

Are Programmes that Allow Job Centres Considerable Freedom to Choose the Exact Design Effective?

Tamara Harrer^{*}, Andreas Moczall[†], Joachim Wolff[‡]

Preliminary version, please do not cite!

Abstract:

In 2009 Germany reformed its active labour market policy (ALMP) portfolio. The number of programmes was reduced and local job centres were given greater latitude in the design and assignment of participants to particular interventions. Among the affected programmes were the several types of short training measures and the outsourcing of placement services to private providers. These two programme types were combined, along with a number of smaller programmes, into the new “schemes for activation and integration” (SAI). The aim of that reform was to allow for more innovative and better approaches to reducing individual impediments to employability and finding work for participants, and to streamline and simplify the process of awarding public contracts. The policy makers allow a considerable freedom to design the scheme. In turn they expect that the public employment services implement measures that adequately address individual employment impediments of disadvantaged welfare recipient. Among the innovations is the ability to combine training and placement in employment in one single programme assignment. Participants of such a programme combination can thus immediately apply the new knowledge they have gained during the classroom sessions.

We evaluate these reshaped schemes with propensity score and difference-in-difference matching using large and rich administrative data. Although both recently-employed unemployment insurance recipients and welfare benefit recipients can participate, we solely consider welfare benefit recipients that were unemployed at the end of 2009 and how their comparatively smaller employment and larger benefit receipt probabilities change as a result of SAI participation in the first quarter of 2010. We distinguish assignments taking place in firms from those taking place in other settings and separate the results by gender, region, age and unemployment duration. Our results indicate that while SAI substantially

^{*} Tamara Harrer, Institute for Employment Research (IAB), Research Unit C2: Basic Income Support and Activation. Postal address: Regensburger Straße 104, 90478 Nuremberg, Germany. Email: tamara.harrer@iab.de, Phone: +49 (911) 179-3468.

[†] Andreas Moczall, Institute for Employment Research (IAB), Research Unit C2: Basic Income Support and Activation. Postal address: Regensburger Straße 104, 90478 Nuremberg, Germany. Email: andreas.moczall@iab.de, Phone: +49 (911) 179-3066.

[‡] Joachim Wolff, Institute for Employment Research (IAB), Research Unit C2: Basic Income Support and Activation. Postal address: Regensburger Straße 104, 90478 Nuremberg, Germany. Email: joachim.wolff@iab.de, Phone: +49 (911) 179-1248.

improve the employability of their participants in the three years after participation, in-firm training more so than training in other settings, the improvements are no greater than those previously found for comparable pre-reform programme types.

Keywords: welfare recipients, activation, propensity score matching

JEL classification: J68, C14, I38, H43

1 Introduction

From the 1980s onwards, the German economy became increasingly plagued by high and persistent unemployment. The German government responded with a series of wide-ranging reform laws that were enacted between 2003 and 2005, the so-called “Hartz” reforms. The fourth of these reform laws received the most attention, as it introduced a system of rights and duties (“Fördern und Fordern”) for welfare recipients receiving the new means-tested “basic income support for job-seekers”. From the beginning of the fourth Hartz reform, active labour market programmes (ALMPs) played a major role in the efforts to integrate unemployed welfare recipients into the labour market.

The year 2009 saw the first major revision of the ALMP portfolio (“Neuaustrichtung der arbeitsmarktpolitischen Instrumente”). Based on the lessons learned from policy evaluations and from the experience of public employment services (PES) who implemented ALMPs, programmes were streamlined and simplified for greater effectiveness. For example, short training programmes had existed in several varieties, most of them subject to rather detailed and specific provisions and procedures. By abolishing rarely-applied programmes and replacing administrative minutiae with a generalized description of means and objectives, the PES gained some degrees of freedom in designing interventions in ways most appropriate for any particular individual. One such programme is “schemes for activation and integration” (SAI), which replaced short training programmes and the contracting-out of placement services. The reform bill’s draft calls for these new schemes to be applied in such a way that it suits the specific needs of welfare recipients to raise their chances of finding a job or even to directly integrate them into a job (German Federal Parliament 2008: 25).

Whether these new flexible “schemes for activation and integration” perform well in raising the employment prospects of welfare recipients is one research question of this study. A second, related research question is whether they perform considerably better than the programmes that they replaced. If so, it would provide important evidence for providing public employment services with a substantial degree of freedom in implementing active labour market policies. Therefore, we will compare the results of our study to a benchmark of results from similar studies on the participation effects of former short-term training and contracting out placement services for welfare recipients.

Calmfors (1994) discusses how participants may benefit from ALMPs. SAI’s short training and private placement services may teach participants to search for jobs more effectively, and may lead to better matches between job-seekers and employers. During the programme, participants usually reduce their job search activities, resulting in lower employment rates until the programme ends (“lock-in effect”).

Well-assigned participants may benefit from a subsequently higher job take-up rate and thus more exits from welfare receipt. They may also become more attached to the labour market, leading to fewer drop-outs of the labour force and thus fewer exits from welfare receipt even without holding a job. ALMPs involving pressure to quickly find jobs (work tests) may also lead to an increased take-up of lower-quality jobs that are insufficient to leave welfare. However, for participants assigned to programmes that do not address their specific employment impediments participation might negatively affect their employment rate and lead to increased exits from welfare without finding employment. Therefore, it is not clear whether on average participation effects will lead to a reduced or an increased rate of welfare receipt and/or earnings among the (former) participants of SAI.

Our analysis is based on a large sample drawn from administrative data. We study the (entire) stock of unemployed welfare recipients at the end of the year 2009. As a treatment group we regard those sample members that entered schemes for activation and integration during the first three months of the year 2010. The analysis distinguishes between two sub-schemes: (pure) in-firm training and schemes of short classroom training and/or placement services, both organised by private providers. As a control group or waiting group we defined all unemployed welfare recipients in our stock sample that did not enter these programmes in the first quarter of 2010, but may have entered later. We estimate, using methods of propensity score matching, the effects of participation in the two programmes on a large number of outcomes. Among them we consider working in unsubsidized contributory (regular) employment, receiving welfare benefit and real earnings. Our observation period ends in the year 2013. As the participation in the scheme under consideration lasts often only for a few weeks, this period is long enough to make statements about medium- up to long-term effects after programme termination.

The paper is structured as follows: In section 2 we briefly discuss the welfare-to-work systems as introduced from the 1990s onwards. The section points out that flexibility in programme design and the effectiveness of such “flexible” programmes is a new topic of evaluation studies. Section 3 introduces the institutional framework of the welfare benefit regime and of the schemes for activation and integration. A review on key results of empirical studies on the employment effects of participating in related programmes is provided in section 4, with a special emphasis on studies on the German programmes that were replaced by schemes for activation and integration. The methods are discussed in section 5 and section 6 describes the data base and the sample studied. Our estimates are presented in section 7. The final section 8 concludes.

2 Decentralised welfare-to-work policies and localising programme design

In the 1990s traditional welfare systems and employment policies were not prepared to provide intensive support for people who were detached from the labour market, but rather concentrated on passive income support and on helping people who were temporarily unemployed (OECD 1999: 10). Welfare-to-Work reforms in a number of countries attempted to overcome this problem. Some countries introduced such reforms already during the 1990s (e.g., in the United States or United Kingdom) others followed after the start of the millennium like Germany with the introduction of the Basic Income Support for Job Seekers or Social Code (SC) II in the year 2005.

Although the welfare-to-work policies differed considerably across countries there was also some common ground: One important common issue is that welfare-to-work policies are ultimately delivered at the local level (OECD 1999: 26) and one of the key advantages of decentralised implementation of such policies is that they can be adapted to local needs and circumstances (OECD 1999, Nativel et al. 2002).¹ However, there are also some disadvantages to decentralised welfare-to-work policies such as an uneven provision of services or a lack of accountability, as local agencies might not pursue the same goals of the central government and might not implement policies in the way intended by the central government (OECD 1999, Lundin/Skedinger 2006).

Decentralisation may relate to different aspects. The local flexibility of (active) labour market policies may relate to the design of the programmes, allocating budgets, defining target groups, setting performance criteria, collaborating with other actors and outsourcing (Froy/Giguère 2009a: 37). Under the Social Code II introduced in 2005 German job centres can choose the programme mix at the local level and have discretion of the use of their budgets within the limits of the goals for national policy delivery outsourcing (Froy/Giguère 2009a: 40). Froy and Giguère (2009a: 40) point out that in OECD countries local actors do not have a significant role in the design of labour market policies and programmes. In Germany a reform of the year 2009 though introduced a programme that allows for a flexible design and hence a significant role of the PES in programme design concerning short training and contracting out placement services. Therefore, we will study the effectiveness of this new programme, “schemes for activation and integration”, that allows job centres to design participations according to the needs of the welfare recipients and that replaced former training schemes and schemes of private placement services that did not allow for such a flexibility. We can therefore compare our results to the results on the effectiveness of

¹ Of course there exist more advantages, e.g., the capacity to co-ordinate locally a range of different national, regional and local policies affecting employment or the ability to mobilise relevant local actors (OECD, 1999).

the former programmes to provide an idea whether introducing design flexibility raised the effectiveness of participation by welfare recipients. To our knowledge this is the first study that provides such a comparison. And the topic is of considerable interest as researchers have not yet identified a solid relationship between decentralisation and policy effectiveness (Froy/Giguère 2009b: 23). We now turn to the details of the new programme.

3 Schemes for activation and integration under the basic income support for job-seekers and unemployed

Schemes for activation and integration - currently Article 45 of Social Code III - are a set of active labour market programmes which target unemployed people, employed people at risk of unemployment, and people searching for an apprenticeship. Most of those schemes apply both to unemployed people receiving the unemployment insurance (UI) benefit or the unemployment benefit (UB) II – the means-tested welfare benefit - according to the SC II.² The SAI do not focus on specific groups of unemployed people. In fact and due to their flexibility, the instruments can be designed in various ways to satisfy individual needs of different groups participating in the labour market. Though the instruments target almost all types of unemployed, we consider only those schemes available to UB II recipients (currently Article 16 of SC II in conjunction with Article 45 of SC III).

Introduced in 2009, SAI replaced a number of older instruments, in particular short-term training programmes. The main aim of the reform was to reduce the overall number of programmes, and to give local case workers greater flexibility in developing and choosing measures that will be successful at integrating unemployed job-seekers into the labour market. The schemes can be run by either (private) training providers and placement services (schemes by providers – SP) or employers (in-firm training – IFT). These two types differ in aims and locations.

IFT takes place in an establishment and can last up to four weeks.³ IFT replaces the previous in-firm training measures and aims to get participants accustomed to the regular life of an employee and the (specific) employment situation in the company. Participants can test their practical skills and identify potential deficits. IFT therefore helps determine the

² The German unemployment benefit system distinguishes between two types of unemployment benefits. People who become unemployed and who paid for a sufficient period of time contributions to the unemployment insurance system receive 60 per cent (67 per cent for parents) of their last net wage as UI benefit. Its rules are regulated in the SC III. The duration of benefit receipt depends on the length of contribution record during the four years prior to unemployment. Its maximum duration is one year for those aged younger than 50 years and rises for older age-groups up to a maximum of two years. A second benefit is the means-tested welfare benefit of the SC II. This basic income support is available not just for long-term unemployed, but also for employed persons and unemployment insurance recipients if their household income is below the poverty line.

³ Due to a reform in April 2012, the design of SAI was modified. For observing participation in 2010, we describe the schemes' design before the reform. Further information about changes in design is available on request.

participants' professional skills and suitability for specific occupations. In practice it can be compared to unpaid internships because participants do not receive payment but are instructed and supervised by qualified employees of the firm hosting the programme, allowing them to learn on the job and improve their skills. Unless the participant is kept as a regular worker, the firm hosting the programme must provide a report detailing which skills the participants has acquired.

In contrast, SP focuses on improving (personal) skills, reducing individual impediments to employability, and finding work for participants. SP replaces the previous classroom training measures and the delegation of placement services to private providers (Article 37 SC III, prior to 2009). Among these schemes are different types regarding specific content. *Guiding into apprenticeships and into work*, for example, teaches participants to choose suitable job offers and write CVs and application letters to improve their chances of finding a job. *Determining, Reducing and Removing Employment Impediments* focuses on finding out which particular attributes define the individual's disadvantage, improving participants' skills, and gaining knowledge about certain occupational fields and individual opportunities on the labour market. Other SP types offer placement services both in employment and vocational training, support gaining skills as well as information about self-employment, and stabilisation in case of employment that has already begun. Stabilisation schemes motivate participants and teach how to avoid conflicts regarding work and family life in order to avoid further unemployment and benefit receipt; they last for up to six months.

Schemes can be combined if necessary, and providers may combine elements from different IFT types to suit individual needs. Any SP may be accompanied by four weeks of work to apply the knowledge gained during the training programmes. Here, participants can either enter some kind of internship like IFT or they complete types of workshops organized by the provider and focussing on practical skills.

While IFT can last up to four weeks, SPs in principle are not limited in duration. However, if they teach skills for a particular profession, the schemes can last up to only eight weeks. Training measures on personal competences, like writing CVs and application letters, are not limited. In our sample, the average duration of SP is 2.5 months.

SAI are quantitatively important. Even upon their introduction in 2009, more than 550,000 UB-II-recipients took part (Table 1). Since then, the inflow increased up to over 960,000 participants in 2010 and stabilised at around 785,000 inflows each year between 2012 and 2015. Accordingly, the costs rose from about 200 million euros in 2009 up to more than 600 million in 2010 and then stabilised between 450 and nearly 530 million euros during the last years (Table 2). Most inflows are into SP (72% to 80%); only 20% to 28% of the inflows enter IFT (Table 1). Not surprisingly, the total programme expenditure is higher for SP than for IFT. While IFTs cost less than 10 million euros each year (except in 2010), the

annual average costs for SPs are 486 million euros (Table 2). With respect to the average costs per participant and month and in contrast to SP, IFTs are rather cheap programmes. The costs do not exceed 100 euros while SPs are more expensive with about 870 euros. However, these costs range from about 400 euros to more than 1,200 euros per participant and month depending on the particular programme type (Table 2).

Table 1: Inflow into different schemes for activation and integration of unemployed people receiving UB II from 2009 to 2015 (in 1,000)

	2009	2010	2011	2012	2013	2014	2015
SAI - schemes for activation and integration							
total	550.8	960.9	781.2	750.6	775.4	820.1	797.6
West Germany	390.8	693.6	561.1	540.1	559.5	593.8	581.2
% female	39.4	39.4	40.6	41.7	41.7	41.9	41.5
East Germany	159.9	267.4	220.1	210.5	215.9	226.3	216.3
% female	43.4	42.9	43.4	43.3	43.8	44.1	44.6
SP - schemes by providers							
total	382.1	744.3	578.5	581.6	608.0	652.1	633.8
West Germany	286.2	556.2	435.1	437.0	455.9	489.0	478.6
% female	41.4	41.2	42.5	42.9	42.9	43.0	42.7
East Germany	95.9	188.1	143.4	144.6	152.1	163.0	155.2
% female	44.0	43.9	44.1	43.1	43.5	43.8	44.1
IFT - in-firm training							
total	168.7	216.7	202.7	169.0	167.5	168.0	163.7
West Germany	104.6	137.4	126.0	103.1	103.6	104.8	102.6
% female	34.0	32.2	33.9	36.4	36.3	36.7	36.1
East Germany	64.1	79.3	76.7	65.9	63.8	63.3	61.1
% female	42.4	40.4	42.3	43.8	44.4	44.9	45.7

Source: Department of Statistics of the Federal Employment Agency, calculation from the Data Warehouse.

Table 2: Expenditure for the schemes for activation and integration (only for UB II recipients) ¹⁾

	2009	2010	2011	2012	2013	2014
<i>Total expenditure (in 1,000 euros)</i>						
SAI	207,024	603,197	465,639	400,324	465,975	527,447
SP	-	592,570	458,447	398,051	462,147	523,328
IFT	-	10,627	7,193	2,274	3,827	4,120
<i>Expenditure per participant and month (in euros)</i>						
SAI	344	426	402	772	423	967
SP	-	441	420	1,022	441	1,245
IFT	-	145	109	18	71	33

¹⁾ without data from ALP job centres.

Source: Department of Statistics of the Federal Employment Agency.

4 Previous research

Because schemes for activation and integration were introduced as a replacement for former short-term training measures and outsourcing of placement services to third parties, we discuss results of various studies looking at the effectiveness of training measures and

placement services by private providers. Of special interest is the German research on participants receiving means-tested benefit as well as research of comparable ALMPs in other countries. All but one of the studies in our discussion analysed micro data to estimate treatment effects on the treated and applied methods of propensity score matching, probability weighting and timing of events approaches. Although the discussed studies are experimental as well as non-experimental, all of them estimate effects on some employment outcome like working in an unsubsidized contributory job. The reason is that studies wanted to show whether the programmes are effective in the sense of raising the chances of integrating participants into work.

A number of studies regard a sample of short-term training participants who entered the programme in 2005 and a group of controls using administrative data on unemployed welfare recipients in Germany. In general, short-term training participation increases the employment rate of participants in Germany (Wolff/Jozwiak 2007, Kopf 2013, Hartig et al. 2008, Zabel 2013), whereas the type of training matters: in-firm training participation is far more effective for its participants in terms of integrating them into work than classroom training participation is (Wolff/Jozwiak 2007). If UB II recipients participate in short training programmes taking place in companies, their probability of being employed increases up to 19 percentage points a few months after programme start. In contrast, classroom training participation does increase the participants' employment rates up to three percentage points, but is accompanied by lock-in-effects occurring in the short run (Wolff/Jozwiak 2007). Furthermore, the classroom training participation on average does not help leaving welfare benefit receipt while for participants in in-firm training the treatment reduces substantially their rate of dependency on welfare benefit (Wolff/Jozwiak 2007). The effects of participating in short-term training programmes also differ by type of programme and by varying goals. Kopf (2013) and also Bernhard and Kopf (2014) detected predominantly negative effects on employment chances and the probability of not receiving means-tested benefits if UB II recipients participated in classroom application training, work tests, and classroom combinations. In contrast, participants benefited by lowering their risk of unemployment and welfare benefit receipt if they attended classroom aptitude tests, skill training, and within-company combinations according to the results of Kopf (2013).

Depending on age and former employment, the study of Wolff and Jozwiak (2007) found some variation in the participation effects in short-term training. 20 months after programme start, benefit recipients younger than 25 years who attended classroom training have the same employment rates as their controls not attending a training programme. Younger participants in in-firm training do have higher employment rates than their controls but they benefit less than older ones: For participants older than 50 years in East Germany, the treatment effect on regular employment is twice as high as for participants younger than

25 years (Wolff/Jozwiak 2007). Looking at participants at the start of the year 2005 and therefore immediately after the introduction of the UB II, Wolff and Jozwiak (2007) also found the following: Participants who did not hold a regular job for two to four years are characterised by higher than average estimated treatment effects on the employment rate than all of the other groups. Therefore, they benefit more from participation than those not employed for one year, last employed before 2001 and who have never worked.

Job search assistance (JSA) is an ALMP comparable to former German application training measures and therefore often conducted together with other short-term training measures. Research from other countries on the effects of these ALMPs on participants' employment chances and earnings is heterogeneous in the design of various studies as previously described and also the direction of the estimated effects. While meta-analyses attest positive effects of JSA (Thomsen 2009, Card et al. 2010), studies exist showing negative or inconsistent results (Gerfin/Lechner 2002, van den Berg/van der Klauuw 2006). Similar to Germany, researchers analysing participation in short-term training found positive effects only for particular target groups and types of training: E.g. for Switzerland, Prey (1999) detects increased employment chances for people participating in German language courses but not for attending courses improving computer applications. Swedes participation in training measures leads to shorter unemployment spells, even if lock-in effects are considered according to results of Vikström (2015). In Austria especially women benefit in higher into-job transitions from participation in job-search programmes (Weber/Hofer 2003). Classroom training also has positive effects on earnings of participants according to findings for the United States (US) (Ashenfelter 1978). In the Netherlands, Gorter and Kalb (1996) conducted a social experiment with unemployed receiving unemployment benefits. They compared participants of the Counseling and Monitoring programme, which is an intensive JSA programme, with non-participants. Their study shows that the programme reduces the time taken to find a job because the participants write more applications than their controls. However, the employment opportunities do not differ between participants and their controls (Gorter/Kalb 1996).

SAI can consist in the outsourcing of placement services to third parties. Previous research on outsourcing of placement services to third parties in Germany is available for the time before (WZB/infas 2006) and after the introduction of SC II in 2005 (Bernhard/Wolff 2008, Krug/Stephan 2016). The various studies regard different groups of unemployed people: Research before 2005 focused on unemployed people in general and hence unemployment insurance and unemployment assistance recipients. Studies conducted after 2005 look at unemployed people with four months of unemployment duration and who are declared as hard-to-place by the case workers (Krug/Stephan 2016) and also unemployed welfare benefit recipients (Bernhard/Wolff 2008). The studies in our discussion use

administrative data from the German Federal Employment Agency and are non-experimental except for the randomized field experiment by Krug and Stephan (2016).

The outsourcing of placement services was meant to increase the effectiveness of the process by competing private providers. However, the results often indicate hardly or small effects: Neither unemployed nor UB II recipients benefit to a great extent from private placement services (WZB/infas 2006, Krug/Stephan 2016, Bernhard/Wolff 2008). According to some of the results, private placement services do not help the unemployed find regular employment or maintain found contributory jobs in a better way than public placement services do (WZB/infas 2006, Krug/Stephan 2016). Krug and Stephan (2016) even find that 18 months after assignment, the accumulated days in unemployment of unemployed people assigned to public placement services are by one to two months lower than for unemployed people assigned to private providers. However, two-thirds of this reduction of time spent in unemployment for people assigned to public placement services is attributable to withdrawal from the labour market. Therefore, it is hard to estimate which form of placement services is more effective.

Bernhard and Wolff (2008) show heterogeneous results among UB II recipients. During the first four months after assignment, unemployed people assigned to private providers are less frequently employed (up to four percentage points) and more at risk of unemployment (up to ten percentage points) than comparable unemployed welfare recipients not assigned to private placement services. After four months, the effects become insignificant in East Germany. However, the treatment by private placement services positively affects the employment opportunities for West German participants by up to four percentage points. Although assignment to private providers raises the employment opportunities in the medium term, they did not reduce the participants' dependency on welfare benefit receipt (Bernhard/Wolff 2008).

Analogue to research on contracting out placement services in Germany, studies from other countries are both experimental and non-experimental. Most of them estimate the effects of contracting out placement services on the basis of micro data but there are also studies using regional data (e.g. Hasluck et al. 2003). Hasluck et al. (2003) used difference-in-difference methods to examine variations in the relative unemployment outflows between target groups and non-target groups in the British Employment Zone (EZ). Introduced in 2000, the EZ is an instrument to overcome the relatively high levels of long-term unemployment persisting in some regions in Great Britain whereas the country's labour market was characterised by a general fall in the number of unemployed people. In 15 areas the main programme for long-term unemployed adults, the New Deal 25plus, was replaced by the new instrument. EZ participants now had more freedom in their personal choices and more responsibility of entering employment on their own. Hasluck et al. (2003) analysed the

reduction in the stock of long-term unemployed in EZ and comparison areas during the first 15 months after programme start. A few months after programme start they found small positive effects because long-term unemployment fell at a faster rate in EZ than in the comparison areas. Additionally, participants who found work were less likely to re-enter unemployment if they have lived in EZs (Hasluck et al. 2003). In contrast to these quite positive effects of programme participation, evaluation studies of private placement services in Australia indicate neutral and critical evidence. Although the Australian outsourcing of placement services displays significant efficiency gains like reduced service costs and greater reach, the quality of the integration process reduced: Job-seekers still were limited in choosing private providers and the level of service decreased (Struyven/Steurs 2005).

According to a net impact analysis with Swedish data (Benmarker et al. 2013), contracting out positively influences the probability of employment only for migrants in the short run. At the same point of time participants younger than 25 years of experience negative effects. In Michigan private placement services neither improve nor deteriorate the participants' employment opportunities according to the results of Carcagno et al. (1982). Behaghel et al. (2012) examined the effect of assignment to public and private placement services with intense assistance compared to assignment to placement services without such immense support in France. While people assigned to public placement services and receiving intense assistance immediately benefit from the programme in terms of increased exit rates to employment, assignment to private providers is inefficient at the beginning of the counselling period. After six months, the assignment of unemployed people to private providers affects positively the exit rates into employment. Although these effects seem to remain stable over time, the effects of contracting out placement services are always somewhat smaller than the effects of assignment to public placement services with intense assistance during the whole observation window of twelve months after assignment. Therefore, placement services with intense assistance helped participants getting regular employment, but public services performed better than private ones (Behaghel et al. 2012).

5 Methods

Studying ALMP effects on participants requires dealing with the fundamental problem of unobservable outcomes as formalised in the Roy-Rubin model (Roy, 1951; Rubin, 1974): We can never observe at one and the same point in time the potential outcome Y of an individual i that we study in both states that we consider: with the treatment under consideration, $Y_i(1)$, and without the treatment, $Y_i(0)$. If we define D as an indicator being one for people receiving the treatment under consideration and zero for people not receiving the treatment, then we can write the effect of interest, the treatment effect on the treated as:

$$\tau_{ATT} = E[Y_i(1) - Y_i(0)|D_i = 1] = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1]$$

The second part on the left-hand side of this equation is unobservable, as we cannot observe the potential outcomes in a situation of no treatment for the treated (at a point in time after the treatment has started).

To consistently estimate the effect we will apply matching methods (Rosenbaum/Rubin, 1983, Sianesi, 2004, Caliendo/Kopeinig, 2008). The idea of these methods is the following: We find for every participant non-participants who are at risk of receiving the treatment and who in terms of observable characteristics, X , prior to the treatment are very similar to the participant. We can then compute the difference between an observed outcome of a participant and the average of the observed outcomes of the matched non-participants for each of the participants in our sample. Under certain assumptions the average of these differences over all matched pairs provides a consistent estimate of the treatment effect on the treated. We measure the outcomes from the beginning of the programme participation, which is a standard approach in the literature (e.g. Gerfin/Lechner 2002, Sianesi 2004, Wunsch/Lechner 2008, Wolff et al. 2015). However, we do not have a start month of programme participation in the sample of non-participants. Therefore, we assigned such a month to each control individual by randomly drawing it from the distribution of the three programme start months of the participant sample. In turn, we can also compute the outcome values after this (random) programme start for the potential controls.

Relying on many observed pre-treatment characteristics in our data to implement an exact matching on all these characteristics is not suitable. We follow Rosenbaum and Rubin (1983) who propose an approach to match treatment and control individuals by using a one-dimensional function of these characteristics, the propensity score, $P(X)$. In our application it will be an estimate of the individuals' probability to participate in the programme under consideration. The parameters of this probability are estimated by probit models. We estimate such models for each sub-programme and sub-sample that we consider.

That the conditional independence assumption (CIA) holds is crucial for identifying the treatment effects on the treated. It states that outcomes both in the case of treatment and non-treatment are independent of the assignment to treatment given the propensity score. In order to make the assumption hold, we use a large set of different characteristics that determine our probit models of participating in SP or IFT. The administrative data from which we drew the sample will be described in the section that follows. Here we turn to some key variables available.

Let us turn to the participation equation and the pre-treatment characteristics that we consider as covariates. The parameters of the probit participation equation are always estimated separately for four groups: men and women living in East or living in West Germany. The reason is that with respect to the assignment of the programmes there are

considerable gender differences and differences between the labour markets of the two regions (in particular between the regional unemployment rates, but also the composition of the unemployed). This can be controlled for by regarding these four groups separately. As pre-treatment covariates we consider a large number of variables. Among them are the *socio-demographic characteristics* as a first set of covariates: age, schooling degree and occupational qualification, nationality, disability status, whether a person lives with his/her partner, number of own children by age of the children, family status and a classification of the person (e.g., if the person lives with a partner with children aged less than 18 years or with children already aged at least 18 years, etc.) and the number of adult household members for different age-groups.

As a second set of covariates we include very detailed information on the *past performance of the individuals in the labour market*. In particular we include information on the last job: occupational status, last daily real wage⁴, type of last job (contributory employment, vocational training and minor employment), sector, and the time since the end of the last regular job. Moreover, we control for the cumulated duration of different labour force states and of benefit receipt during the last five years (regular employment, minor employment, selected active labour market programmes and UB II) and some further details. Moreover, we control for welfare benefit sanctions within the last year. A third set of covariates for people living with their partner, we control in the programme participation equation for *important characteristics of the partner*. As one outcome in our analysis is welfare receipt, it is important to control for such variables, as the probability of exiting from welfare receipt, a means-tested benefit for which household income matters, the partner's job finding probability and potential earnings play an important role. In turn, similar to the individual's covariates we control for sociodemographic characteristics and characteristics of the last job and labour market history of the partner of the individual, but in somewhat less detailed way.

We include for a fourth set of covariates: *real equivalent income* from welfare benefit and real equivalent income from other income sources than welfare benefit as well as the *types of income* (e.g. from earnings or other benefits that at least one household member received) that were available to a household in December 2009. These variables characterise to what extent the welfare recipient's household is still dependent on welfare benefit, as in principle the welfare benefit is paid as an addition to already available income, so that the household members can reach the legal minimum standard of living. Finally, as a last set of determinants of the participation equation, we include *district level labour market*

⁴ We compute all real variables in our analysis by dividing them with the consumer price index which is normalized to 1 in the year 2010.

indicators: the unemployment rate, long-term unemployment rate, rate of welfare recipients and the inflow rates into SP and IFT.

Apart from the CIA to hold, we need to ensure that there is common support, i.e. the probability of participating in SP or in IFT (our propensity scores) of participants and non-participants in our sample is larger than zero and lower than one. Moreover, the participation probabilities of the participants and the non-participants, from which we chose matched controls, should overlap. In practice that means that for all observed estimates of the participation probability of the participants we should find non-participants in the control sample with similar estimates of the participation probability. To guarantee that this is the case we compare the distribution of the participation probability of a participant group with the distribution of participation probability of its potential control group.

To assure that there are no differences in observables after matching between the treated and matched controls, we proceed as follows. We check by t-tests that the means of the determinants of the participation equation of the treated and of the matched controls do not differ significantly. Moreover, we use as summary measures the mean standardized absolute bias and Pseudo- R^2 of the participation equation estimated with a sample of matched pairs (see, e.g. Caliendo/Kopeinig, 2008), that help to determine whether some selection on unobservables might still lead to inconsistent results.

We estimate the results using different matching algorithms and use the results on mean standardized absolute bias to decide which results we present in our paper. First of all, we apply three algorithms of nearest neighbour matching with one neighbour and replacement. The first of these used no calliper. For the other two we computed callipers. By specifying a calliper, the nearest neighbours that the matching algorithm chooses from the control sample for a participant in absolute terms cannot deviate in their propensity score from the one of the participant by more than the calliper. Hence, callipers can be specified to ensure that no matched pairs will exist with a large deviation of their propensity scores. To determine two reasonable calliper values we proceed as follows. We carry out the nearest neighbour one-to-one matching with replacement. We use the results distribution of the absolute difference of the propensity scores of the resulting matched pairs to determine its 90th and 99th percentile. These two percentile values are then used as callipers when we apply algorithms of nearest neighbour one-to-one matching with replacement and with a calliper. Moreover, we use or plan to use these two values as callipers in other applications with callipers. We also estimate the average treatment effects on the treated by nearest neighbour one-to-five matching algorithms with replacement without and with the two callipers. Regarding the results of all these algorithms that we have run so far, the preferable results according the mean standardised absolute bias are achieved by nearest neighbour one-to-five matching with replacement and without a calliper. So we present these results.

We additionally plan to run two radius-calliper matching algorithms with same callipers that we considered for the other algorithms to check for the robustness of our results.

Our analysis relies on the large set of pre-treatment characteristics and in particular many characteristics on the labour market history and the partner's skills and labour market history. These covariates are very likely highly correlated with unobserved characteristics like individual talents and individual motivation to search for work. Therefore, if our matching approach achieves a balancing of the covariates, it is likely that we also assured that the treated and their matched controls do not differ with respect to important unobservable determinants of the propensity score and of the outcomes regarded in our analysis. For some variable we also carried out difference-in-difference matching, in particular for real earnings, the cumulated days in regular employment and the cumulated days in welfare receipt in each of the first years after programme start.

6 Data

For our analyses we use rich administrative data from the German Federal Employment Agency. The data contain information about (registered) German job-seekers and benefit recipients, including their employment and unemployment histories, ALMP participation, and some sociodemographic characteristics. The inclusion of household identifiers allows us to determine which individuals belong to each welfare recipient's household so that we can include information on welfare benefit payments by type, monthly earnings, and unearned income (i.e. other benefits, alimony payments, rents or capital income). Further, the data provide information about children, a person's partner and the partner's labour market history. This rich data basis allows us to control for a large variety of pre-treatment characteristics in the selection equations as described in the previous section.

6.1 The Sample

Our sample was drawn as follows: Individuals belong to our sample if they were registered unemployed and received welfare benefit on 31 December 2009. However, there were three reasons for excluding specific unemployed people. First, the unemployed registered with job centres that are run entirely by municipalities ("zugelassene kommunale Träger", Approved Local Providers - ALP) and not jointly by municipalities and the Federal Employment Agency ("gemeinsame Einrichtungen", Joint Local Agencies - JLA)⁵ were excluded from the

⁵ Before 2005, "social assistance" welfare benefits were administered solely by the local communities, placement services and "unemployment assistance" solely by local branches of the Federal Employment Agency. 2005's consolidation of both "assistance" types into one welfare benefit brought about two types of institutional arrangements: in most jurisdictions, "Joint Local Agencies" (joint between local community and Federal Employment Agency) provide benefit and placement services to welfare recipients. In a limited number of jurisdictions, the local community alone provides these services as "Approved Local Providers". The simultaneous existence of both arrangements served as a policy experiment to determine which model is better (Holzner and Munz, 2013: 85).

analysis⁶. ALP differ from JLA in data quality and in the detail of the characteristics concerning the labour market. In particular in the first three years after the Hartz IV reform JLA data on ALMP participation and welfare receipt information was partly characterised by reporting errors. Second, individuals were excluded from the analysis if they were participating in one of the labour market programmes of interest at the end of December 2009. Third, we considered only people aged 17 to 61 years. Older welfare recipients might retire during our observation window of 45 months after programme start. People aged younger than 17 would have been represented in our sample with a very low number of 351 people and an extremely low share of these people being assigned to the programmes under consideration. In turn this specific group in our sample was too small to be studied.

Our two treatment groups were composed of the total inflow into SAI from January to March 2010 of people who were unemployed and on welfare at the end of December 2009. The potential controls represent a 20 per cent random sample of unemployed welfare benefit recipients at the end of December 2009. They did not participate in any SAI between January and March 2010 but they may have entered other ALMPs in this period of time.

For the two waiting groups, we computed a hypothetical programme start month that was randomly drawn from the distribution of programme start months of the treatment groups, making it possible to compute outcomes from the month of programme start onwards for all individuals in our sample. People were excluded from the analysis if they already found contributory jobs or exited from welfare receipt between the end of December 2009 and their hypothetical programme start month. We created two samples from the remaining observations. In the sample for SP, all individuals participating in IFT between January and March 2010 were excluded. In our analysis of IFT, people participating in SP are not part of the sample.

The size of the treatment groups ranges from about 10,000 (East German women) up to more than 35,000 (West German men) for SP, and from less than 4,000 (East German women) up to approximately 8,000 (West German men) in IFT. For the eight main groups the relevant control group being at least 2.9 times as large as the treated group, the propensity score matching procedure should find a considerable amount of comparable controls for each treated person (Table 3).⁷

⁶ In December 2009 and therefore the month of our sampling date, 12.8 per cent of the stock of unemployed people receiving UB II were registered unemployed with ALP (Department of Statistics of the Federal Employment Agency, calculation from the Data Warehouse).

⁷ The size of the control group sample might differ for the two programmes for a number of reasons. First, we eliminated individuals who exited from welfare or had a transition into contributory job between the end of December 2009 and their random programme start from the control group. One reason for the differences is hence a different distribution of the random programme starts of SP and IFT during the first three months of the year 2010. And second, in the sample of potential controls for SP, all individuals participating in IFT during the first three months of 2010 were excluded. The same holds for the sample of potential controls for IFT if individuals participated in SP during these three

Table 3: Sample size of treated and potential controls

	SP - schemes by providers				IFT - in-firm training			
	East Germany men	West Germany women	East Germany men	West Germany women	East Germany men	West Germany women	East Germany men	West Germany women
number of treated	13,583	10,329	35,404	24,157	4,552	3,623	8,095	4,646
number of potential controls	58,430	46,596	102,616	88,593	53,071	44,621	88,159	84,852
ratio between controls and treated	4.3	4.5	2.9	3.7	11.7	12.3	10.9	18.3
samples by age								
<i>aged 17-<25</i>								
number of treated	2,348	1,721	4,582	3,197	448	382	716	357
number of potential controls	3,971	3,364	6,644	5,572	1,613	1,096	3,197	1,965
ratio between controls and treated	1.7	2.0	1.5	1.7	3.6	2.9	4.5	5.5
<i>aged 25-<50</i>								
number of treated	8,708	6,653	22,374	16,548	3,035	2,152	5,692	2,838
number of potential controls	36,776	29,638	63,847	59,862	32,191	23,657	55,519	44,968
ratio between controls and treated	4.2	4.5	2.9	3.6	10.6	11.0	9.8	15.8
<i>aged 50 and more</i>								
number of treated	1,818	1,675	4,832	3,610	465	424	851	544
number of potential controls	14,175	12,038	22,280	19,751	13,356	10,809	20,323	15,859
ratio between controls and treated	7.8	7.2	4.6	5.5	28.7	25.5	23.9	29.2
samples by time since last job								
<i>never worked since 2000</i>								
number of treated	3,226	3,950	6,212	8,480	486	687	856	975
number of potential controls	16,118	19,493	22,377	36,348	14,704	18,540	19,371	34,377
ratio between controls and treated	5.0	4.9	3.6	4.3	30.3	27.0	22.6	35.3
<i>up to one year</i>								
number of treated	2,668	1,518	8,580	4,166	1,641	1,117	2,912	1,381
number of potential controls	8,427	5,054	17,797	9,987	7,444	4,824	14,706	9,480
ratio between controls and treated	3.2	3.3	2.1	2.4	4.5	4.3	5.1	6.9
<i>more than one year up to five years</i>								
number of treated	4,389	2,444	13,333	6,172	1,807	1,222	3,313	1,533
number of potential controls	17,147	10,439	36,156	21,117	15,478	10,005	31,108	20,387
ratio between controls and treated	3.9	4.3	2.7	3.4	8.6	8.2	9.4	13.3
<i>more than five years</i>								
number of treated	3,300	2,417	7,279	5,339	618	597	1,014	757
number of potential controls	16,738	11,610	26,286	21,141	15,445	11,252	22,974	20,608
ratio between controls and treated	5.1	4.8	3.6	4.0	25.0	18.8	22.7	27.2

6.2 Selected characteristics of the sample members

We present some selected descriptive statistics to show the differences in the observable characteristics among the treatment groups and control groups. Not surprisingly and according to the job centres' incentives from special rules of placing especially welfare benefit recipients younger than 25 years into ALMPs, the share of younger recipients is in

months. As there are more people participating in SP than in IFT, the sample of potential controls is smaller for IFT than SP.

many cases more than twice as high in both treatment groups as the corresponding share in the control groups (see Table 4). These differences are larger in East Germany than in West Germany and also larger for IFT than for SP. This could reflect the use of IFT as an instrument to in some cases test the younger welfare recipients' willingness and capability to work. In other cases it is used as sort of unpaid internship. In contrast to younger welfare recipients, people aged more than 50 years are not a specific target group of SAI. While about one quarter of the control group is older than 50 years, only about 15 to 17 per cent of the participants in SP and about eleven to more than twelve per cent in IFT are of the same age. The shares of participation in SP and IFT for this older age group do not differ considerably between East and West Germany but are slightly higher for older women than older men.

Further target groups defined by the Federal Employment Agency are persons with health restrictions, foreigners, and persons without secondary schooling degree. Neither persons with disability status⁸ nor foreigners are strongly targeted by any of the training measures. Both for men and women and in both regions the share of people with disability status is lower than their share in the control group. In East Germany, for SP this difference is highest as the share of people with disability status in the treatment group is less than half as high than in the control group. The shares displayed in the table imply that regarding IFT, men and women do not differ in their inflow rates. However, people with disability status in East Germany have smaller chances to participate in IFT than disabled people in West Germany. Compared with Germans it is less likely for foreigners to participate in SP. A lower share of foreigners among participants than among the controls demonstrate this in particular for East Germany and with one exception: in West Germany for foreign men the shares imply that they were assigned more often to SP than men of German nationality. Furthermore, the displayed shares imply that foreigners are even less represented among the IFT than among SP participants. This is also true for people without secondary schooling degree compared to groups with higher schooling, especially for women. Compared with their share in the control group, females without a schooling degree are less represented in training measures, especially regarding IFT participation: While the difference in the participation share between treatment and control females is 2.6 to 4.8 percentage points for SP, the difference between IFT treatment and control females is 6.8 to 14.2 percentage points. The share of males without a secondary schooling degree does not differ between the control group and the SP treatment group but the IFT group.

About 66 to 70 per cent of controls do not have a partner. This share is higher for most of the programme participants (except West German males in SP). Especially female

⁸ As our data are not informative about health restrictions in general, we take the disability status into account, which is available.

participants do not have a partner: the share of females without a partner and participating in IFT is nine to 15 percentage points higher than their share in the waiting group. Women and West German males without a partner and without children are overrepresented in both treatment groups. Especially the share of single and childless females in West Germany is up to 16 percentage points higher in the IFT treatment group than in the potential control group.

Schemes for activation and integration do not focus on hard to place individuals among UB II recipients according to our results. IFT neglects UB II recipients who were never employed during the last five years even more than SP training measures do. The share of people in the waiting group who were never employed in an unsubsidized contributory job between 2005 and 2009 is around 63 per cent for females, about 54 per cent for males in East Germany and approximately 46 per cent for males in West Germany. In contrast, these shares are between three to eight percentage points lower in the SP treatment group while they are between 22 to 30 percentage points lower in IFT participant group. Looking at the different types of SAI, we find IFT participants to have a mean cumulated duration of employment periods of nearly 400 days and therefore about 160 days longer cumulated duration of employment periods than participants in SP. In the waiting group, the mean cumulated duration of employment periods is approximately 190 days. This is a hint that schemes for activation and integration, and especially in-firm training, focus more on needy unemployed with good perspectives on the labour market than hard to place individuals.

The average (equivalent) benefit levels in December 2009 differ only slightly between the control group and the treatment groups. They range from about 644 to 751 euros per months in the waiting group. The IFT treatment group receives on average less benefit payments than the SP treatment group. This is a hint that the households of IFT participants are less dependent on welfare because they already have more income from other sources than welfare than the households in the potential control group. However, the benefit level is lower in East Germany than in West Germany, which may be related to lower rents in East as opposed to West Germany, and higher for women than for men.

Table 4: Selected descriptive statistics of treatment and control group by gender and region (in %)

	Control (waiting)				SP – schemes by providers				IFT – in-firm training			
	East Germany		West Germany		East Germany		West Germany		East Germany		West Germany	
	men	women	men	women	men	women	men	women	men	women	men	women
age 17-19	1.2	1.4	1.7	1.6	2.7	2.9	3.3	3.7	1.3	1.6	2.8	4.1
age 20-24	6.9	7.0	6.2	5.7	15.2	14.0	10.2	9.8	16.0	18.7	11.6	12.8
age 25-29	14.6	13.4	13.1	12.0	18.2	15.3	15.1	13.6	26.0	20.3	21.1	18.4
age 30-34	12.3	11.1	13.0	13.2	13.2	12.2	14.1	13.9	16.4	12.5	16.0	13.8
age 35-39	10.9	10.9	12.7	14.1	10.8	11.1	13.5	13.9	10.4	10.9	13.7	12.3
age 40-44	13.3	13.7	14.8	15.5	12.2	13.3	14.5	15.2	10.4	11.7	12.8	13.6
age 45-49	15.4	15.7	15.0	14.5	13.2	14.2	13.6	13.8	8.8	11.8	10.9	12.5
age 50-54	13.1	13.6	12.8	12.1	9.6	11.1	9.9	10.1	7.2	8.3	7.4	8.6
age 55-61	12.2	13.2	10.8	11.4	4.9	6.1	5.8	5.9	3.6	4.2	3.7	3.9
disability	4.1	3.3	5.1	3.7	2.0	1.5	3.9	2.9	2.6	2.1	3.1	3.0
Foreign nationality	9.1	9.4	21.8	23.4	5.6	5.1	24.6	22.1	1.1	3.9	19.3	12.3
no secondary schooling degree	10.6	9.4	15.8	19.8	10.4	6.8	15.0	15.2	5.4	2.6	9.3	5.6
no partner	70.3	65.9	67.4	67.6	75.2	69.1	67.0	72.6	74.2	75.2	71.6	83.1
no child	82.4	61.1	76.6	53.1	83.4	58.7	74.5	54.0	82.5	63.0	78.7	63.4
no partner, no child	68.4	42.8	65.3	39.6	73.3	44.0	65.1	42.9	72.3	51.3	69.9	55.5
never regularly employed for some time in the period 01/2005-12/2009	54.2	64.6	45.7	62.5	47.3	61.0	37.7	56.6	23.9	34.9	22.8	36.8
mean duration of cumulated employment in the period 01/2005 -12/2009 (in days)	193	155	246	161	234	180	298	195	421	378	426	340
equivalent UB II in December 2009 (in euros)	644	703	684	751	628	697	677	746	631	684	670	723

7 Results and discussion

7.1 Main effects on regular employment and welfare receipt probabilities

Regular employment effects

Table 5: Average participation effects on holding a regular unsubsidized job (in percentage points)¹⁾

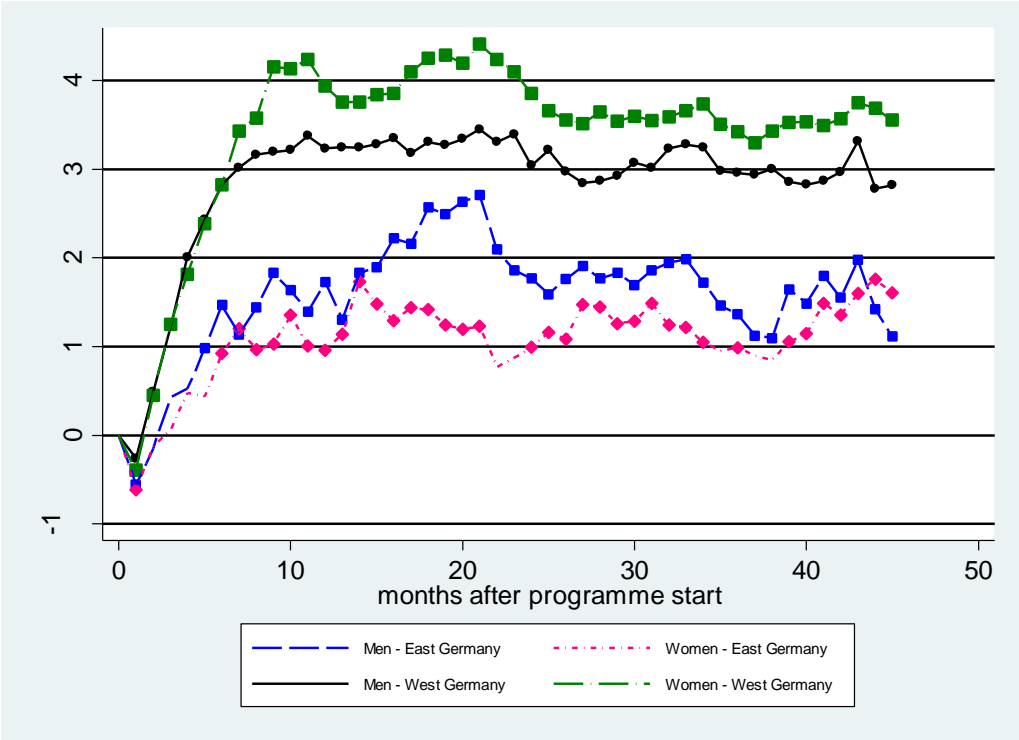
Months after programme start	SP – schemes by providers								IFT – in-firm training							
	East Germany				West Germany				East Germany				West Germany			
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women		
1	-0.56	***	-0.62	***	-0.26	**	-0.39	***	11.37	***	10.53	***	9.23	***	10.53	***
6	1.47	***	0.93	**	2.82	***	2.83	***	18.06	***	17.91	***	15.84	***	16.16	***
12	1.73	***	0.96	*	3.23	***	3.94	***	19.53	***	23.91	***	17.37	***	20.22	***
20	2.63	***	1.20	*	3.34	***	4.19	***	19.97	***	21.59	***	16.49	***	17.69	***
24	1.77	***	0.99	*	3.04	***	3.85	***	18.67	***	21.15	***	15.22	***	16.41	***
36	1.36	**	0.98	*	2.96	***	3.42	***	16.68	***	18.11	***	13.73	***	15.13	***
45	1.11	*	1.60	**	2.82	***	3.55	***	15.67	***	17.87	***	13.31	***	14.17	***

¹⁾ Statistical significance on the 0.1%, 1% and 5 % levels is denoted by ***, **, *.

Table 5 displays our main results on the effects of programme participation on working in a regular job at different months after programme start; all the displayed estimates are well determined. Both SP and IFT increase the probability of holding a regular unsubsidized job, with effects being much stronger for IFT participants than for SP participants. While IFT's effect jumps up as early as one month after programme start (Figure 1b), SP exhibits small lock-in effects of up to 0.6 percentage points (Figure 1a). They disappear soon afterwards, leading to positive effects for both West and East Germans two to three months after programme start (not shown in Table 5). SP effects reach their peak between the first and second year after programme start but remain on a similarly high level as late as three years after programme start. At that point in time, West German male SP participants profit from a positive employment effect of almost three percentage points, or 12.1 per cent higher than the employment probability in the absence of treatment (24.4 per cent among their matched controls). In East Germany after three years, the effects for men are 1.4 percentage points and a 6.4 per cent increase of the employment probability in the absence of treatment (21.3 per cent among their matched controls). During the entire observation period of up to 45 months after programme start, it becomes clear that West German women in particular profit from SP participation. These measures might serve as a stepping stone into the labour market after a long absence due to child-rearing and care activities.

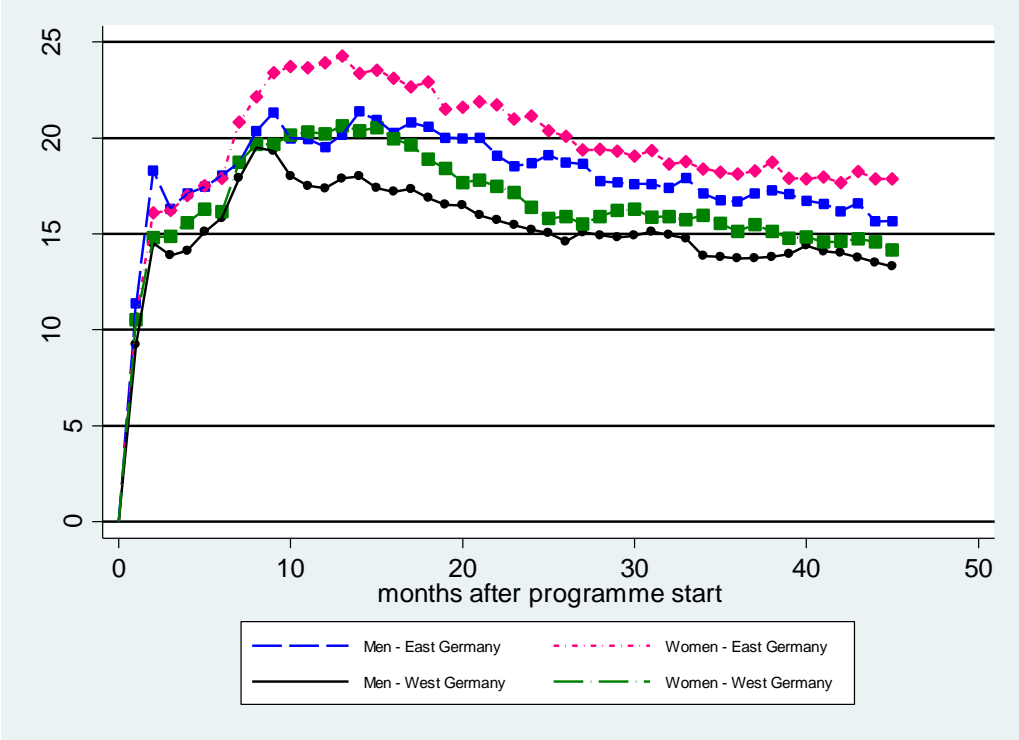
IFT participation exhibits far larger effects on the chance of having regular employment. IFT unfolds its strongest effect during the first year after programme start and decreases afterwards while staying high: three years after programme start, West German women and men show a 15.2 and 16.4 percentage points higher probability of being regularly employed than their respective control groups. Given an employment probability among the matched controls of 28.7 and 31.7 per cent, this amounts to an increase by 52.7 and 43.3 per cent that is a result of participation. Among East German men, the employment effect amounts to 16.7 percentage points three years after programme start, or a 53.8 per cent increase given an employment probability of 31 per cent among the statistical twins. The effect is even higher among East German women, amount to 18.1 percentage points or an increase of 61.8 per cent compared to the control group's probability of 29.3 per cent. In our observation window of up to 45 months after programme start these effects, are quite stable after three years.

Figure 1a: Estimates of the effect of SP on regular employment probability¹⁾



¹⁾ Squares indicate effects significant on a level of at least 5%. Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

Figure 1b: Estimates of the effect of IFT on regular employment probability¹⁾

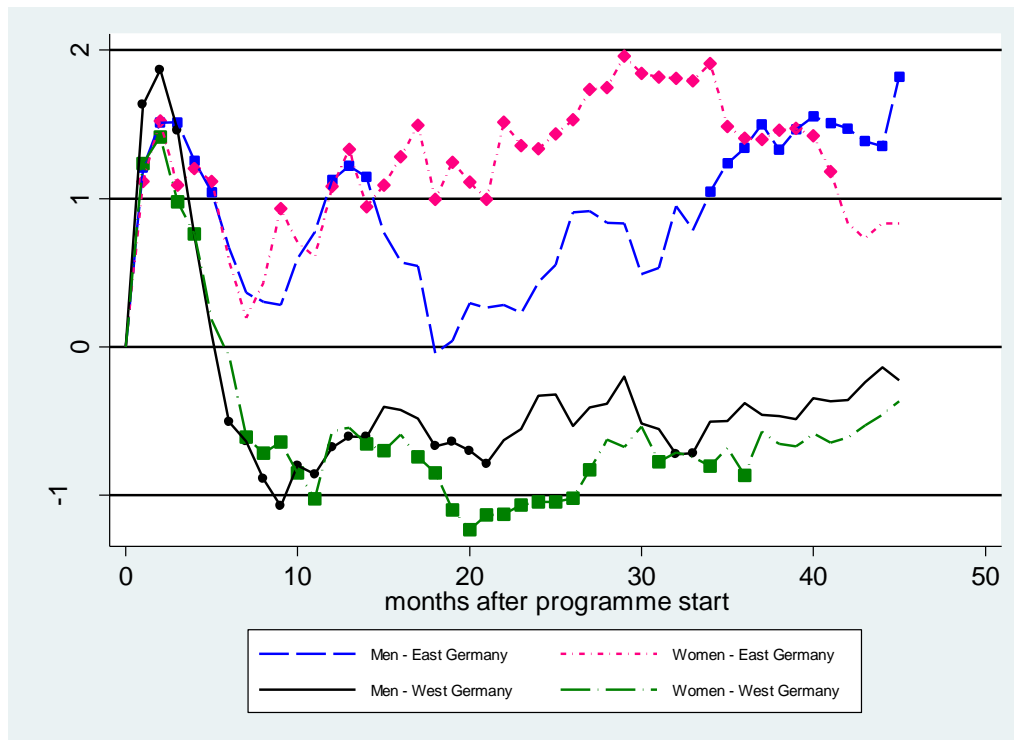


¹⁾ Squares indicate effects significant on a level of at least 5%. Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

Effects on welfare benefit receipt

These observed effects on regular unsubsidized employment can in part be transferred to the probability of leaving welfare receipt. The following figures show how the estimated effects on UB II receipt probability develop over time, for SP participants as well as for IFT participants. Each figure depicts the differences between the four main groups of male and female participants in West as well as in East Germany.

Figure 2a: Estimates of the effect of SP on UB II receipt¹⁾



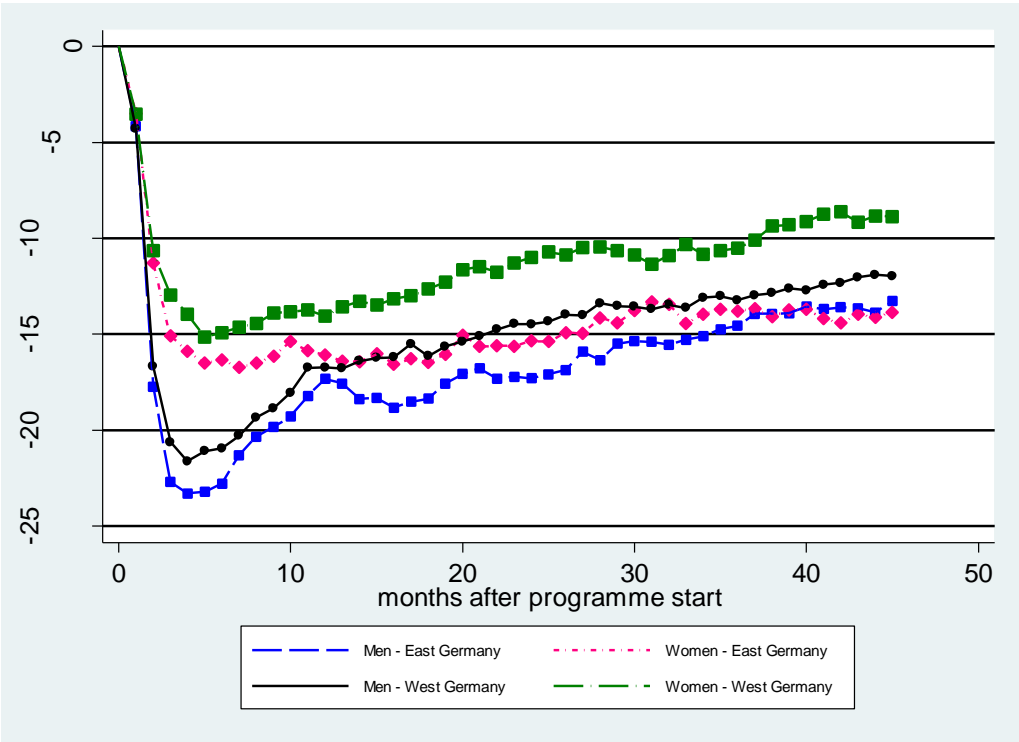
¹⁾ Squares indicate effects significant on a level of at least 5%.

Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

As with the effects on regular employment, SP participation causes a small lock-in effect on welfare receipt in all four main groups (e.g. two months after programme start of a magnitude of 1.4 to 1.9 percentage points). Up to five months after programme start, participants are slightly more likely to receive UB II than matched non-participants (see Figure 2a). Starting in the sixth month, regional and, starting in the twelfth month, gender differences become apparent. While East German participants are generally more likely to receive UB II than their control group counterparts, this probability somewhat decreases among West Germans. About one year after programme start, effects among East and West German men lose their statistical significance, leaving no observable effects for male participants and non-participants. Only three years after programme can small differences between East and West German men be seen. While in the West, participants' outcomes continue to be indistinguishable from those of matched non-participants (UB II recipients amounting to 60.9

per cent in the control group at that point in time), East German SP participants are 1.3 percentage points more likely to receive UB II (the control group share being 67.6 per cent three years after programme start). The respective trends continue among women. While SP participation raises the probability of receiving UB II of East German female participants three years after programme start (UB II share of their matched control group: about 73 per cent), it is 0.9 percentage points lower for West German women than in the control group (UB II share of their matched control group: 66.8 per cent). This demonstrates one thing: although SP participation increases the probability of having a regular job, it does not, on average, help participants leave the benefit receipt. There may be several causes for this: one might be that due to the participation fewer individuals leave UB II receipt without having taken up regular employment, e.g. because fewer individuals leave welfare and probably stop by moving to family members who might support them. It also could be due to participants taking up part-time or low-paid employment that generates an income that is insufficient to cover the cost of living.

Figure 2b: Estimates of the effect of IFT on UB II receipt¹⁾



¹⁾ Squares indicate effects significant on a level of at least 5%. Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

Figure 2b shows that IFT participation decreases the UB II benefit receipt probability in all four subsamples throughout the entire observed time period. The effect is strongest for East German men, whereas West German women are comparatively unlikely to leave the benefit receipt. The effect reaches its maximum six months after programme start (seven months

among East German women) and becomes weaker afterwards. After three years, participants are still 10.5 to 14.6 percentage points more likely than comparable non-participants to not receive welfare benefits. The share of UB II recipients is between 51.8 and 61.9 per cent in the control groups.

Effects on real earnings

The real earned annual income of SP participants is positively affected by their participation (maximum effect at about 136 to 493 euros in the second year after programme start; see Table 6). These effects are far lower than the effects of IFT participation on IFT participants. IFT participation increases participants' real earned income by a greater magnitude; the effect reaches its maximum already in the first year after programme start with 3,119 to 4,281 euros (no figure given), indicating that many IFT participants are kept as regular workers by the establishments which conducted the programme (see Wolff/Jozwiak 2007).

The effects of both programme types somewhat decrease over time, yet stay clearly positive throughout the observed time period. The mechanism behind their magnitudes differs between the programme types. The findings of SP effects show stronger regional differences, as for West Germans participant the effects exceed considerably those for East German participants (effects of 428/476 euros in West versus 129.7/199.7 euros in East Germany among men/women, compared to comparison group in the third year after programme start). IFT on the other hand exhibits gender differences at the same point in time. Participating men in East/West Germany earn 2,940/2,952 euros more than their comparison groups; the effects are also greater than those of East/West German women (2,354/2,247 euros). This runs contrary to the effects on the probability of finding regular employment: IFT participation positively and more strongly affects the employment probability of female participants than of male participants, even as our results imply higher earnings effects for male as opposed to female participants. One reason for this could be that women more often work part-time and/or at low wages. Another possible cause may lie in gender-specific occupational segregation, meaning that women more often work in occupations that are paid badly such as child and elderly care or cleaning professions (Hausmann/Kleinert 2014).

Table 6: Average participation effects on real annual earned income (in euros)¹⁾

Year after programme start	East Germany		West Germany	
	Men	Women	Men	Women
SP - schemes by providers				
1 st	114 **	43	406 ***	337 ***
2 nd	236 ***	136 *	484 ***	493 ***
3 rd	200 **	130 *	476 ***	428 ***
IFT - in-firm training				
1 st	4,104 ***	3,324 ***	4,281 ***	3,119 ***
2 nd	3,324 ***	2,692 ***	3,238 ***	2,548 ***
3 rd	2,940 ***	2,354 ***	2,952 ***	2,247 ***

¹⁾ Statistical significance on the 0.1%, 1%, and 5% levels is denoted by ***, **, *.

Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

As was mentioned earlier, SAI draw on the former training measures, leading to some similarities between SP and IFT and out-of-firm and in-firm training measures. As a consequence, our results mirror similar findings of both instrument types regarding the effects on regular employment. Wolff and Jozwiak (2007) also found lock-in effects for in-school training measures which, at four to five months, last longer than those of SP investigated here. The effect magnitudes are similar as well. In a similar fashion, the authors found distinctly larger effects for in-firm than for in-school training measures, both on the regular employment probability and the probability of leaving the welfare benefit receipt (Wolff/Jozwiak 2007). SP however also replaced the predecessor programme of bringing in private placement services. Their effects were found to be of a similar magnitude (in the context of SC II, see Bernhard/Wolff 2008) to the SP effects presented here.

7.2 Effects by age

German law mandates particularly intense activation for persons below the age of 25 while in practice allowing for reduced search and activation efforts of older unemployed persons (in particular among those aged 58 and older). The chances of finding regular employment are commonly found to decline with age, in part due to diminishing expected returns on human capital investment. On the other hand, higher age will necessarily be correlated with more labour market experience, but also longer non-employment duration. Different age groups may also be subject to different kinds of specific interventions within the SP/IFT policy framework, which may further contribute to heterogeneity in outcomes.

We split each of the four main samples into three subsamples each according to age on December 31, 2009: 17–24 years (“young”), 25–49 years (“middle aged”), and 50-61 years (“old”). These thresholds reflect those explicitly specified in the German Social Codes II and III; in particular, the threshold at 50 years triggers eligibility for special kinds of subsidies.

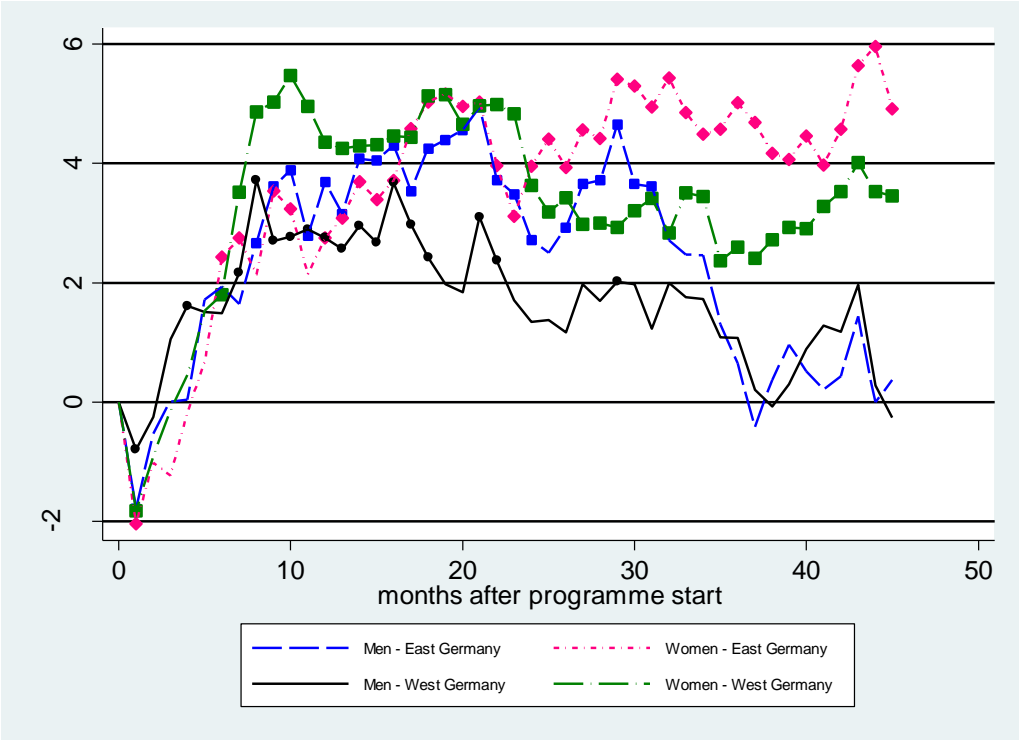
Another important threshold at age 58 was not considered due to the small number of observations above it. It would be important because persons aged 58 years and older could receive unemployment benefits without actively searching for work (Article 65 of SC II).

According to our estimates SP participation has positive effects on the chances of holding a regular job in all age groups. The youngest age group is characterised by a slightly larger effect of three to five percentage points than in the other age groups, an effect that holds throughout the observation period for women but collapses among men 36 months after programme start (Figure 3a). For young women therefore a more sustained improvement of their employment prospects is achieved due to SP participation than for young men. The West-East difference is not very pronounced.

The middle age group exhibits an entirely different pattern similar to that of the main samples: a very pronounced difference between West and East Germany, with West German men and women improving their employment prospects by a maximum of four to 4.5 percentage points (at different points in time, see Figure 3b), and East Germans experiencing a maximum improvement of only 1.5 to slightly below three percentage points. Unlike what was seen among the youngest age group, the differences between the genders are not as pronounced, although there are local peaks after twelve months for West German women and around 20 months for East German men.

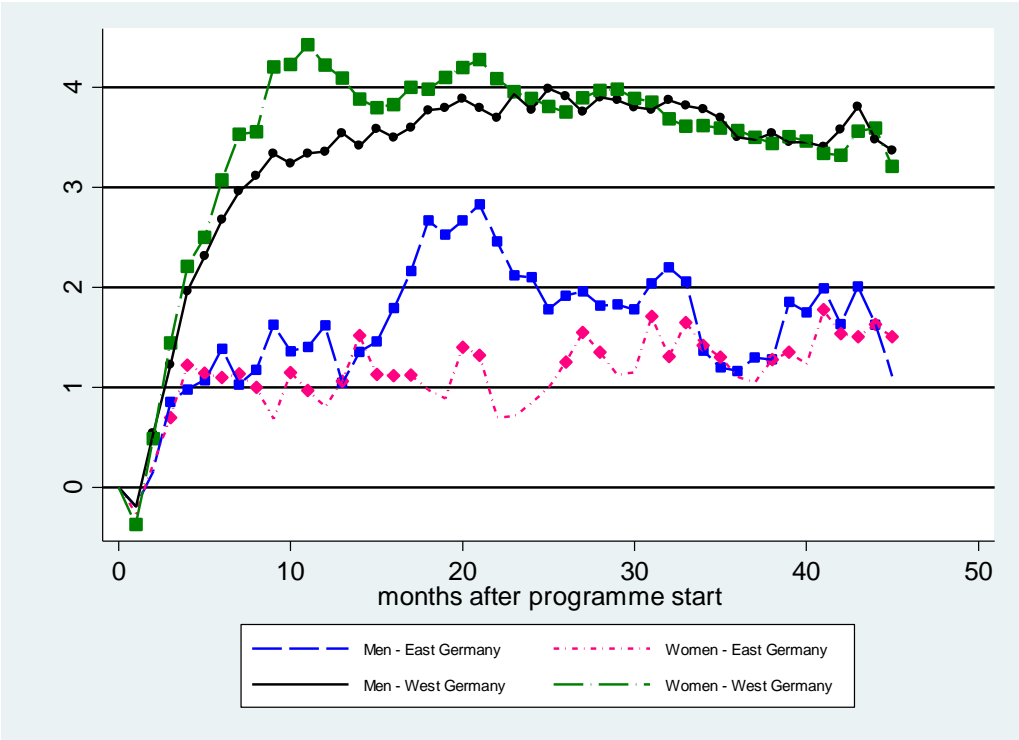
The pattern in the oldest age group is basically similar to the middle age group. The most important difference is that old East German women do not improve their employment prospects at all as a result of SP participation while East German men do. Their improvement only begins after a rather long time of about 18 months after programme start, an effect that holds to the end of the observation period (Figure 3c).

Figure 3a: Estimates of the effect of SP on regular employment - ages 17-24¹⁾



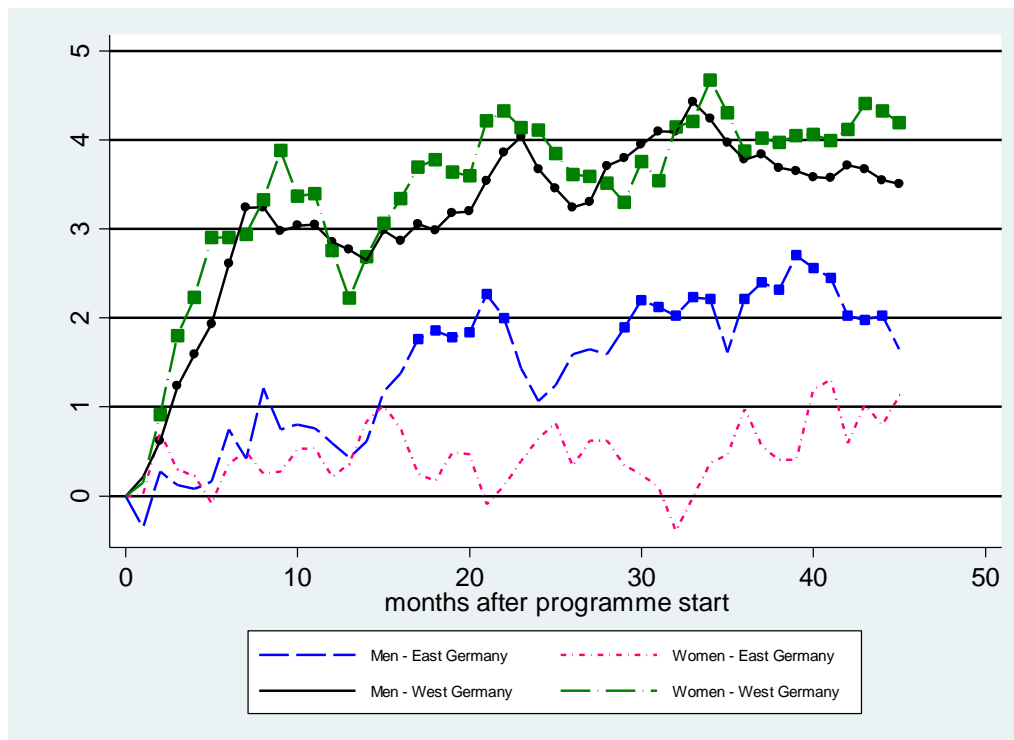
¹⁾ Squares indicate effects significant on a level of at least 5%. Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

Figure 3b: Estimates of the effect of SP on regular employment - ages 25-49¹⁾



¹⁾ Squares indicate effects significant on a level of at least 5%. Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

Figure 3c: Estimates of the effect of SP on regular employment - ages 50-61¹⁾

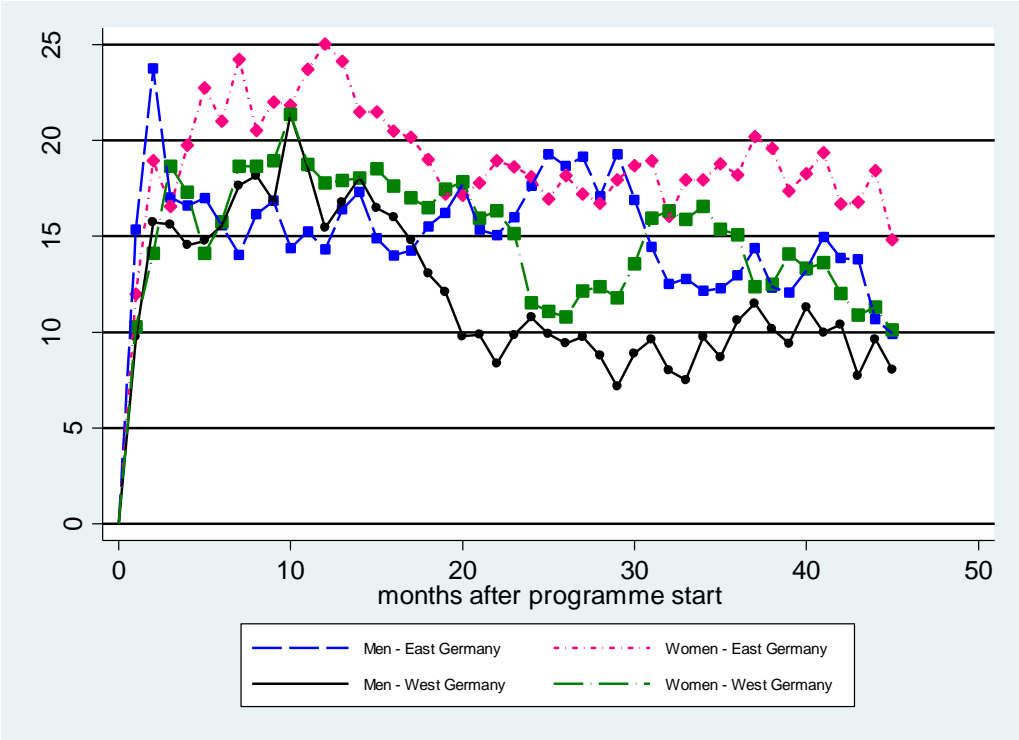


¹⁾ Squares indicate effects significant on a level of at least 5%.

Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

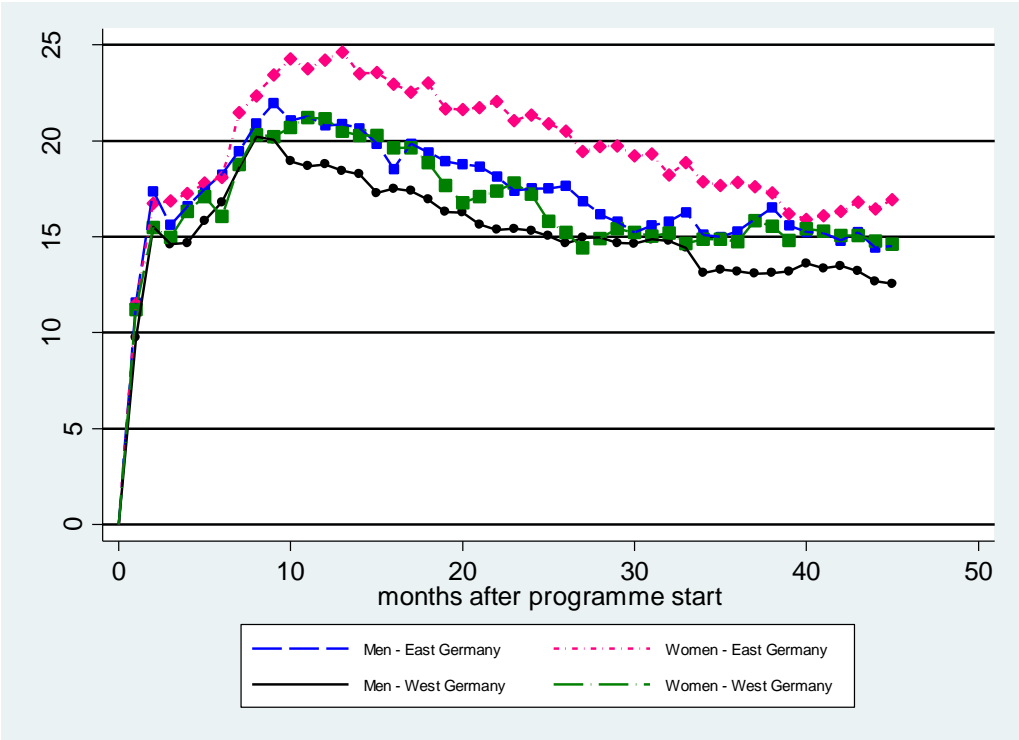
For all age groups our estimates imply higher chances of regular employment due to IFT participation. Age makes a difference however in how this effect is spread throughout the observed time period. In the youngest age group, the effect rises very quickly in the first three months after programme start to a local maximum of between 15 to 23 percentage points (Figure 4a). In the middle-aged group, the maximum is reached only after about twelve months (Figure 4b), and in the oldest age group, the increase at the twelve month threshold is very strongly pronounced (Figure 4c). This time period of the first twelve months after IFT programme start is when IFT participation is followed by additional support subsidized employment with EGZ (Eingliederungszuschuss, hiring subsidy) and EGS (Einstiegsgeld, integration allowance).

Figure 4a: Estimates of the effect of IFT on regular employment - ages 17-24¹⁾



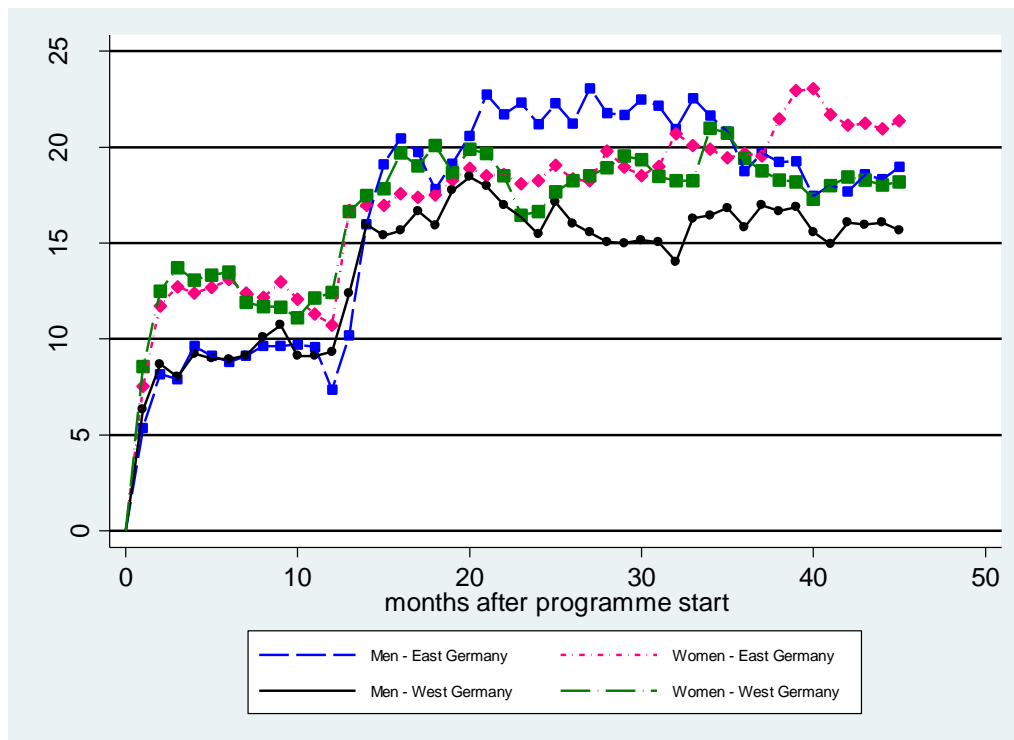
¹⁾ Squares indicate effects significant on a level of at least 5%. Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

Figure 4b: Estimates of the effect of IFT on regular employment - ages 25-49¹⁾



¹⁾ Squares indicate effects significant on a level of at least 5%. Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

Figure 4c: Estimates of the effect of IFT on regular employment - ages 50-61 ¹⁾



¹⁾ Squares indicate effects significant on a level of at least 5%.

Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

Therefore, if young people find a job after IFT participation, it is very often a regular job, whereas in older age groups, subsidized employment plays a very important bridge function into regular employment. And the subsidized employment often ends after around eleven months after IFT start for the group of 50-61 so that afterwards their regular employment rate jumps up as the employers often continue to employ them as unsubsidised employees.

7.3 Effects by time since last unsubsidized contributory job

As with age, we further split the four main samples into four subsamples each according to the number of days since a person's last employment ended and our sampling date, December 31, 2009: up to one year non-employed, more than one to five years non-employed, more than five years to ten years non-employed, and more than ten years or has never worked at all ("never worked since 2000").

Table 7 shows the effects of SP and IFT participation on the regular employment probability by the duration of non-employment. The last category, "never worked since 2000", is a special category, as it includes both those with extremely long unemployment experience as well as those who have never worked before because they have just entered the labour market, be it from child-rearing or after finishing full-time education.

West German SP participants always improve their regular employment opportunities regardless of non-employment duration, women more so than men (2.5 to 5.3 percentage

points among women versus 1.7 to 3.8 percentage points among men). The effect is less stable after 45 months for men whose last job was within the previous year.

More diverse are the SP results in East Germany: among those with a non-employment duration of at most five years, only men improve their regular employment prospects by at most 3.5 percentage points, and the effect is also only strongly statistically different from zero among those who have not worked for one to five years. Only after 36 months do women with the same non-employment duration slightly improve their employment chances by 2.8 percentage points. For those not employed for five to ten years, effects decline over time among East German women while remaining stable among East men, both starting at around two percentage points. Effects are small and unstable among those who have not worked since 2000.

Table 7: Average participation effects on holding a regular unsubsidized job (in percentage points) ¹⁾

Months after programme start	SP – schemes by providers				IFT – in-firm training			
	East Germany		West Germany		East Germany		West Germany	
	Men	Women	Men	Women	Men	Women	Men	Women
	<u>less than one year since last job</u>							
6	0.398	1.381	2.534 ***	2.718 **	15.917 ***	15.957 ***	14.844 ***	16.177 ***
12	2.901 *	2.191	3.634 ***	4.850 ***	16.319 ***	24.043 ***	16.397 ***	19.218 ***
24	2.758 *	1.368	2.977 ***	3.745 ***	15.222 ***	15.867 ***	12.499 ***	12.875 ***
45	1.571	1.368	1.851 *	3.909 *	10.360 ***	10.907 ***	8.518 ***	11.803 ***
	<u>one to five years since last job</u>							
6	3.538 ***	0.638	3.502 ***	4.283 ***	21.353 ***	20.312 ***	18.164 ***	17.359 ***
12	2.931 ***	0.597	3.638 ***	5.307 ***	21.608 ***	22.346 ***	18.828 ***	20.967 ***
24	2.813 ***	0.565	3.748 ***	4.279 ***	19.523 ***	22.494 ***	18.158 ***	18.248 ***
45	1.199	2.758 *	3.761 ***	2.918 ***	17.227 ***	19.491 ***	15.445 ***	13.922 ***
	<u>five to ten years since last job</u>							
6	1.976 ***	2.028 ***	2.737 ***	3.110 ***	15.916 ***	17.037 ***	12.261 ***	14.690 ***
12	1.267 *	1.498 *	3.223 ***	4.294 ***	19.287 ***	21.650 ***	17.552 ***	17.966 ***
24	1.867 **	1.945 *	3.297 ***	4.976 ***	20.843 ***	22.795 ***	17.512 ***	17.252 ***
45	2.031 *	1.796	3.437 ***	3.829 ***	22.010 ***	19.192 ***	16.308 ***	16.988 ***
	<u>never worked since 2000</u>							
6	0.031	1.408 ***	1.668 ***	2.535 ***	19.212 ***	16.569 ***	14.790 ***	15.426 ***
12	1.297 *	1.216 *	2.783 ***	3.283 ***	20.124 ***	24.692 ***	17.313 ***	19.856 ***
24	1.440 *	0.162	3.089 ***	3.875 ***	19.378 ***	21.466 ***	16.308 ***	16.164 ***
45	0.422	1.373 *	2.680 ***	3.545 ***	17.427 ***	22.405 ***	15.818 ***	13.518 ***

¹⁾ Statistical significance on the 0.1%, 1%, and 5% levels is denoted by ***, **, *.

Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

For IFT participants, the difference in non-employment durations has a greater impact on how stable the effect is over time rather than the maximum magnitude of the effect. Among West German men who lost their job within the last year, the initial effect on regular employment probability collapses from a maximum value of 16.4 percentage points after twelve months to only 8.5 percentage points after 45 months. For the three other non-employment groups among West German men the effects tend to be higher already six to twelve months after programme start and remain relatively stable thereafter. A similar pattern holds for the other groups. Therefore in-firm training causes a far less stable regular

employment effect for those who are already close to the labour market by having left their previous job within the last year. Directing the programme more frequently to different groups of long-term non-employed individuals could raise its effectiveness.

7.4 Match quality

As Table 8 shows, the large number of control group observations results in sufficient common support both between the IFT and SP treatment and control groups, as the number of off-support treatment group observations is always in the single digits. Comparing the distributions of the propensity scores for the control group and groups, we find sufficient mass among non-participants for regions of the propensity score with mass among participants in both treatment groups.⁹ Therefore, the assumption of overlap holds. Following a strategy of minimizing the Mean Standardized Bias, we chose a matching algorithm of five nearest neighbour matching with replacement but without a calliper. The Mean Standardized Bias after matching is always below the value of one, the post-matching adjusted R-square is close to zero, and based on t-statistics, no significant differences remains in any covariate between the treatment and matched control groups. These figures stand for an excellent match quality, resulting in any bias due to selection on observable characteristics being almost completely removed from our effect estimates. Even for the age and non-employment duration sub-groups, the mean standardized bias after matching never rises above a value of 2.5 (see Appendix: Table A1 and Table A2).

Table 8: Matching quality in the four main samples

Outcome/statistic	East Germany		West Germany	
	Men	Women	Men	Women
SP - schemes by providers				
Treatment group obs. on support	13,581	10,328	35,397	24,157
Treatment group obs. off support	2	1	7	0
Control group observations	58,430	46,596	102,616	88,593
Mean Standardized Bias after matching	0.399	0.554	0.228	0.312
Adjusted R ² after matching	0.001	0.001	0.000	0.001
IFT - in-firm training				
Treatment group obs. on support	4,551	3,619	8,088	4,645
Treatment group obs. off support	1	4	7	1
Control group observations	53,071	44,621	88,159	84,852
Mean Standardized Bias after matching	0.631	0.829	0.421	0.519
Adjusted R ² after matching	0.002	0.002	0.001	0.001

Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

⁹ We did not display the figures that demonstrate this for the main groups of male and female participants in East and in West Germany. They are available on request.

8 Summary and Conclusions

Short term training and contracting out placement services to private providers were programmes that were highly regulated in Germany. They were replaced by the SAI in the year 2009, which introduced an opportunity for the PES to flexibly design such tools. Policy makers had the intention to create a programme that can be easily designed such that it helps to overcome individual employment impediments of welfare recipients. Therefore, we analysed whether this new programme is effective in terms of enhancing the employment prospects of unemployed people who receive welfare benefits who compared with unemployment insurance recipients are the most vulnerable group of unemployed people. A key question in this context is whether the more flexible design of these programmes implies higher effects than participation in the former programmes of short term training and of contracting out placement services.

We studied a large sample of unemployed welfare recipients at the end of the year 2009 that was drawn from administrative records. We estimated the effects for participants from this sample who started their participation in one of two different types of SAI during the first quarter of 2010. For SP participants our estimates imply the following: We generally find that lock-in effects are negligible. SP participation enhances the employment prospects of participants raising their employment rates by up to more than four percentage points with lower effects in East than in West Germany. Often we find that the employment effects decline somewhat after a peak, but are still significant at the end of our observation window of 45 months after programme start. By and large these implications also hold when we separate our findings by different age-groups and for different groups of participants according to the time since their last job ended. There is no convincing evidence that SP participation effects on the employment outcome generally increase or decrease with the age of participants. When we regarded the four non-employment groups of people non-employed for less than one year, one up to five years, five up to ten years and more than ten years or never worked in an unsubsidized contributory job, the results tend to be highest for long-term non-employed people with more than one up to five years of non-employment.

Our results for the outcome UB II (welfare) receipt suggest that SP participations are generally lower than the employment effects and for East German participants sometimes even negative of an order of up to -1 percentage point. We would not interpret this as a hint for an ineffective programme. One reason for this result might be that due to SP participation participants are more likely to remain in the labour force and on welfare without finding a job for longer periods. Moreover, we find positive effects of SP participation on annual real earnings of an order of magnitude from 43 up to almost 500 euros during the first three years after programme start.

A key question was whether our estimates of SP effects are in line with an interpretation that this quite flexible programme is more effective than the former classroom training programmes and the placement services by private providers. Therefore, we compared our results to two studies that analysed a sample of welfare recipients who were unemployed at the end of January 2005. They studied effects of short-term training (Wolff/Jozwiak 2007) and contracting out placement services (Bernhard/Wolff 2008) on the participants regular employment rate with administrative data and propensity score matching. The estimates of the effects on the regular employment rate in both studies do not differ substantially from our estimates on SP participation effects.

The estimation results on IFT are much larger than those for SP, which is not surprising as the direct contact to an employer might be followed by a contract. The effect on the employment rate for the four main groups of men and women in East and West Germany according to our estimates reach an order of magnitude of more than level of up to more than 23 percentage points. They tend to be higher for women than for men and decline somewhat remaining solidly positive even 45 months after programme start. There is a clear difference between the employment effect estimates by age groups. In the long term they tend to be highest for those aged at least 50 year with about 15 up to more than 20 percentage points and lowest for those aged 17-24 years with less than ten up to about 15 percentage points. Similar to the SP effects the IFT participation effects are highest for non-employed people with more than one up to five years of non-employment compared with the other non-employment groups considered. Finally, the estimates of the annual earnings effects for IFT participants are quite considerable ranging from about 2,200 up to about 4,300 euros during the first three years after programme start for the four main groups analysed. Moreover, we also find that participation in IFT considerably reduces the participants' dependency on welfare receipt.

A comparison to the IFT results of Wolff and Jozwiak (2007) of a participant sample of previously unemployed welfare recipients in the year 2005 imply that the results are quite similar. One exception are the results for West German women: For them the findings implied an employment effect of only about 13 percentage point 20 months after programme start and hence in the medium term. The corresponding effect in our study is substantially higher with 17.7 percentage points.

The studies of Wolff and Jozwiak (2007) on participation effects of short training programmes and Bernhard and Wolff (2008) on participation effects of contracting out placement services analysed the same type of unemployed welfare people in Germany. We cannot find remarkable differences between their and our findings. These two studies regarded both a different time period and hence a different composition of programme participants and cyclical conditions. Nevertheless, we would still interpret the lack of

considerable differences between their and our results as a hint that the more flexible implementation of such policies did not lead to a substantial improvement of the participation effects of short training or the use of private placement services for welfare recipients. Hence, the evidence cannot confirm that the effectiveness of these policies considerably improved due to the 2009 reform.

The results that we presented are still preliminary and incomplete. We still have to discuss results on further outcomes and include a further important outcome in our analysis, a proxy for participation in the labour force. Moreover, we will carry out some robustness analyses: We plan to estimate the effects additionally by radius-calliper matching and radius matching on the propensity score with bias adjustment (Huber et al. 2015).

Future research could regard in more detail the different options of the SP part of the SAI, which are not defined by their exact implementation but by their goal such as *Guiding into apprenticeships and into work* or *Determining, Reducing and Removing Employment Impediments* as described in section three. Presumably, the effects considerably vary with the goals set by these options. Moreover, it is important to study the SAI at a later point in time. We chose to study programme participation already one year after the introduction of the SAI, to avoid studying a period in which the job centres were gaining first experiences with implementing the programme. Nevertheless, it is possible that setting the PES free to implement such a new policy may still be followed by a period in which they handle it often according to well-known rules of the former programmes that were replaced. Three to four years after the introduction this should be different.

References

- Ashenfelter, Orley (1978): Estimating the Effect of Training Programs on Earnings. *The Review of Economics and Statistics* 60(1): 47–57.
- Behaghel, Luc; Crépon, Bruno; Gurgand, Marc (2012): Private and public provision of counseling to job seekers: Evidence from a large controlled experiment. *IZA Discussion Paper* 6518.
- Benmarker, Helge; Grönqvist Erik; Öckert, Björn (2013): Effects of contracting out employment services: Evidence from a randomized experiment. *Journal of Public Economics* 98: 68–84.
- Bernhard, Sarah; Kopf, Eva (2014): Courses or individual counselling: does job search assistance work? *Applied Economics* 46(27): 3261–3273.
- Bernhard, Sarah; Wolff, Joachim (2008): Contracting out placement services in Germany. Is assignment to private providers effective for needy job-seekers? *IAB-Discussion Paper* 05/2008.
- Caliendo, Marco; Kopeinig, Sabine (2008): Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys* 22(1): 31–72.
- Calmfors, Lars (1994): Active Labour Market Policy and Unemployment – A Framework for the Analysis of Crucial Design Features. *OECD Economic Studies* No. 22. Paris, OECD.
- Carcagno, George J; Cecil, Robert; Ohls, James C. (1982): Using Private Employment Agencies to Place Public Assistance Clients in Jobs. *The Journal of Human Resources* 17(1): 132–143.
- Card, David; Kluge, Jochen; Weber, Andrea (2010): Active Labour Market Policy Evaluations: A Meta-Analysis. *The Economic Journal* 120(548): F452–F477.
- Froy, Francesca; Giguère, Sylvain (2009a): *Which Countries Have Most Flexibility in the Management of Labour Market Policy? An OECD Comparison*, Chapter 2 in Froy, Francesca and Sylvain, Giguère (eds.): *Flexible Policy for More and Better Jobs*, Organisation of Co-operation and Economic Development.
- Froy, Francesca; Giguère, Sylvain (2009b): *A New Framework for Labour Market Policy in a Global Economy*, Chapter 1 in Froy, Francesca and Sylvain, Giguère (eds.): *Flexible Policy for More and Better Jobs*, Organisation of Co-operation and Economic Development.
- Gerfin, Michael; Lechner, Michael (2002): A microeconomic evaluation of the active labour market policy in Switzerland. *The Economic Journal* 112(482): 854–893.
- German Federal Parliament (2008): Gesetzentwurf der Bundesregierung – Entwurf eines Gesetzes zur Neuausrichtung der arbeitsmarktpolitischen Instrumente. [Draft bill of the Federal government – Draft of an Act to readjust labour market policy instruments] *Bundestagsdrucksache* 16/1081.
- Gorter, Cees; Kalb, Guyonne R.J. (1996): Estimating the Effect of Counseling and Monitoring the Unemployed Using a Job Search Model. *The Journal of Human Resources* 31(3): 590–610.
- Hartig, Martina; Jozwiak, Eva; Wolff, Joachim (2008): Trainingsmaßnahmen: Für welche unter 25-jährigen Arbeitslosengeld-II-Empfänger erhöhen sie die Beschäftigungschancen? [Training measures: for which unemployment benefit II recipients below the age of 25 do they increase employment chances?] *IAB-Forschungsbericht* 6/2008.
- Hasluck, Chris; Elias, Peter; Green, Anne (2003): *The Wider Labour Market Impact of Employment Zones*. Warwick: Institute for Employment Research.
- Hausmann, Ann-Christin; Kleinert, Corinna (2014): Berufliche Segregation auf dem Arbeitsmarkt, Männer- und Frauendomänen kaum verändert. [Occupational segregation on the labour market. Hardly any change in male versus female fields.] *IAB-Kurzbericht* 9/2014.

- Holzner, Christian; Munz, Sonja (2013): Should local public employment services be merged with local social benefit administrations? *Journal for Labour Market Research* 46(2): 83–102.
- Huber, Martin; Lechner, Michael; Steinmayr, Andreas (2015): Radius matching on the propensity score with bias adjustment: tuning parameters and finite sample behaviour. *Empirical Economics* 49(1): 1–31.
- Kopf, Eva (2013): Short training for welfare recipients in Germany: which type works? *International Journal of Manpower* 34(5): 486–516.
- Krug, Gerhard; Stephan, Gesine (2016): Private and public placement services for hard-to-place unemployed: results from a randomized field experiment. *ILR Review* 69(2): 417–500.
- Lundin, Martin; Skedinger, Per (2006): Decentralisation of active labour market policy: The case of Swedish local employment service committees. *Journal of Public Economics* 90(4–5): 775–98.
- Nativel, Corinne; Sunley, Peter; Martin, Ron (2002): Localising welfare-to-work? Territorial flexibility and the New Deal for Young People. *Environment and Planning C: Government and Policy* 20(6): 911–932.
- OECD (1999): *The local dimension of welfare-to-work – an international survey*. OECD Proceedings, Organisation of Co-operation and Economic Development, Paris.
- Prey, Hedwig (2009): Evaluation of Training Programs in St. Gallen, Switzerland. *Schweiz. Zeitschrift für Volkswirtschaft und Statistik* 136 (3): 417–432.
- Rosenbaum, Paul R.; Rubin, Donald B. (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70(1): 41–55.
- Roy, Andrew Donald (1951): Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers* 3(2): 135–145.
- Rubin, Donald B. (1974): Estimating Causal Effects to Treatments in Randomised and Nonrandomised Studies. *Journal of Educational Psychology* 66(5): 688–701.
- Sianesi, Barbara (2004): An Evaluation of the Swedish System of Active Labour Programmes in the 1990s. *The Review of Economics and Statistics* 86(1): 133–155.
- Struyven Ludo; Steurs, Geert (2005): Design and redesign of a quasi-market for the reintegration of jobseekers: empirical evidence from Australia and the Netherlands, *Journal of European Social Policy* 15(3): 211–229.
- Thomsen, Stephan (2009): Job Search Assistance Programs in Europe: Evaluation Methods and Recent Empirical Findings. *FEMM Working Paper* 18/2009.
- van den Berg, Gerard J.; van der Klauuw, Bas (2006): Counseling and Monitoring of Unemployed Workers: Theory and Evidence from a Controlled Social Experiment. *International Economic Review* 47(3): 895–936.
- Vikström, Johan (2015): Evaluation of sequences of treatments with application to active labor market policies. *Working Paper* 5/2015.
- Weber, Andrea; Hofer, Helmut (2003): Active Job-search Programs a Promising Tool? A Microeconomic Evaluation for Austria. *Economics Series* 131.
- Wolff, Joachim; Jozwiak, Eva (2007): Does short-term training activate means-tested unemployment benefit recipients in Germany? *IAB-Discussion Paper* 29/2007.
- Wolff, Joachim; Nivorozhkin, Anton; Bernhard, Stefan (2015): You can go your own way! The long-term effectiveness of a self-employment programme for welfare recipients in Germany. *International Journal of Social Welfare*, Online First.
- Wunsch, Conny; Lechner, Michael (2008): What Did All the Money Do? On the General Ineffectiveness of Recent West German Labour Market Programmes. *Kyklos* 61(1): 134–174.
- WZB/infas (2006): *Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission*. Modul 1a: Neuausrichtung der Vermittlungsprozesse. Endbericht 2006.

[Evaluation of the measures for implementing the proposals of the Hartz commission. Module 1a: Redesign of placement procedures. Final report 2006.] Berlin, Bonn.

Zabel, Cordula (2013): Effects of participating in skill training and workfare on employment entries for lone mothers receiving means-tested benefits in Germany. *IAB-Discussion Paper 3/2013*.

Appendix

Table A1: Matching quality in the three age subsamples

Outcome/statistic	East Germany		West Germany	
	Men	Women	Men	Women
SP - schemes by providers				
<u>age 18-24 years</u>				
Treatment group obs. on support	2,347	1,710	4,577	3,187
Treatment group obs. off support	1	11	5	10
Control group observations	3,971	3,364	6,644	5,572
Mean Standardized Bias after matching	0.802	1.140	0.809	0.922
Adjusted R ² after matching	0.002	0.003	0.002	0.002
<u>age 25-49 years</u>				
Treatment group obs. on support	8,706	6,645	22,373	16,546
Treatment group obs. off support	2	8	1	2
Control group observations	36,776	29,638	63,847	59,862
Mean Standardized Bias after matching	0.391	0.440	0.276	0.377
Adjusted R ² after matching	0.001	0.001	0.000	0.001
<u>age 50+ years</u>				
Treatment group obs. on support	1,817	1,675	4,831	3,609
Treatment group obs. off support	1	0	1	1
Control group observations	14,175	12,038	22,280	19,751
Mean Standardized Bias after matching	1.272	1.040	0.749	0.690
Adjusted R ² after matching	0.003	0.003	0.001	0.002
IFT - in-firm training				
<u>age 18-24 years</u>				
Treatment group obs. on support	446	379	715	354
Treatment group obs. off support	2	3	1	3
Control group observations	1,613	1,096	3,197	1,965
Mean Standardized Bias after matching	2.151	2.580	1.717	2.182
Adjusted R ² after matching	0.010	0.015	0.006	0.009
<u>age 25-49 years</u>				
Treatment group obs. on support	3,034	2,151	5,685	2,838
Treatment group obs. off support	1	1	8	0
Control group observations	32,191	23,657	55,519	44,968
Mean Standardized Bias after matching	0.750	0.806	0.582	0.796
Adjusted R ² after matching	0.002	0.003	0.001	0.002
<u>age 50+ years</u>				
Treatment group obs. on support	465	421	851	540
Treatment group obs. off support	0	3	0	4
Control group observations	13,356	10,809	20,323	15,859
Mean Standardized Bias after matching	1.592	1.800	1.299	1.683
Adjusted R ² after matching	0.005	0.007	0.004	0.006

Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.

Table A2: Matching quality in the four non-employment subsamples

Outcome/statistic	East Germany		West Germany	
	Men	Women	Men	Women
SP - schemes by providers				
<u>less than one year since last job</u>				
Treatment group obs. on support	2,661	1,506	8,580	4,165
Treatment group obs. off support	7	12	0	1
Control group observations	8,427	5,054	17,797	9,987
Mean Standardized Bias after matching	0.746	1.073	0.582	0.643
Adjusted R ² after matching	0.001	0.002	0.001	0.001
<u>one to five years since last job</u>				
Treatment group obs. on support	4,387	2,444	13,330	6,169
Treatment group obs. off support	2	0	3	3
Control group observations	17,147	10,439	36,156	21,117
Mean Standardized Bias after matching	0.844	0.873	0.465	0.716
Adjusted R ² after matching	0.001	0.001	0.001	0.001
<u>five to ten years since last job</u>				
Treatment group obs. on support	3,299	2,416	7,279	5,338
Treatment group obs. off support	1	1	0	1
Control group observations	16,738	11,610	26,286	21,141
Mean Standardized Bias after matching	0.813	0.930	0.524	0.570
Adjusted R ² after matching	0.001	0.002	0.001	0.001
<u>more than ten years since last job or never worked</u>				
Treatment group obs. on support	3,222	3,948	6,210	8,480
Treatment group obs. off support	4	2	2	0
Control group observations	16,118	19,493	22,377	36,348
Mean Standardized Bias after matching	0.989	0.831	0.809	0.691
Adjusted R ² after matching	0.002	0.002	0.001	0.001
IFT - in-firm training				
<u>less than one year since last job</u>				
Treatment group obs. on support	1,641	1,113	2,909	1,381
Treatment group obs. off support	0	4	3	0
Control group observations	7,444	4,824	14,706	9,480
Mean Standardized Bias after matching	1.235	1.114	0.667	1.163
Adjusted R ² after matching	0.005	0.005	0.002	0.004
<u>one to five years since last job</u>				
Treatment group obs. on support	1,803	1,219	3,311	1,530
Treatment group obs. off support	4	3	2	3
Control group observations	15,478	10,005	31,108	20,387
Mean Standardized Bias after matching	1.096	1.258	0.699	1.020
Adjusted R ² after matching	0.003	0.006	0.002	0.005
<u>five to ten years since last job</u>				
Treatment group obs. on support	617	594	1,013	757
Treatment group obs. off support	1	3	1	0
Control group observations	15,445	11,252	22,974	20,608
Mean Standardized Bias after matching	1.580	1.339	1.177	1.613
Adjusted R ² after matching	0.010	0.007	0.005	0.009
<u>more than ten years since last job or never worked</u>				
Treatment group obs. on support	482	682	856	975
Treatment group obs. off support	4	5	0	0
Control group observations	14,704	18,540	19,371	34,377
Mean Standardized Bias after matching	1.950	1.312	1.538	1.474
Adjusted R ² after matching	0.008	0.006	0.008	0.008

Matching was performed using nearest-neighbour-matching with five neighbours, with replacement and without calliper.