlin.

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

Diaz-Serrano, L. and O'Neill, D. (2004). *The Relationship between Unemployment and Risk-Aversion*. IZA Discussion Paper No. 1214, IZA, Bonn.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J. and Wagner, G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association*, **9** (3), 522–550.

Ekelund, J., Johannson, E., Järvelin, M. and Lichtermann, D. (2005). Self-employment and risk aversion - evidence from psychological test data. *Labour Economics*, **12** (5), 649–659.

Evans, D. and Leigthon, L. (1989). Some empirical aspects of entrepreneurship. *American Economic Review*, **79** (3), 519–535.

Fairlie, R., Karlan, D. and Zinman, J. (2015). Behind the GATE experiment: Evidence on effects of and rational for subsidized entrepreneurship training. *American Economic Journal: Economic Policy*, **7** (2), 125–161.

Federal Statistical Office (2012). *Verdienste und Arbeitskosten, Arbeitnehmerverdienste 3. Vierteljahr 2012, Fachserie 16, Reihe 2.1.* Wiesbaden.

Federal Statistical Office (2013). *Verdienste und Arbeitskosten, Nettoverdienste, Modellrechnung 2012, Fachserie 16, Reihe 2.5.* Wiesbaden.

Feinberg, R. M. (1977). Risk aversion, risk, and the duration of unemployment. *Review of Economics & Statistics*, **59** (3), 264–271.

Fritsch, M. (2008). How does new business development affect regional development? Introduction to the special issues. *Small Business Economics*, **30** (1), 1–14.

—, Kritikos, A. and Rusakova, A. (2012). *Who Starts a Business and Who is Self-Employed in Germany.* DIW Discussion Papers No. 1184, DIW, Berlin.

Georgellis, Y. and Wall, H. J. (2005). Gender differences in self-employment. *International Review of Applied Economics*, **19** (3), 321–342.

Hamilton, B. H., Papageorge, N. W. and Pande, N. (2015). *The Right Stuff? Personality and Entrepreneurship.* Working paper, Johns Hopkins University.

Hansemark, O. C. (2003). Need for achievement, locus of control and the prediction of business start-ups: A longitudinal study. *Journal of Economic Psychology*, **24** (3), 301–319.

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** (3), 411–482.

Heineck, G. and Anger, S. (2010). The returns to cognitive abilities and personality traits in Germany. *Labour Economics*, **17** (3), 535 – 546.

Heunks, F. J. (1998). Innovation, creativity and success. *Small Business Economics*, **10** (3), 263–272.

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.

John, O., Donahue, E. and Kentle, R. (1991). The big five inventory - versions 4a and 54. University of California, Berkeley, Institute of Personality and Social Research.

Judge, A. T. and Bono, J. E. (2001). Relationship of core self-evaluations traits -self-esteem, generalized self-efficacy, locus of control, and emotional stability- with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology*, **86** (1), 80–92.

Judge, T. A., Erez, A., Bono, J. E. and Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology*, **83** (3), 693–710.

Koellinger, P. D. and Thurik, R. (2012). Entrepreneurship and the business cycle. *The Review of Economics and Statistics*, **94** (4), 1143–1156.

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

McClelland, D. C. (1965). Need achievement and entrepreneurship: A longitudinal study. *Journal of Personality and Social Psychology*, **1** (4), 389–392.

McCrae, R. and Costa, P. (2008). The five-factor theory of personality. In O. P. John, R. Robins and L. Pervin (eds.), *Handbook of Personality: Theory and Research*, New York: Guilford, pp. 159–181.

McGee, A. D. (2015). How the perception of control influences unemployed job search. *Industrial and Labor Relations Review*, **68** (1), 184–211.

Meager, N. (1996). From Unemployment 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.

Niefert, M. (2010). Characteristics and determinants of start-ups from unemployment: Evidence from German micro data. *Journal of Small Business and Entrepreneurship*, **23** (3), 409–429.

Nolte, H., Weischer, C., Wilkesmann, U., Maetzel, J. and Tegethoff, H. G. (1997). *Kontrolleinstellungen zum Leben und der Zukunft - Auswertung eines neuen sozialpsychologischen Itemblocks im Sozio-oekonomischen Panel*. Diskussionspapier, Fakultät für Sozialwissenschaft Ruhr-Universität Bochum.

Oberschachtsiek, D. and Ullrich, B. (2010). *The link between career risk aversion and unemployment duration: evidence of non-linear and time-depending pattern*. Working Paper Series in Economics No. 189, University of Lüneburg.

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.

—, Kolodziejczyk, P. and Lázár, G. (1998). The net impact of active labour programmes in Hungary and Poland. *International Labour Review*, **137** (3), 321–346.

Pannenberg, M. (2010). Risk attitudes and reservation wages of unemployed workers: Evidence from panel data. *Economic Letters*, **106** (3), 223–226.

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.

Pissarides, C. A. (1974). Risk, job search, and income distribution. *Journal of Political Economy*, **82** (6), 1255–1267.

Rauch, A. and Frese, M. (2000). Psychological approaches to entrepreneurial success: A general model and an overview of findings. In C. Cooper and I. Robertson (eds.), *International Review of Industrial and Organizational Psychology*, vol. 15, Wiley, pp. 101–142.

— and — (2007). Let's Put the Person Back into Entrepreneurship Research: A Meta-Analysis on the Relationship Between Business Owners' Personality Traits, Business Creation, and Success. *European Journal of Work and Organizational Psychology*, **16** (4), 353–385.

Rodriguez-Planas, N. (2010). Channels through which public employment services and small-business assistance programs work. *Oxford Bulletin of Economics and Statistics*, **72** (4), 458–485.

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

Rosenbaum, P. and Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, **70** (1), 41–55.

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

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

Rotger, G. P., Gortz, M. and Storey, D. J. (2012). Assessing the effectiveness of guided preparation for new venture creation and performance: Theory and practice. *Journal of Business Venturing*, **27** (4), 506–521.

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.

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.

Skriabikova, O. J., Dohmen, T. and Kriechel, B. (2014). New evidence on the relationship between risk attitudes and self-employment. *Labour Economics*, **30**, 176–184.

Stewart, W. and Roth, P. (2001). Risk Propensity Differences Between Entrepreneurs and Managers: A Meta-Analytic Review. *Journal of Applied Psychology*, **86** (1), 145–153.

Storey, D. (1994). *Understanding the Small Business Sector*. London: Routledge.

— (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., Haapanen, M. and Ritsilä, J. (2008). Evaluation of investment subsidies: When is deadweight zero? *International Review*, **22** (5), 585–600.

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

Wren, C. and Storey, D. J. (2002). Evaluating the effect of soft business support upon firm performance. *Oxford Economic Papers*, **54** (2), 334 – 365.

Zhao, H., Seibert, S. E. and Lumpkin, G. (2010). The relationship of personality to entrepreneurial intentions and performance: A meta-analytic review. *Journal of Management*, **36** (2), 381–404.

Tables and Figures

Table 1: The role of personality

	Decision to start a business (1)	Survival as entrepreneur (2)	Exit from unemployment (3)	Interaction with programme effect (4)
Big five				
Conscientiousness	$^{+}/_{-}$	+	+	$^{+}/_{-}$
Extraversion	+	+	$^{+}/_{-}$	$^{+}/_{-}$
Agreeableness	$^{+}/_{-}$	$^{+}/_{-}$	$^{+}/_{-}$	$^{+}/_{-}$
Neuroticism	−	−	−	$^{+}/_{-}$
Openness	+	$^{+}/_{-}$	+	$^{+}/_{-}$
Locus of control	+	+	$^{+}/_{-}$	$^{+}/_{-}$
Readiness to take risks	+	inverse u	−	$^{+}/_{-}$

Note: The table summarises the hypotheses about the direction of the effects of personality characteristics on start-up (1), business survival (2), exit from unemployment (3), and the interaction with programme effectiveness (4) as discussed in Section 2.2.
+ denotes a positive effect, − denotes a negative effect, and $^{+}/_{-}$ denotes no or an ambiguous effect.

Table 2: Definition of the estimation sample

	Participants (1)	Non-participants (2)
Respondents in first interview	2,306	2,307
Random subsample (50%)	1,143	1,390
Respondents in second interview	632	789
Consent to link administrative data	617	776
Estimation sample	589	699
Men	367	439
Women	222	260

Note: Number of observations. The first interviews were conducted in November and December 2010, the second interviews in August through October 2012.

Table 3: Selected descriptive statistics

	Men			Women		
	Particip.	Non-part.	p-val.	Particip.	Non-part.	p-val.
	(1)	(2)	(3)	(4)	(5)	(6)
A. Selected individual characteristics[a]						
Age at start-up (years)	40.92	41.02	0.89	41.05	40.42	0.47
Married	0.64	0.56	0.02	0.55	0.55	0.94
East Germany	0.20	0.24	0.18	0.26	0.27	0.76
Upper secondary school	0.52	0.54	0.45	0.55	0.60	0.35
Lifetime unemployment (share)[b]	0.04	0.08	0.00	0.05	0.08	0.00
Dependent employed before unempl.	0.63	0.51	0.00	0.51	0.46	0.26
Monthly unemployment benefit (euros)	1,093	1,080	0.77	803	793	0.80
B. Personality measures[c]						
Big five[d]						
Conscientiousness	5.95	5.89	0.36	6.19	6.18	0.88
	(0.82)	(0.88)		(0.76)	(0.78)	
Extraversion	5.63	5.47	0.04	6.07	5.79	0.00
	(1.08)	(1.12)		(1.00)	(1.05)	
Agreeableness	5.93	5.97	0.59	6.34	6.28	0.41
	(0.96)	(1.08)		(0.76)	(0.78)	
Neuroticism	3.83	3.88	0.56	4.27	4.76	0.00
	(1.35)	(1.30)		(1.41)	(1.24)	
Openness	4.86	4.69	0.07	5.34	5.06	0.01
	(1.36)	(1.33)		(1.27)	(1.27)	
Locus of control[d]	5.48	5.25	0.00	5.43	4.99	0.00
	(0.80)	(0.86)		(0.83)	(0.87)	
Readiness to take risks[e]	6.33	6.06	0.05	5.82	5.70	0.51
	(1.87)	(2.01)		(2.10)	(1.97)	
C. Short-term labour market outcomes (21 months after start-up)						
Self-employed	0.853	0.114	0.00	0.797	0.096	0.00
Self- or regular employed	0.943	0.731	0.00	0.901	0.673	0.00
Unemployed or in ALMP	0.052	0.257	0.00	0.054	0.188	0.00
Net earned income (Euro/month)[f]	2,332	1,381	0.00	1,279	853	0.00
	(2,158)	(1,672)		(1,200)	(865)	
	[2,000]	[1,200]		[1,000]	[750]	
D. Long-term labour market outcomes (40 months after start-up)						
Self-employed	0.774	0.128	0.00	0.689	0.096	0.00
Self- or regular employed	0.929	0.786	0.00	0.865	0.692	0.00
Unemployed or in ALMP	0.033	0.123	0.00	0.059	0.081	0.34
Net earned income (Euro/month)[f]	2,813	1,836	0.00	1,611	978	0.00
	(2,397)	(2,125)		(2,257)	(919)	
	[2,500]	[1,500]		[1,054]	[900]	
Number of observations	367	439		222	260	

Note: Reported are sample averages and p-values for t-tests of equal means. Standard deviations are denoted in parentheses, medians in brackets.

[a] The full list of individual characteristics used in the subsequent propensity score matching estimations can be found in Table B.1.1 in the Supplementary Appendix.

[b] Shares are calculated by dividing the cumulative time spent in unemployment in the past by the total time spent in the labour market (as approximated by age-15).

[c] For details on the construction of the personality variables, see Table B.1.2 in the Supplementary Appendix.

[d] The big five and locus of control are measured on a scale from 1 to 7, where higher values indicate a stronger degree of the respective trait or a more internal locus of control.

[e] Risk is measured on a scale from 0 to 10, where higher values indicate a higher willingness to take risk.

[f] Income measures are based on slightly lower numbers of observations due to item non-responses.

Table 4: Descriptive statistics with respect to business outcomes

	Men		Women	
	Outcome variable		Outcome variable	
	21 mo. after start-up	40 mo. after start-up	21 mo. after start-up	40 mo. after start-up
	(1)	(2)	(3)	(4)
Net earned income (euros/month)[a]	2,609	3,189	1,488	1,988
	(2,228)	(2,477)	(1,326)	(2,601)
	[2,000]	[2,500]	[1,300]	[1,450]
Working time (hours/week)[a]	51.0	50.5	42.0	40.8
	(13.5)	(14.4)	(16.9)	(18.0)
	[50.0]	[50.0]	[40.0]	[40.0]
Net earned hourly income (euros)[a]	12.52	15.03	9.01	13.17
	(10.36)	(10.96)	(7.36)	(15.61)
	[10.62]	[12.44]	[7.34]	[9.22]
At least one employee	0.391	0.437	0.307	0.353
Number of full-time equivalent employees (if > 0)[b]	3.6	3.6	2.5	2.3
	(10.1)	(5.2)	(4.5)	(4.7)
	[1.3]	[2.0]	[1.3]	[1.0]
Filed patent application	0.018	0.018	0.007	0.013
Filed application to protect corporate ID	0.074	0.102	0.039	0.059
Number of observations	284	284	153	153

Note: Reported are sample averages for all participants self-employed 40 months after start-up. Standard deviations are denoted in parentheses, medians in brackets.

[a] Income and working time measures are based on slightly lower numbers of observations due to item non-responses.

[b] Full-time equivalent employees are calculated as the weighted sum of full-time employees (weight 1), part-time employees (weight 0.5), and other employees (weight 0.25). Apprentices are not considered in the calculations.

Table 5: Matching quality indicators

	Men		Women	
	Before matching (1)	After matching (2)	Before matching (3)	After matching (4)
A. Number of variables with significant differences in means[a]				
at 1%-level	10	0	6	0
at 5%-level	18	0	9	0
at 10%-level	25	2	12	0
B. Number of variables with absolute standardised bias[b]				
$< 1\%$	5	12	7	11
1% until $< 3\%$	8	24	14	26
3% until $< 5\%$	9	18	6	16
5% until $< 10\%$	22	14	21	19
10% until $< 15\%$	14	4	12	0
$\geq 15\%$	14	0	12	0
Mean absolute standardised bias in %	10.02	3.66	9.22	3.57
Median absolute standardised bias in %	6.88	2.97	6.57	2.93
C. (Re)Estimation of the propensity score[c]				
Pseudo-R^2	0.2202	0.0305	0.2027	0.0309
p-value of joint significance test	0.0000	1.0000	0.0000	1.0000
Total number of variables	72	72	72	72
Participants off support		2		15

Note: Reported are indicators for covariate balancing before and after matching using a bandwidth of 0.24 for the extended specification (cf. Table A.1 in the Appendix). The matching quality indicators for all optimal bandwidths from Table B.1.3 are reported in the Supplementary Appendix, Table B.1.4.

[a] Equality of means is tested based on t-tests.

[b] The standardised bias is calculated as the difference of sample means for participants (P) and (matched) non-participants (NP) as a percentage of the square root of the average of sample variances in both groups $SB = 100 \cdot (\overline{X}_P - \overline{X}_{NP})/(0.5 \cdot V_P(X) + 0.5 \cdot V_{NP}(X))^{0.5}$ following Rosenbaum and Rubin (1985).

[c] The pseudo-R^2 and the p-value of joint significance test stem from a probit (re)estimation of the propensity score on the (un)matched sample (Sianesi, 2004).

Table 6: Matching estimation results

	Propensity score specification			Comparison	
	Auxiliary $P(X \setminus W)$ (1)	Standard $P(X)$ (2)	Extended $P(X+U)$ (3)	Auxiliary vs. extended (4)	Standard vs. extended (5)
A. Treatment effects for men					
Short-term labour market outcomes (21 months after start-up)					
Self- or regular employed	0.2065^{***}	0.1396^{***}	0.1336^{***}	0.0730^{***}	0.0061
	{0.0000}	{0.0000}	{0.0000}	{0.0069}	{0.6245}
Net earned income (euros/mo.)[a]	901.13^{***}	634.77^{***}	600.55^{***}	300.58^{**}	34.22
	{0.0000}	{0.0000}	{0.0010}	{0.0344}	{0.6433}
Long-term labour market outcomes (40 months after start-up)					
Self- or regular employed	0.1373^{***}	0.0921^{***}	0.0800^{***}	0.0573^{**}	0.0122
	{0.0000}	{0.0000}	{0.0010}	{0.0102}	{0.2612}
Cumulated ($\sum_{i=0}^{40}$, months)	10.71^{***}	8.12^{***}	7.76^{***}	2.95^{***}	0.36
	{0.0000}	{0.0000}	{0.0000}	{0.0005}	{0.3499}
Net earned income (euros/mo.)[a]	865.79^{***}	774.85^{***}	736.64^{***}	129.15	38.21
	{0.0000}	{0.0000}	{0.0010}	{0.4208}	{0.6945}
Participants off support	1	2	2		
Number of observations	806	806	806		
B. Treatment effects for women					
Short-term labour market outcomes (21 months after start-up)					
Self- or regular employed	0.2284^{***}	0.1864^{***}	0.1669^{***}	0.0615	0.0195
	{0.0000}	{0.0000}	{0.0000}	{0.1200}	{0.4548}
Net earned income (euros/mo.)[a]	418.85^{***}	343.96^{***}	313.48^{***}	105.37	30.48
	{0.0000}	{0.0010}	{0.0030}	{0.2330}	{0.6396}
Long-term labour market outcomes (40 months after start-up)					
Self- or regular employed	0.1749^{***}	0.1231^{***}	0.1056^{***}	0.0693^{*}	0.0175
	{0.0000}	{0.0010}	{0.0090}	{0.0806}	{0.5134}
Cumulated ($\sum_{i=0}^{40}$, months)	11.56^{***}	9.44^{***}	8.56^{***}	3.00^{**}	0.88
	{0.0000}	{0.0000}	{0.0000}	{0.0250}	{0.3258}
Net earned income (euros/mo.)[a]	632.85^{***}	624.88^{***}	608.57^{***}	24.28	16.31
	{0.0400}	{0.0040}	{0.0040}	{0.7876}	{0.8063}
Participants off support	1	11	15		
Number of observations	482	482	482		
C. Propensity score specification					
Socio-demographics	✓	✓	✓		
Intergenerational information	✓	✓	✓		
Regional labour market	✓	✓	✓		
Human capital		✓	✓		
Labour market history		✓	✓		
Personality			✓		

Note: Reported are average treatment effects on the treated (ATT) as the difference in mean outcomes between participants and matched non-participants using Epanechnikov kernel propensity score matching with optimal bandwidth based on leave-one-out cross-validation for the indicated specification. Following Huber *et al.* (2014), p-values are bootstrapped and based on 999 replications. p-values for the differences in ATTs are based on bootstrapped robust Hausman tests with 999 replications (see Cameron and Trivedi, 2009, for details). All p-values are reported in braces. ***/**/* indicate significance at the 1/5/10% level.

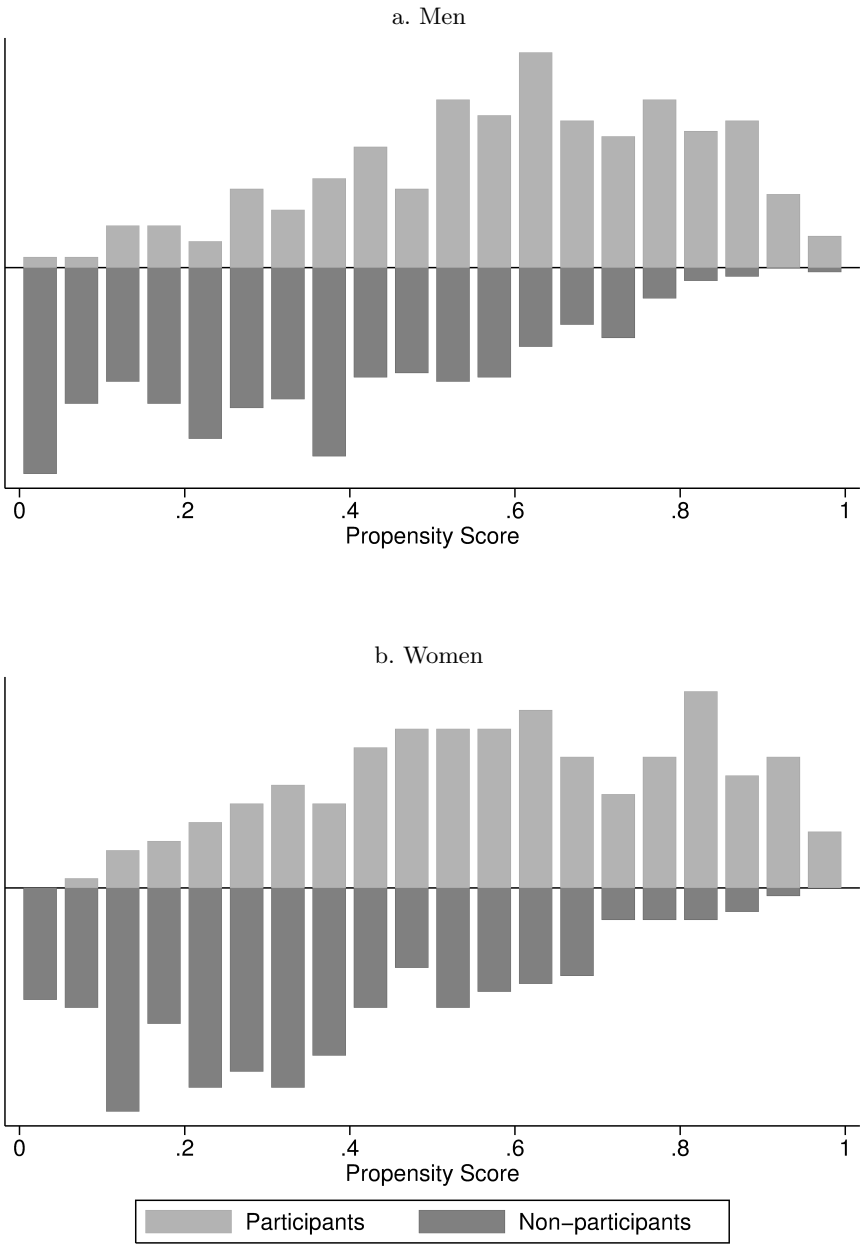
[a] Income measures are based on slightly lower number of observations due to item non-responses.

Table 7: Balancing of personality and control variables

	Descriptive raw difference (1)	Propensity score specification: Auxiliary $P(X \setminus W)$ (2)	Standard $P(X)$ (3)	Extended $P(X+U)$ (4)
A. Men				
Mean absolute standardised bias in %				
Unaffected covariates ($X \setminus W$)				
Socio-demographics	9.69	5.72	5.46	5.47
Intergenerational information	8.45	5.19	3.52	5.03
Regional labour market	4.46	2.65	2.96	2.75
Potentially affected variables (W)				
Human capital	11.52	13.21	4.05	4.88
Labour market history	12.44	10.98	3.21	2.91
Usually unobserved variables (U)				
Personality variables	11.75	10.72	9.77	3.14
Standardised bias in %				
Big five				
Conscientiousness	6.56	6.18	4.73	4.54
Extraversion	14.74	14.17	9.71	2.31
Agreeableness	-3.80	-0.98	1.12	0.28
Neuroticism	-4.11	-1.74	3.35	6.03
Openness	12.94	12.85	19.63	3.60
Locus of control	27.71	25.16	18.19	4.26
Readiness to take risks	14.20	14.97	13.18	-0.39
Read. to take risks squared	-9.92	-9.67	-8.24	3.74
B. Women				
Mean absolute standardised bias in %				
Unaffected covariates ($X \setminus W$)				
Socio-demographics	4.04	3.30	2.07	2.53
Intergenerational information	4.88	2.66	1.60	4.12
Regional labour market	5.52	4.24	1.60	3.76
Potentially affected variables (W)				
Human capital	9.41	9.84	3.55	3.48
Labour market history	10.41	10.46	2.72	3.41
Usually unobserved variables (U)				
Personality variables	20.15	20.90	21.08	4.79
Standardised bias in %				
Big five				
Conscientiousness	1.36	3.41	3.71	-1.24
Extraversion	26.91	27.68	21.27	6.96
Agreeableness	7.48	7.43	6.43	4.34
Neuroticism	-36.63	-35.39	-33.83	-9.99
Openness	22.36	23.15	27.92	3.90
Locus of control	51.33	52.67	54.42	4.56
Readiness to take risks	6.08	7.45	8.20	1.37
Read. to take risks squared	9.03	10.03	12.86	5.93

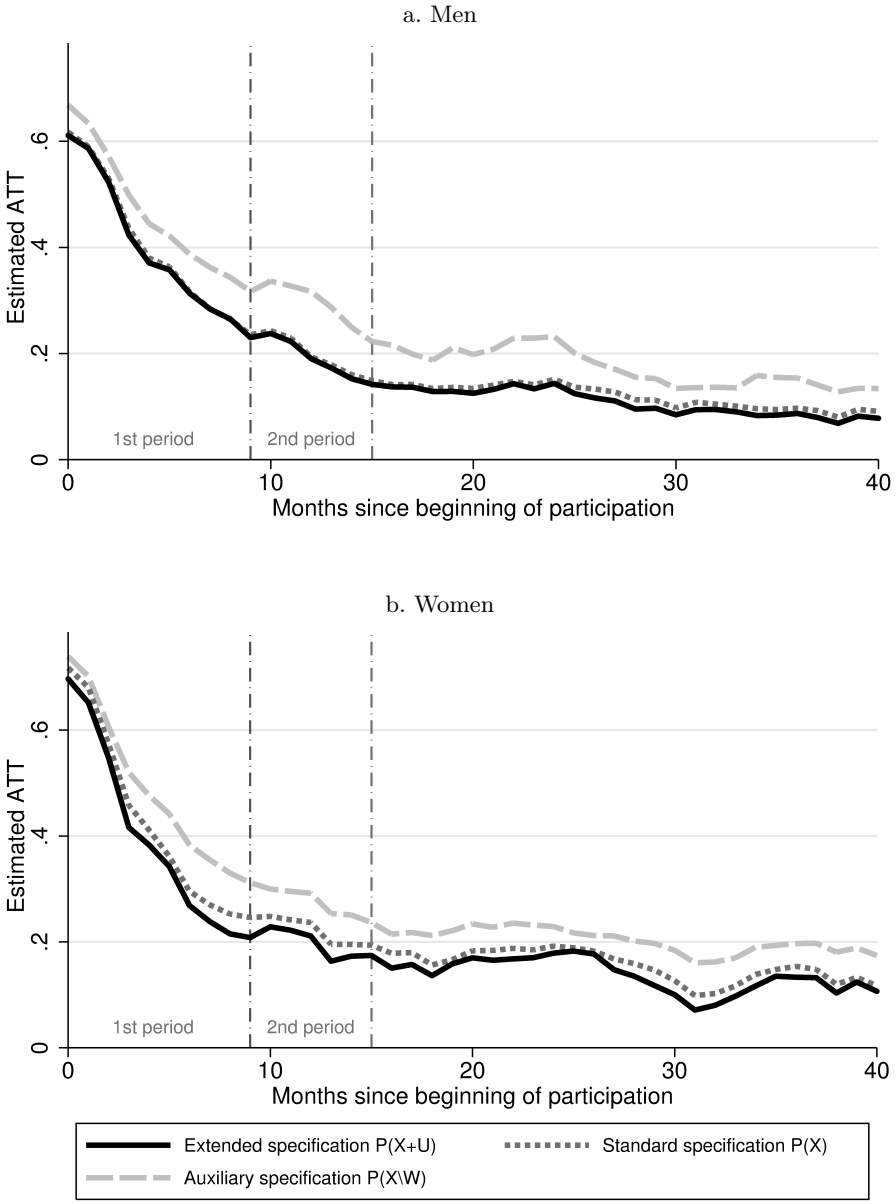
Note: Reported are the mean standardised biases for each covariate block calculated over the absolute standardised biases of all covariates in the block. The standardised bias is the difference of sample means for participants and non-participants as a percentage of the square root of the average of sample variances in both groups (Rosenbaum and Rubin, 1985). For each personality variable, the table reports the individual standardised bias. The numbers are reported for the unmatched sample (descriptive raw difference) and the matched sample using the indicated propensity score specification and a bandwidth of 0.24. The results of the underlying probit estimations are reported in Table A.1 in the Appendix.

Figure 1: Propensity score distributions



Note: Depicted are propensity score distributions for participants and non-participants based on the propensity score probit specification including variables on socio-demographics, intergenerational information, regional labour market, human capital, labour market history, and personality characteristics (extended specification). The results of the underlying probit estimations are reported in Table A.1 in the Appendix.

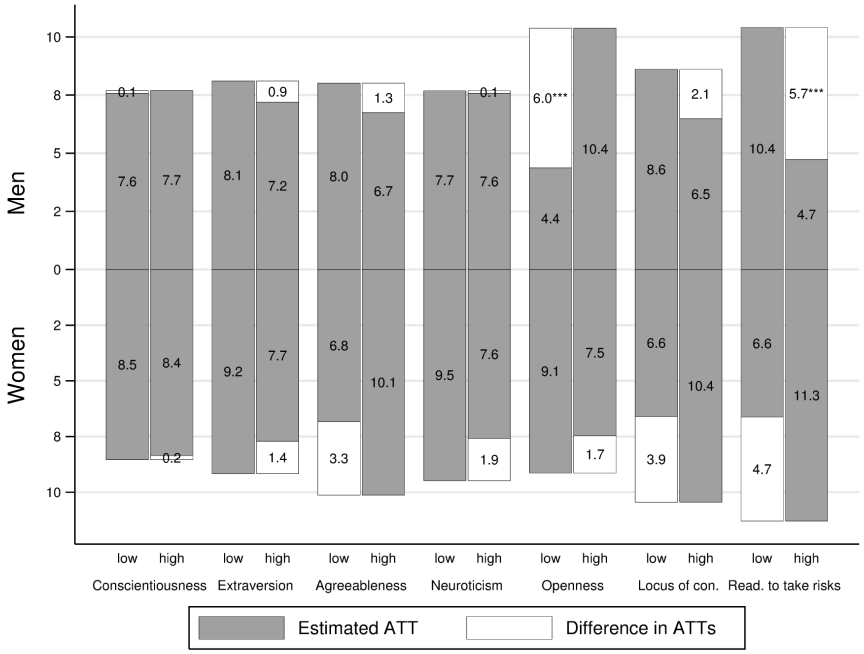
Figure 2: Programme effects over time



Outcome: Self- or regular employment

Note: Depicted are average treatment effects on the treated (ATT) as the difference in mean outcomes between participants and matched non-participants over time using Epanechnikov kernel propensity score matching. The first funding period consisted of 9 months during which an amount equivalent to the individual's last unemployment benefit and a lump sum of 300 euros per month for social security coverage were paid. During the optional second period of 6 months, the subsidy was reduced to the lump sum. The specifications include variables on socio-demographics, intergenerational information, and regional labour market (auxiliary specification, $P(X \setminus W)$), plus variables on human capital and labour market history (standard specification, $P(X)$), and additionally personality characteristics (extended specification, $P(X + U)$).

Figure 3: Treatment effect heterogeneity analysis



Outcome: Cumulated effect in self- or regular employment (in months)

Note: Depicted are average treatment effects on the treated (ATT) for individuals with low and high degrees of the respective personality measure as well as the differences in ATTs between the two groups based on post-matching weighted regressions, separately for men (upper bars) and women (lower bars). ***/**/* indicate significant differences at the 1/5/10% level. For details, see Section 5.3 and Table A.2 in the Appendix.

A Appendix

Table A.1: Propensity score estimation

	Men			Women		
	Propensity score specification			Propensity score specification		
	Auxiliary $P(X \setminus W)$ (1)	Standard $P(X)$ (2)	Extended $P(X+U)$ (3)	Auxiliary $P(X \setminus W)$ (4)	Standard $P(X)$ (5)	Extended $P(X+U)$ (6)
Socio-demographics						
Age at start-up						
(ref.: younger than 30 years)						
30 to less than 40 years	0.196	0.203	0.174	-0.109	-0.244	-0.378
40 to less than 50 years	0.189	0.056	0.021	-0.029	-0.275	-0.425*
50 years and older	-0.034	-0.062	-0.025	-0.010	-0.220	-0.291
East Germany	-0.557***	-0.712***	-0.687***	-0.094	-0.089	-0.394
German citizen	-0.052	-0.324	-0.368	0.005	0.019	-0.056
Health restrictions	-0.151	0.025	0.072	-0.029	0.077	0.231
Married	0.216**	0.028	0.054	-0.042	-0.036	-0.060
Children						
(ref.: no children)						
Children under 10	-0.025	0.106	0.150	0.021	-0.020	0.092
Children 10 years or older	-0.133	0.165	0.189	0.227	0.354*	0.424*
Single parent	0.794*	0.900*	1.012**	-0.089	-0.051	0.044
Intergenerational information						
Highest schooling degree of father						
(ref.: other/no degree)						
Lower secondary school	0.119	0.033	0.012	-0.068	-0.120	-0.119
Middle secondary school	-0.101	-0.181	-0.173	-0.204	-0.149	-0.238
Upper secondary school	0.302*	0.227	0.206	0.055	0.020	0.007
One or both parents born abroad	-0.053	-0.009	-0.023	-0.121	-0.197	-0.244
Father and/or mother is/was self-employed	0.126	0.251**	0.259**	-0.056	0.010	-0.028
Father employed when respondent 15 years old	0.132	0.167	0.170	-0.041	-0.045	-0.018
Regional labour market						
Regional cluster (ref.: type Ia)						
Type Ib	0.134	-0.107	-0.095	-0.158	-0.334	-0.355
Type IIa	-0.036	-0.170	-0.209	0.048	0.219	0.171
Type IIb	0.123	0.360	0.403	0.009	0.054	0.259
Type IIc	0.090	0.235	0.249	-0.232	-0.131	-0.270
Type IIIa	0.012	0.0001	0.028	-0.144	-0.160	-0.216
Type IIIb	-0.022	0.049	0.106	0.100	0.249	0.360
Type IVa	0.158	0.059	0.108	-0.011	-0.220	-0.345
Type IVb	-0.057	-0.115	-0.147	-0.158	-0.108	-0.059
Type IVc	0.204	0.076	0.103	0.680	0.408	0.259
Type Va	0.806**	0.934**	0.978**	-0.292	-0.264	0.077
Type Vb	0.771***	0.982***	0.961***	0.012	-0.164	0.138
Type Vc	0.64*	0.599	0.641*	0.147	0.311	0.553
Human capital						
Highest schooling degree						
(ref.: other/no degree)						
Lower secondary school		6.170	6.621		-0.336	0.057
Middle secondary school		5.849	6.242		-0.529	-0.262
Upper secondary school		5.931	6.337		-0.866	-0.662
Professional education						
(ref.: other/no training)						
Vocational training		-0.132	-0.185		-0.374	-0.403
Professional/vocational academy		-0.158	-0.206		0.284	0.355
Technical college/university degree		-0.557**	-0.572**		0.103	-0.014
Professional qualification						
(ref.: unskilled workers)						
Skilled workers		0.237	0.211		0.149	0.144
Top management		-0.171	-0.245		-0.214	-0.119

(Table continued on next page)

(Table A.1 continued)

	Men			Women		
	Propensity score specification			Propensity score specification		
	Auxiliary $P(X \setminus W)$ (1)	Standard $P(X)$ (2)	Extended $P(X+U)$ (3)	Auxiliary $P(X \setminus W)$ (4)	Standard $P(X)$ (5)	Extended $P(X+U)$ (6)
Labour market history						
Lifetime unemployment (share)[b]		-15.544***	-15.762***		-6.888***	-6.727***
Lifetime unemployment (share[b], squared)		22.843***	23.506***		6.795*	6.152
Second to last year before start-up						
Months employed		0.003	0.007		0.019	0.037*
Months in labour market program		-0.010	-0.001		0.091*	0.106**
Next to last year before start-up						
Months employed		0.004	0.011		-0.028	-0.035
Months in labour market program		0.022	0.019		-0.115**	-0.084
Last year before start-up						
Months employed		-0.033*	-0.035*		-0.016	-0.022
Months in labour market program		0.089*	0.094*		-0.005	-0.017
Employment status before unemployment (ref.: other)						
Dependent employment		0.164	0.163		0.011	-0.004
Disable to work/unemployable		-0.706***	-0.769***		-0.511**	-0.308
Occupational group before unemployment (ref.: other)						
Manufacturing		0.034	0.018		0.818*	1.022**
Technical profession		0.222	0.268		0.126	0.202
Services		0.277	0.279		0.410	0.529
Daily income from last employment (euros)		-0.002	-0.002		-0.004**	-0.004*
Duration of last unemployment spell (ref.: less than 1 month)						
1 to less than 3 months		-0.781***	-0.788***		-0.749**	-0.889**
3 to less than 6 months		-0.514**	-0.549**		-0.736**	-0.929**
6 to less than 12 months		-0.438*	-0.470*		-0.885**	-1.063***
12 months and above		-0.559*	-0.566*		-0.162	-0.333
Monthly unemployment benefit (ref.: less than 300 euros)						
300 to less than 600 euros		0.041	0.166		-0.282	-0.259
600 to less than 900 euros		-0.202	-0.151		-0.170	-0.155
900 to less than 1200 euros		-0.029	0.005		-0.189	-0.206
1200 to less than 1500 euros		-0.122	-0.004		0.217	0.215
1500 euros and above		-0.195	-0.139		0.221	0.073
Remaining unemployment benefit entitlement (ref.: less than 3 months)						
3 to less than 6 months		0.048	0.033		0.448**	0.374*
6 to less than 9 months		-0.210	-0.251		-0.128	-0.137
9 to less than 12 months		-0.026	-0.073		-0.239	-0.287
12 months and above		-0.101	-0.135		-0.085	-0.151
Number of placement offers		0.009	0.009		-0.003	-0.002
Personality traits						
Big five						
Conscientiousness			-0.044			-0.055
Extraversion			-0.001			0.067
Agreeableness			-0.046			0.015
Neuroticism			0.026			-0.117
Openness			0.155***			0.203***
Locus of control			0.142**			0.354***
Readiness to take risks			0.062			-0.073
Readiness to take risks squared			-0.047			0.032
Constant	-0.566*	-4.359	-4.742	0.127	2.371**	2.291*
p-value of joint significance test of all personality variables			{0.017}			{0.000}
log-Likelihood	-534.4	-444.2	-434.6	-326.3	-286.8	-264.9
Pseudo-R^2	0.038	0.200	0.218	0.019	0.138	0.204
Hitrate (in %)	59.93	72.58	71.71	54.56	66.81	70.75
Number of observations	806	806	806	482	482	482

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

[a] All personality variables are standardised. The big five and locus of control are initially measured on a scale from 1 to 7, where higher values indicate a stronger degree of the respective trait respectively a more internal locus of control. Risk is initially measured on a scale from 0 to 10, where higher values indicate a higher readiness to take risk.

[b] Shares are calculated by dividing the cumulative time spent in unemployment in the past by the total time spent in the labour market (as approximated by age-15).

Table A.2: Regression results on effect heterogeneity

Difference in ATTs:	Outcome variable				
	Self or regular employment			Net earned income (euros/mo.)	
High - low degree	after 21 mo.	after 40 mo.	cumulated	after 21 mo.	after 40 mo.
A. Men					
Conscientiousness					
Difference in ATTs	-0.0202	0.0166	0.11	198.42	451.58
	{0.7798}	{0.7898}	{0.9650}	{0.6196}	{0.2593}
Extraversion					
Difference in ATTs	-0.0586	0.0518	-0.92	-58.63	-367.63
	{0.3554}	{0.3704}	{0.6496}	{0.8779}	{0.3904}
Agreeableness					
Difference in ATTs	-0.0198	-0.0535	-1.26	311.69	-130.20
	{0.7588}	{0.3794}	{0.5556}	{0.4454}	{0.7497}
Neuroticism					
Difference in ATTs	0.0268	-0.0047	-0.10	-166.82	-298.78
	{0.6747}	{0.9359}	{0.9459}	{0.6837}	{0.5115}
Openness					
Difference in ATTs	0.1615**	0.0879	6.00***	218.26	768.58*
	{0.0170}	{0.1431}	{0.0050}	{0.5375}	{0.0751}
Locus of control					
Difference in ATTs	-0.0332	-0.0784	-2.13	-573.91	-456.51
	{0.5906}	{0.1602}	{0.2583}	{0.1772}	{0.3303}
Readiness to take risks					
Difference in ATTs	-0.1790***	-0.1184**	-5.68***	-321.30	-509.66
	{0.0060}	{0.0270}	{0.0020}	{0.3874}	{0.2062}
B. Women					
Conscientiousness					
Difference in ATTs	0.0606	-0.0408	-0.18	-286.83	-703.88
	{0.5526}	{0.6767}	{0.9469}	{0.3033}	{0.1041}
Extraversion					
Difference in ATTs	-0.0279	-0.0679	-1.44	488.25	174.35
	{0.7858}	{0.4835}	{0.6456}	{0.1171}	{0.7518}
Agreeableness					
Difference in ATTs	0.0448	0.0596	3.30	-98.23	67.30
	{0.6537}	{0.5736}	{0.2823}	{0.7407}	{0.8899}
Neuroticism					
Difference in ATTs	-0.0454	-0.0999	-1.88	102.34	-279.05
	{0.6216}	{0.2633}	{0.5455}	{0.7227}	{0.4645}
Openness					
Difference in ATTs	-0.0094	-0.1020	-1.67	-505.40*	-427.49
	{0.9359}	{0.3654}	{0.6797}	{0.0621}	{0.2553}
Locus of control					
Difference in ATTs	0.0232	0.1605	3.86	187.42	-87.64
	{0.8038}	{0.1051}	{0.2432}	{0.5415}	{0.8278}
Readiness to take risks					
Difference in ATTs	0.1302	0.1749*	4.68	309.79	346.05
	{0.1792}	{0.0851}	{0.1471}	{0.2823}	{0.4124}

Note: We construct a dummy variable indicating a high degree for each personality variable which takes on the value one if the value exceeds the median in the participant group and zero otherwise, separately for men and women. Presented are coefficients from a post-matching weighted regression of the outcome variables on a constant, the treatment dummy, the constructed dummy variables for all personality variables, and the interaction terms between treatment dummy and constructed personality dummies. Reported are the coefficients on the interaction terms, see Section 5.3 for details. p-values are bootstrapped based on 999 replications and denoted in braces. ***/**/* indicate significance at the 1/5/10% level.

B Supplementary Appendix

Content:

Section B.1 contains additional tables.

Section B.2 contains additional information to Section 3 in the paper. It provides details on the data collection, implementation of the survey, and an analysis of panel attrition.

Section B.3 provides further estimation results concerning sensitivity analyses for the matching results in Section 5 in the paper. A short description of the applied sensitivity tests is included.

Section B.4 contains an additional sensitivity analysis for the matching results of the extended specification testing the robustness with respect to classical measurement error in the personality variables.

B.1 Additional Tables

Table B.1.1: Descriptive statistics for control variables

	Men			Women		
	Particip. (1)	Non-part. (2)	*p*-val. (3)	Particip. (4)	Non-part. (5)	*p*-val. (6)
Socio-demographics						
Age at start-up						
Average age (years)	40.92	41.02	0.89	41.05	40.42	0.47
Younger than 30 years	0.13	0.17	0.08	0.14	0.14	0.92
30 to less than 40 years	0.31	0.27	0.28	0.30	0.33	0.38
40 to less than 50 years	0.36	0.30	0.05	0.37	0.34	0.4
50 years and older	0.20	0.26	0.06	0.20	0.19	0.79
East Germany	0.20	0.24	0.18	0.26	0.27	0.76
German citizen	0.95	0.95	0.94	0.97	0.97	0.99
Health restrictions	0.04	0.05	0.31	0.03	0.03	0.85
Married	0.64	0.56	0.02	0.55	0.55	0.94
Children						
No children	0.63	0.67	0.21	0.57	0.59	0.65
Children under 10	0.23	0.18	0.07	0.24	0.26	0.57
Children 10 years or older	0.13	0.14	0.68	0.19	0.15	0.21
Single parent	0.02	0.01	0.07	0.13	0.13	0.88
Intergenerational information						
Highest schooling degree of father						
Lower secondary school	0.45	0.43	0.60	0.36	0.36	0.98
Middle secondary school	0.15	0.21	0.05	0.18	0.23	0.22
Upper secondary school	0.26	0.21	0.05	0.32	0.28	0.36
Other/no degree	0.14	0.16	0.31	0.14	0.13	0.78
One or both parents born abroad	0.17	0.18	0.50	0.16	0.18	0.57
Father and/or mother is/was self-employed	0.33	0.29	0.28	0.31	0.31	0.90
Father employed when respondent 15 years old	0.90	0.88	0.34	0.88	0.89	0.74
Regional labour market						
Regional cluster						
Type Ia	0.14	0.13	0.79	0.18	0.17	0.76
Type Ib	0.10	0.09	0.57	0.09	0.10	0.59
Type IIa	0.07	0.07	0.79	0.06	0.05	0.68
Type IIb	0.09	0.13	0.07	0.09	0.08	0.83
Type IIc	0.07	0.06	0.93	0.05	0.06	0.53
Type IIIa	0.16	0.16	0.94	0.14	0.15	0.66
Type IIIb	0.04	0.05	0.47	0.06	0.05	0.54
Type IVa	0.09	0.08	0.53	0.06	0.05	0.82
Type IVb	0.07	0.08	0.46	0.06	0.07	0.63
Type IVc	0.04	0.03	0.50	0.04	0.01	0.07
Type Va	0.04	0.03	0.39	0.05	0.07	0.20
Type Vb	0.08	0.07	0.32	0.08	0.08	0.99
Type Vc	0.04	0.04	0.90	0.07	0.05	0.53
Human capital						
Highest schooling degree						
Lower secondary school	0.19	0.16	0.33	0.11	0.09	0.46
Middle secondary school	0.30	0.28	0.65	0.32	0.31	0.70
Upper secondary school	0.52	0.54	0.45	0.55	0.60	0.35
Other/no degree	0.00	0.01	0.02	0.01	0.00	0.47

(Table continued on next page)

(Table B.1.1 continued)

	Men			Women		
	Particip. (1)	Non-part. (2)	p-val. (3)	Particip. (4)	Non-part. (5)	p-val. (6)
Professional education						
Vocational training	0.44	0.38	0.13	0.41	0.49	0.10
Professional/vocational academy	0.18	0.13	0.08	0.12	0.05	0.01
Technical college/university degree	0.33	0.43	0.00	0.41	0.41	0.95
Other/no training	0.06	0.06	0.86	0.05	0.05	0.86
Professional qualification						
Unskilled workers	0.68	0.56	0.00	0.62	0.56	0.19
Skilled workers	0.25	0.28	0.33	0.27	0.28	0.88
Top management	0.08	0.16	0.00	0.11	0.16	0.09
Labour market history						
Lifetime unemployment (share)[a]	0.04	0.08	0.00	0.05	0.08	0.00
Second to last year before start-up						
Months employed	9.84	8.92	0.00	8.38	8.25	0.78
Months in labour market program	0.20	0.49	0.01	0.45	0.56	0.50
Next to last year before start-up						
Months employed	10.30	9.90	0.11	8.77	9.14	0.35
Months in labour market program	0.14	0.24	0.13	0.24	0.42	0.21
Last year before start-up						
Months employed	9.76	9.70	0.80	8.72	8.87	0.69
Months in labour market program	0.41	0.34	0.36	0.70	0.67	0.86
Employment status before unemployment						
Dependent employment	0.63	0.51	0.00	0.51	0.46	0.26
Disable to work/unemployable	0.03	0.08	0.00	0.05	0.13	0.00
Other	0.19	0.35	0.00	0.26	0.36	0.02
Occupational group before unemployment						
Manufacturing	0.21	0.22	0.69	0.05	0.03	0.41
Technical profession	0.08	0.08	0.92	0.02	0.02	0.92
Services	0.69	0.66	0.39	0.90	0.91	0.57
Other	0.02	0.04	0.15	0.04	0.03	0.93
Daily income from last employment (euros)	90.29	86.89	0.39	51.74	57.55	0.18
Duration of last unemployment spell						
Average number (months)	4.01	4.62	0.08	5.59	5.47	0.83
Less than 1 month	0.18	0.06	0.00	0.09	0.03	0.01
1 to less than 3 months	0.32	0.39	0.03	0.26	0.27	0.83
3 to less than 6 months	0.22	0.24	0.47	0.26	0.31	0.26
6 to less than 12 months	0.23	0.24	0.62	0.29	0.33	0.32
12 months and above	0.06	0.07	0.44	0.10	0.07	0.13
Monthly unemployment benefit						
Average amount (euros)	1,092.85	1,080.48	0.77	802.85	793.12	0.80
Less than 300 Euros	0.11	0.08	0.12	0.11	0.07	0.13
300 to less than 600 euros	0.08	0.09	0.81	0.22	0.28	0.13
600 to less than 900 euros	0.17	0.24	0.02	0.31	0.31	0.90
900 to less than 1200 euros	0.21	0.22	0.83	0.18	0.17	0.94
1200 to less than 1500 euros	0.15	0.13	0.46	0.13	0.08	0.14
1500 euros and above	0.27	0.24	0.36	0.06	0.08	0.43
Remaining unemployment benefit entitlement						
Average number (months)	8.05	7.26	0.02	6.18	6.55	0.40
Less than 3 months	0.16	0.21	0.04	0.24	0.27	0.46
3 to less than 6 months	0.17	0.16	0.72	0.24	0.16	0.03
6 to less than 9 months	0.11	0.17	0.01	0.14	0.18	0.28
9 to less than 12 months	0.24	0.23	0.70	0.15	0.22	0.08
12 months and above	0.33	0.23	0.00	0.22	0.17	0.19
Number of placement offers	1.66	2.07	0.25	1.64	2.00	0.41
Number of observations	367	439		222	260	

Note: Reported are sample averages. p-values are based on t-tests of equal means.
[a] Shares are calculated by dividing the cumulative time spent in unemployment in the past by the total time spent in the labour market (as approximated by age-15).

Table B.1.2: Measurement of personality variables

Big five: *To what degree do the following statements apply to you personally?*
Please answer on the basis of a scale ranging from 1 "does not apply at all" to 7 "applies perfectly".

I see myself as someone who ...

Q1. does a thorough job.
Q2. is communicative, talkative.
Q3. worries a lot.
Q4. tends to be lazy.
Q5. is outgoing, sociable.
Q6. values artistic experiences.
Q7. gets nervous easily.
Q8. does things effectively and efficiently.
Q9. is considerate and kind to others.
Q10. has an active imagination.

Aggregated big five indices scaled from 1 to 7:
Conscientiousness = [Q1+R(Q4)+Q8]/3
Extraversion = [Q2+Q5]/2
Agreeableness = [Q9]
Neuroticism = [Q3+Q7]/2
Openness = [Q6+Q10]/2

Locus of control: *To what degree do you personally agree with the following statements?*
Please answer on the basis of a scale ranging from 1 "do not agree at all" to 7 "agree completely".

Q1. How my life takes course is entirely dependent on me.
Q2. What one achieves is, in the first instance, a question of destiny and luck.
Q3. I often experience that others make decisions about my life.
Q4. Success is gained through hard work.
Q5. When I encounter difficulties in life, I often doubt my abilities.
Q6. I have little control over things which happen in my life.

Aggregated locus of control index scaled from 1 to 7:
Locus of control = [Q1+R(Q2)+R(Q3)+Q4+R(Q5)+R(Q6)]/6

Readiness to take risks: *To what degree are you ready to take risks in general?*
Please answer on the basis of a scale ranging from 0 "not at all ready" to 10 "perfectly ready".

Note: R() indicates reverse coding. The choice of items in our data set is closely related to the German "Socio-Economic Panel" (SOEP) survey. The SOEP uses a list of 15 items for the big five (wave 2009) based on John *et al.* (1991) (see Dehne and Schupp, 2007) and 10 items for locus of control (wave 2010) based on Nolte *et al.* (1997) where all items are surveyed using a 7-point Likert scale. Due to budget constraints, however, we had to restrict the number of items to 10 for the big five and to 6 for the locus of control in our survey. The readiness to take risks is surveyed using the general risk question with an 11-point scale (Dohmen *et al.*, 2011) which is also implemented in the SOEP.

Table B.1.3: Kernel bandwidth choice

	Optimal kernel bandwidth	
	Men	Women
	(1)	(2)
Outcome variables 21 months after start-up		
Self- or regular employed	0.21	0.27
Net earned income (euros/month)	0.26	0.31
Outcome variables 40 months after start-up		
Self- or regular employed	0.28	0.29
Cumulated effect ($\sum_{i=0}^{40}$, in months)	0.24	0.24
Net earned income (euros/month)	0.17	0.38

Note: Reported are the optimal kernel bandwidth for each outcome variable according to a leave-one-out cross-validation.

Table B.1.4: Matching quality indicators

	Before	After kernel matching with bandwidth				
	matching	0.17	0.21	0.24	0.26	0.28
Men	(1)	(2)	(3)	(4)	(5)	(6)
A. Number of variables with significant differences in means[a]						
at 1%-level	10	0	0	0	0	0
at 5%-level	18	0	0	0	0	1
at 10%-level	25	1	0	2	1	1
B. Number of variables with absolute standardised bias[b]						
$< 1\%$	5	10	16	12	16	9
1% until $< 3\%$	8	23	19	24	22	27
3% until $< 5\%$	9	20	17	18	14	19
5% until $< 10\%$	22	17	17	14	16	13
10% until $< 15\%$	14	2	3	4	4	4
$\geq 15\%$	14	0	0	0	0	0
Mean absolute standardised bias in %	10.21	3.74	3.60	3.66	3.74	3.77
Median absolute standardised bias in %	6.89	3.28	3.13	2.97	2.76	2.95
C. (Re)Estimation of the propensity score[c]						
Pseudo-R^2	0.2202	0.0294	0.0285	0.0305	0.0340	0.0347
p-value of joint significance test	0.0000	1.0000	1.0000	1.0000	1.0000	0.9999
Total number of variables	72	72	72	72	72	72
	Before	After kernel matching with bandwidth				
	matching	0.24	0.27	0.29	0.31	0.38
Women	(1)	(2)	(3)	(4)	(5)	(6)
A. Number of variables with significant differences in means[a]						
at 1%-level	6	0	0	0	0	0
at 5%-level	9	0	0	0	0	0
at 10%-level	12	0	0	0	0	1
B. Number of variables with absolute standardised bias[b]						
$< 1\%$	7	11	10	9	15	10
1% until $< 3\%$	14	26	23	24	16	22
3% until $< 5\%$	6	16	19	19	20	16
5% until $< 10\%$	21	19	19	19	20	22
10% until $< 15\%$	12	0	1	1	1	1
$\geq 15\%$	12	0	0	0	0	1
Mean absolute standardised bias in %	9.22	3.57	3.66	3.75	3.58	4.12
Median absolute standardised bias in %	6.57	2.93	3.31	3.25	3.29	3.30
C. (Re)Estimation of the propensity score[c]						
Pseudo-R^2	0.2027	0.0309	0.0328	0.0342	0.0363	0.0429
p-value of joint significance test	0.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Total number of variables	72	72	72	72	72	72

Note: Reported are indicators for covariate balancing before and after matching using the indicated optimal bandwidth from Table B.1.3 for the extended specification (cf. Table A.1).
[a] Equality of means is tested based on t-tests.
[b] The standardised bias is calculated as the difference of sample means for participants and non-participants as a percentage of the square root of the average of sample variances in both groups (Rosenbaum and Rubin, 1985).
[c] The pseudo-R^2 and the p-value of joint significance test stem from a probit (re)estimation of the propensity score on the (un)matched sample (Sianesi, 2004).

Table B.1.5: Estimation results for working age population

	Working age men			Working age women		
	Specification		Standard	Specification		Standard
	Standard	Extended	vs. extended	Standard	Extended	vs. extended
	(1)	(2)	(3)	(4)	(5)	(6)
Short-term labour market outcomes (21 months after start-up)						
Self- or regular employed	0.1387***	0.1252***	0.0136	0.1430***	0.1319***	0.0111
	{0.0000}	{0.0010}	{0.3633}	{0.0000}	{0.0000}	{0.7231}
Net earned inc. (euros/month)[a]	557.48***	534.36**	23.12	240.82**	193.90	46.92
	{0.0010}	{0.0110}	{0.7965}	{0.0260}	{0.1061}	{0.6114}
Long-term labour market outcomes (40 months after start-up)						
Self- or regular employed	0.0766***	0.0682**	0.0084	0.1026**	0.0999**	0.0027
	{0.0020}	{0.0100}	{0.4899}	{0.0120}	{0.0300}	{0.9368}
Cumulated ($\sum_{i=0}^{40}$, months)	7.87***	7.37***	0.50	7.97***	7.60***	0.37
	{0.0000}	{0.0000}	{0.2586}	{0.0000}	{0.0000}	{0.7295}
Net earned inc. (euros/month)[a]	659.45***	639.74***	19.70	621.97***	560.72***	61.25
	{0.0030}	{0.0040}	{0.8606}	{0.0090}	{0.0030}	{0.5114}
Participants off support	4	3		22	21	
Number of observations	665	665		412	412	

Note: We exclude all individuals outside working age (30-60 years). The age restriction affects a total of 141 men (50 participants and 91 non-participants) and 70 women (32 participants and 38 non-participants). Reported are estimated average treatment effects on the treated (ATT) as the difference in mean outcomes between participants and matched non-participants using Epanechnikov kernel propensity score matching. Following Huber *et al.* (2014), p-values for ATTs are bootstrapped based on 999 replications while p-values for the differences in ATTs are based on bootstrapped robust Hausman tests with 999 replications (see Cameron and Trivedi, 2009, for details). All p-values are reported in braces. The specifications are presented in more detail in the Appendix, Table A.1. ***/**/* indicate significance at the 1/5/10% level.

[a] Income measures are based on slightly lower number of observations due to item non-responses.

B.2 Detailed Data Description and Analysis of Attrition Bias

Our data set consists of a random sample of unemployed individuals who entered the start-up subsidy in the first quarter of 2009. Our comparison group contains a random sample of other unemployed individuals from the first quarter of 2009 who did not enter the programme during that time period. We combine administrative data provided by the Federal Employment Agency (FEA) with survey data obtained in two interview waves. In the following, we will describe the stepwise attrition procedure leading from the initial full sample (s_{full}) to our final estimation sample (s_{est}). Following this, we will check whether we find evidence for selectivity in terms of observable chracteristics for the specific attrition steps. The corresponding numbers are reported in Table B.2.1.

Table B.2.1: Selectivity of attrition analysis

	Sample averages			p-value of equality test: (2) vs. (3)
	s_{full}	s_{sub}	s_{est}	
	(1)	(2)	(3)	(4)
A. Participants				
Age at start-up	40.4	40.4	41.0	0.261
Male	0.641	0.625	0.623	0.949
East Germany	0.223	0.222	0.222	0.993
Upper secondary school	0.488	0.493	0.531	0.134
Employment status: self- or regular employed				
cumulated $\sum_{i=0}^{20}$	19.2	19.2	19.7	0.040
at 1st interview	0.905	0.904	0.917	0.373
Income at first interview (euros/month)				
Net earned income	1,827	1,846	1,927	0.423
Max. number of observations[a]	2306	1143	589	
B. Non-participants				
Age at start-up	40.1	40.3	40.8	0.267
Male	0.610	0.615	0.628	0.566
East Germany	0.241	0.242	0.252	0.640
Upper secondary school	0.498	0.506	0.564	0.012
Employment status: self- or regular employed				
cumulated $\sum_{i=0}^{20}$	12.0	11.9	12.1	0.435
at 1st interview	0.633	0.635	0.657	0.336
Income at first interview (euros/month)				
Net earned income	1,064	1,103	1,183	0.227
Max. number of observations[a]	2307	1390	699	

Note: Reported are sample averages and p-values for t-tests of equal means (if not denoted otherwise).
[a] The number of observations can deviate from the maximum number of observations due to item non-response.

The first interview wave was conducted in November and December of 2010 for participants and in January through March of 2011 for non-participants. The survey data were collected with computer assisted telephone interviews (CATI). The initial full sample (s_{full}, Table B.2.1, column 1) consists of 2,306 individuals in the participant group (Panel A) and 2,307 individuals

in the non-participant group (Panel B). Due to budget constraints, we collected a set of additional information during the interviews only for a 50% random subsample in both groups (s_{sub}, column 2). The subsamples consist of 1,143 participants and 1,390 non-participants. Since these individuals were chosen randomly from the initial full sample, the attrition is not selective.

Overall, 632 participants and 789 non-participants also took part in the second interview. This second wave was carried out in August through October of 2012. The reasons for panel attrition in the second wave consisted of failure to contact the targeted individuals, inability of the target individual to participate in the interview, and refusal to participate. Respondents also had to give their consent to combining administrative and survey data. Only a small minority (15 participants and 13 non-participants) objected to the data merge. The resulting sample size decreases to 617 participants and 776 non-participants. The number of observations in the final estimation sample (s_{est}, column 3) deviates from these numbers of respondents due to occasional item non-responses in variables relevant for our estimation procedure.

We conduct a selectivity analysis for each attrition step to check whether there is a systematic relationship between the attrition and the outcome variables we consider in our main analysis. We therefore test for each potentially non-random attrition step whether the average outcome variables before and after the attrition differ significantly. If there is selective attrition, there should be systematic differences in the outcome variables before and afterwards.

In addition to outcome variables, we also consider selected socio-demographic features from the first interview because these variables are observed for all individuals. Outcome variables gathered in the second interview cannot be considered because they are unavailable for respondents who participate only in the first but not in the second interview. The selectivity analysis is conducted separately for participants (Table B.2.1, Panel A) and non-participants (Panel B).

Since the selection of the 50% subsample (s_{sub}, column 2) was randomly determined by the survey institute, we test this subsample against our final estimation sample (s_{est}, column 3). The p-values of the corresponding equality tests are reported in column 4. Overall, we do not find strong evidence for a systematic relationship between sample attrition and socio-demographic features or labour market outcomes of the respondents as indicated by the overall small differences in magnitude between the 50% subsample and the final estimation sample as well as by the overwhelming majority of high p-values in column 4. The same is true for non-participants as the numbers in column 4 of Panel B show.

In conclusion, we find no strong evidence for systematic selection. Thus, we do not have to introduce weights in our empirical analysis.

B.3 Sensitivity Analysis

We test the sensitivity of our main estimation results for the extended specification by imposing three alternative common support restrictions, altering the choice of the matching algorithm, and we examine the robustness of the ATTs if we account for time-invariant unobserved heterogeneity by estimating two conditional difference-in-differences approaches. In sum, the results of the robustness checks (reported in Table B.3.1) are very similar to our main analysis.

Trimming the estimation sample: If there is only a limited number of non-participant observations at certain parts of the propensity score distribution, this lack of overlap results in large weights for single non-participant observations in the propensity score matching procedure. This is most likely at the tails of the propensity score distributions. As a consequence, matching estimators will be prone to a large bias and variance. To deal with this problem, Crump *et al.* (2009) propose a systematic approach to optimally trim the sample in order to maximise estimation precision. Based on this approach, we trim the sample by dropping all 113 male observations with an estimated propensity score outside the interval $[0.1089; 0.8911]$, and 57 female observations with an estimated propensity score outside the interval $[0.1056; 0.8944]$ are dropped accordingly. The propensity score matching is then performed on the trimmed sample (see Table B.3.1, lines labelled as 'trimmed sample 1').

Huber *et al.* (2013) propose an alternative two-step trimming procedure which also readjusts the sample after trimming to correct for implicit changes in the reference subpopulations induced by the trimming. The first step is to remove all non-participants from the sample with a weight share larger than a certain threshold. In the second step, the weights of the remaining non-participant observations are normalised again. The correction ensures that the procedure is asymptotically unbiased. In our case, we remove all non-participants with a weight share larger than 1%. Overall, this trimming procedure leads to a reduction of our estimation sample by 11 to 18 male and 24 to 25 female non-participant observations (depending on the optimal bandwidth choice), while for the optimal bandwidth of 0.28 (0.38) for self- or regular employed after 40 months (net earned income after 40 months), no male (female) non-participant observation exceeds the threshold of 1% of the weight share. Estimation results are labelled 'trimmed sample 2' in Table B.3.1.

As an alternative three-step procedure, the sample is additionally corrected for the trimming after the first step by dropping all participant observations with an estimated propensity score larger than the smallest propensity score of the dropped non-participant observations to ensure common support. In a third step, matching weights for non-participants are normalised again.

Table B.3.1: Sensitivity analysis

	Men		Women	
	ATT	p-value	ATT	p-value
	(1)	(2)	(3)	(4)
A. Short-term labour market outcomes (21 months after start-up)				
Self- or regular employed	0.1336***	0.0000	0.1669***	0.0000
trimmed sample 1	0.1423***	0.0040	0.1710***	0.0000
trimmed sample 2	0.1232***	0.0000	0.1948***	0.0000
trimmed sample 3	0.1233***	0.0000	0.1883***	0.0000
radius matching 1	0.1537**	0.0110	0.1540*	0.0581
radius matching 2	0.1468***	0.0000	0.1684***	0.0000
Net earned income (euros/month)	600.55***	0.0010	313.48***	0.0030
trimmed sample 1	625.94**	0.0250	329.42**	0.0210
trimmed sample 2	611.27***	0.0000	394.41***	0.0040
trimmed sample 3	578.91***	0.0070	340.59**	0.0160
radius matching 1	595.06**	0.0190	330.60*	0.0761
radius matching 2	629.25***	0.0010	337.19**	0.0130
B. Long-term labour market outcomes (40 months after start-up)				
Self- or regular employed	0.0800***	0.0010	0.1056***	0.0090
trimmed sample 1	0.0834	0.1431	0.1160**	0.0260
trimmed sample 2	(0.0800***)	(0.0040)	0.1192***	0.0050
trimmed sample 3	(0.0800***)	(0.0040)	0.1149**	0.0200
radius matching 1	0.0865**	0.0320	0.1072	0.1882
radius matching 2	0.0812***	0.0020	0.1312***	0.0010
Cumulated effect ($\sum_{t=0}^{40}$, in months)	7.76***	0.0000	8.56***	0.0000
trimmed sample 1	7.88***	0.0000	8.80***	0.0000
trimmed sample 2	7.27***	0.0000	10.19***	0.0000
trimmed sample 3	7.18***	0.0000	9.66***	0.0000
radius matching 1	9.25***	0.0000	7.98***	0.0010
radius matching 2	8.25***	0.0000	8.81***	0.0000
conditional DID 1	8.20***	0.0000	10.15***	0.0000
conditional DID 2	7.81***	0.0000	8.38***	0.0000
Net earned income (euros/month)	736.64***	0.0010	608.57***	0.0040
trimmed sample 1	741.80***	0.0040	658.38***	0.0030
trimmed sample 2	614.32***	0.0040	(608.57***)	(0.0070)
trimmed sample 3	589.61**	0.0220	(608.57***)	(0.0070)
radius matching 1	685.53**	0.0380	642.96**	0.0210
radius matching 2	745.90***	0.0020	618.82***	0.0040

Note: Presented are estimated average treatment effects on the treated as the difference in mean outcomes between participants and matched non-participants using Epanechnikov kernel propensity score matching for the extended specification with optimal bandwidth based on leave-one-out cross-validation (if not denoted otherwise). Following Huber *et al.* (2014) p-values are bootstrapped and based on 999 replications. ***/**/* indicate significance at the 1/5/10% level.
Trimmed Sample: The treatment effects are estimated based on an optimally trimmed sample following Crump *et al.* (2009) (trimmed sample 1), Huber *et al.* (2013) (trimmed sample 2), and, following the latter, with the additional imposition of common support (trimmed sample 3). For the optimal bandwidth in the male (female) sample for self- or regular employed (net earned income) after 40 months, no observations were trimmed in the second and third procedure so results are identical to the main results and reported in parentheses.
Radius Matching: The treatment effects are estimated using radius matching with bias adjustment following Huber *et al.* (2014, 2013) (radius matching 1) and radius matching using a caliper of 0.1 (radius matching 2).
Conditional DID: The treatment effects are estimated based on conditional difference-in-differences. The reference level is months in regular employment during the ten years prior to start-up (conditional DID 1) and six months prior to start-up (conditional DID 2).
See text for details.

In addition to the trimmed non-participant observations mentioned above, this procedures leads to a reduction of 96 to 104 male participants and 73 to 75 female observations who exceed the smallest propensity score of the dropped non-participants, again depending on the optimal bandwidth choice. In the case where no non-participants were dropped (see above), the participant subsample remains unchanged accordingly and results are reported in parentheses. Estimation results are labelled 'trimmed sample 3' in Table B.3.1.

Altering the matching algorithm/radius matching: As an alternative matching algorithm, we consider the radius matching estimator with a bias adjustment proposed by Huber *et al.* (2014, 2013). It consists of distance-weighted radius caliper matching on the propensity score, where non-participant observations are weighted proportionally to their inverse distance to the participant observations within the caliper. In a second step, this estimator uses the weights from the matching procedure in the first step for a linear regression to correct the estimators for any remaining biases due to mismatches. Estimation results are reported in Table B.3.1 in lines labelled 'radius matching 1.' As an alternative, we consider radius matching with a caliper of 0.1, where all comparison units within the caliper are weighted equally to construct the counterfactual outcome (Table B.3.1, labelled 'radius matching 2').

Conditional difference-in-differences: To test the sensitivity of our results with respect to the presence of additive linear time-constant unobserved heterogeneity, we also perform a conditional difference-in-differences estimation approach for the cumulated employment effects. For the reference level before treatment, we consider the number of months in regular employment during two time periods: the ten years prior to start-up (Table B.3.1, conditional DID1) and the six months prior to start-up (Table B.3.1, conditional DID 2).

Overall, the various robustness checks yield point estimates that are very similar to our results from the main analysis for all outcome measures. In conclusion, the sensitivity checks thus show a consistent picture of robust positive and significant effects of participation in the start-up subsidy programme with respect to labour market reintegration and earned income for both men and women.

B.4 Sensitivity Analysis with Respect to Measurement Error in the Personality Variables

We conduct another sensitivity analysis to investigate whether the estimation results in the extended specification are biased due to classical measurement error in the personality variables following Battistin and Chesher (2014). They show that measurement error in covariates used in treatment effect analyses based on the conditional independence assumption (like propensity score matching) does not necessarily imply attenuation in the estimated treatment effects. If the estimated treatment effects of our extended specification would potentially be attenuated, our estimate of the difference to the standard specification would be an upper bound to the real difference. Instead, the sign of the bias depends on the relationship of the erroneously measured covariate and the propensity score and on its relationship with the counterfactual outcome. Therefore, the potentially measurement error-contaminated personality variables could lead to a positive bias which in turn would lead to an underestimated difference between the specifications without and with the inclusion of our personality variables.

Table B.4.1 reports the estimated approximate biases in the treatment effects of the extended specification containing all personality variables. To facilitate assessment of the magnitudes of the estimated biases, we report as a reference the estimated effects from our main results from Table 6 in the first line of each sub-panel labelled 'ATT' in Table B.4.1. We vary the extent of measurement error in the personality variables from a signal-to-noise ratio of 10% up to 50%. For instance, the estimated treatment effect for the outcome variable "self- or regular employed 21 months after start-up" for males is 13.36 percentage points (Panel A, column 1 to 3, as already reported in Table 6). If the big five variable "conscientiousness" was measured with error that accounted for 10% of the total variance in conscientiousness (column 1), the treatment effect for males would be overestimated by 0.01 percentage points. If the measurement error was responsible for 50% of the total variance in conscientiousness (column 3), the bias for males would amount to 0.07 percentage points. For females, the corresponding treatment effect is 16.69 percentage points while a measurement error of 10% (50%) would lead to an underestimation by 0.04 (0.21) percentage points (column 4 and 6, respectively).

The numbers in Table B.4.1 show that the estimated approximations for the biases due to measurement error in the personality variables are small in magnitude throughout and insignificant at conventional levels without exception for men and women. Thus, we are confident that our results are robust with respect to measurement errors in personality variables. The varying signs of the biases confirm that measurement errors in treatment effect analyses based on the conditional independence assumption do not generally lead to attenuation of the effects.

Table B.4.1: Sensitivity analysis with respect to measurement error in personality variables

	Men			Women		
	Extent of measurement error			Extent of measurement error		
	10%	30%	50%	10%	30%	50%
Estimated approximate bias	(1)	(2)	(3)	(4)	(5)	(6)
A. Short-term labour market outcomes (21 months after start-up)						
Self- or regular employed	ATT = 0.1336			ATT = 0.1669		
Big five						
Conscientiousness	0.0001	0.0004	0.0007	-0.0004	-0.0013	-0.0021
Extraversion	0.0000	0.0000	0.0000	0.0004	0.0012	0.0021
Agreeableness	-0.0001	-0.0003	-0.0005	-0.0001	-0.0002	-0.0003
Neuroticism	-0.0000	-0.0001	-0.0001	-0.0003	-0.0009	-0.0016
Openness	-0.0002	-0.0007	-0.0011	-0.0000	-0.0001	-0.0001
Locus of control	0.0007	0.0020	0.0033	0.0019	0.0057	0.0096
Readiness to take risks	-0.0002	-0.0006	-0.0010	-0.0014	-0.0042	-0.0070
Net earned income (euros/month)	ATT = 600.55			ATT = 313.48		
Big five						
Conscientiousness	0.98	2.95	4.91	-0.16	-0.47	-0.78
Extraversion	-0.01	-0.02	-0.03	-0.11	-0.32	-0.54
Agreeableness	-0.40	-1.20	-1.99	-0.13	-0.39	-0.65
Neuroticism	0.40	1.20	2.00	1.30	3.89	6.49
Openness	-2.24	-6.71	-11.18	0.44	1.32	2.21
Locus of control	3.81	11.43	19.06	4.62	13.87	23.11
Readiness to take risks	-1.00	-3.01	-5.01	-1.38	-4.15	-6.92
B. Long-term labour market outcomes (40 months after start-up)						
Self- or regular employed	ATT = 0.0800			ATT = 0.1056		
Big five						
Conscientiousness	-0.0000	-0.0001	-0.0002	-0.0004	-0.0013	-0.0021
Extraversion	0.0000	0.0000	0.0000	0.0002	0.0007	0.0012
Agreeableness	0.0001	0.0004	0.0007	-0.0000	-0.0001	-0.0002
Neuroticism	-0.0001	-0.0002	-0.0004	-0.0002	-0.0006	-0.0010
Openness	0.0002	0.0005	0.0009	0.0004	0.0011	0.0018
Locus of control	0.0015	0.0045	0.0074	0.0001	0.0003	0.0005
Readiness to take risks	-0.0002	-0.0005	-0.0009	-0.0011	-0.0032	-0.0053
Cumulated effect ($\sum_{t=0}^{40}$, months)	ATT = 7.76			ATT = 8.56		
Big five						
Conscientiousness	-0.0011	-0.0034	-0.0057	-0.0095	-0.0286	-0.0476
Extraversion	0.0001	0.0002	0.0003	0.0142	0.0425	0.0709
Agreeableness	-0.0002	-0.0005	-0.0008	-0.0037	-0.0112	-0.0187
Neuroticism	0.0013	0.0038	0.0063	-0.0047	-0.0142	-0.023
Openness	-0.0078	-0.0235	-0.0392	0.0066	0.0197	0.0329
Locus of control	0.0283	0.0850	0.1417	0.0806	0.2417	0.4028
Readiness to take risks	-0.0063	-0.0190	-0.0317	-0.0526	-0.1577	-0.2629
Net earned income (euros/month)	ATT = 736.64			ATT = 608.57		
Big five						
Conscientiousness	1.79	5.36	8.93	-1.03	-3.09	-5.15
Extraversion	-0.01	-0.04	-0.07	0.23	0.69	1.15
Agreeableness	0.48	1.44	2.39	-0.09	-0.27	-0.46
Neuroticism	0.45	1.35	2.25	-1.01	-3.04	-5.07
Openness	-2.29	-6.89	-11.44	-3.79	-11.37	-18.95
Locus of control	8.95	26.84	44.74	7.17	21.51	35.86
Readiness to take risks	-1.20	-3.59	-5.98	-4.02	-12.07	-20.12

Note: Presented are estimated approximations to measurement error biases in the estimated treatment effect of the extended specification due to classical measurement error in the listed personality variables following Battistin and Chesher (2014). The extent of measurement error is defined as the noise-to-signal ratio. See text for details. Standard errors are bootstrapped and based on 999 replications. ***/**/* indicate significance at the 1/5/10% level.