Employment and Earnings Effects of Awarding Training Vouchers in Germany

Annabelle Doerr

University of Freiburg

Bernd Fitzenberger

Humboldt-University Berlin
IZA, CESifo, IFS, ROA, ZEW

Thomas Kruppe

IAB, Nuremberg

Marie Paul

University of Duisburg-Essen
RGS Econ

Anthony Strittmatter

University of St. Gallen
University of Freiburg

November 28, 2015

Abstract

Participation in intensive training programs for the unemployed in Germany is allocated by awarding training vouchers. Using rich administrative data for all vouchers and on actual program participation, we provide first estimates of the short-run and long-run employment and earnings effects of a voucher award based on a selection-on-observables assumption. Our results imply that after the award, voucher recipients experience long periods of lower labor market success. There are only small positive employment effects and no gains in earnings even four to seven years after the voucher award. However, our findings suggest stronger positive effects both for low-skilled individuals and for degree courses.

JEL-Classification: J68, H43, C21

Keywords: Active Labor Market Policies, Training Vouchers, Treatment Effects Evaluation, Administrative Data

We gratefully acknowledge helpful comments from Joshua Angrist, Thomas Bauer, Joachim Moeller, Anna Salomons, Carolyn Heinrich, Peter Mueser, and Jeffrey Smith, as well as from participants in various seminars. This study is part of the project “Regional Allocation Intensities, Effectiveness and Reform Effects of Training Vouchers in Active Labor Market Policies”, IAB project number 1155. This is a joint project of the Institute for Employment Research (IAB) and the University of Freiburg. We gratefully acknowledge financial and material support from the IAB. The usual caveats apply.

Corresponding Author: Bernd Fitzenberger, Chair of Econometrics, Humboldt-University Berlin, 10099 Berlin, Germany. E-mail: bernd.fitzenberger@wiwi.hu-berlin.de, Tel: +49-30-2093-5711.

1 Introduction

Vocational training for the unemployed is an important part of active labor market policy (ALMP) in many countries. Such programs aim at the enhancement of vocational skills to improve the labor market prospects of the participants. Before 2003, unemployed in Germany were assigned by caseworkers to training programs. Since 2003, participation in training programs for the unemployed has been allocated by awarding training vouchers. Training vouchers in Germany specify the educational objective of the training program for which the voucher can be redeemed. The voucher allows the recipient to choose among a set of eligible training providers, and voucher recipients are free not to redeem the voucher without having to fear sanctions. This paper provides first estimates of the short-run and long-run employment and earnings effects of a voucher award. Using rich administrative data, our estimates control for selection with respect to a large set of observable characteristics.

In contrast to direct assignment or contracting-out, government agencies use vouchers to provide recipients with the financial means to pay for participation in the intended program (e.g. training or education), but not for other goods or services, while leaving the choice of the provider to the recipient.¹ Recipients of training vouchers may choose the training provider among a set of certified private or public providers, which should induce efficiency enhancing competition among providers and which should improve the self-responsibility of the training participants (see e.g. Barnow, 2000, 2009, Hipp and Warner, 2008, for a discussion of the potential advantages and disadvantages of using vouchers for the allocation of vocational training programs). Market access is not limited to incumbent training providers thus making training markets more contestable (Baumol, 1982). Allowing for more choice by the participants could in principle result in better choices, thus increasing the effectiveness of training (Posner et al. 2000), because the unemployed may find a better match. The voucher recipients may put more

¹Training vouchers are not only used in the context of ALMP but also to foster training of employees (see Görlitz, 2010, for a recent evaluation of such training vouchers in Germany). Education vouchers are for the most part used in the schooling system (Posner et al. 2000, and Ladd, 2002, for a review of the literature on education vouchers).

effort into achieving the training goals, if they have a choice, possibly due to a reciprocity motive. Consumer sovereignty includes the freedom of not redeeming the voucher without having to fear a sanction, when there is no suitable training provider or when a job offer has arrived since the voucher award. There is concern that the unemployed may not be sufficiently informed to make good choices in using the training vouchers, a frustrating experience which may hinder their job search. Furthermore, concerns by the unemployed unrelated to the effectiveness of the program may drive the redemption decision. It could be the case that experienced caseworkers have a better understanding of the opportunities in the local labor market, of the training providers that offer the best programs, and of the courses that are the most suitable for a particular unemployed individual. During the years 2003 and 2004, caseworkers were supposed to award a training voucher only when it could be expected that the probability to find a job after training participation was above 70%, possibly resulting in cream-skimming effects.

The Adult and Dislocated Worker Program under the Workforce Investment Act (WIA) in the U.S. and the German Training Vouchers are two important cases that use vouchers for the provision of training.² In the U.S., customers in the WIA program can use the fixed budget provided by the government-funded Individual Training Accounts (ITA) to pay for participation in training. The choice is restricted to eligible training providers who offer occupational skills in demand on the local labor market (see Barnow, 2009, for an overview). However, the customer can choose between training courses, which may differ greatly in content. Training vouchers in Germany differ in various aspects. Vouchers are awarded by caseworkers if they consider training to be helpful for finding a job. A voucher recipient may only choose a course offered by an eligible training provider that fits the training content and the planned duration specified by the caseworker. The voucher recipient may state his or her preference (for example, to become an IT specialist) during an appointment with the caseworker before the voucher is awarded, but ultimately, the caseworker decides upon the content of the training. However, after the award of the voucher, the caseworker is not

²In fact, WIA was a key benchmark for the introduction of the German training vouchers (Hipp and Warner, 2008). However, there are a number of important differences to be discussed below.

permitted to provide guidance regarding the choice of a training course or suitable training provider.

The award of a training voucher allows the recipient to choose among the set of certified training providers and also to decide not to redeem the voucher.

Rinne et al. (2013) estimate the effects of actual participation in training under the voucher system in Germany (earlier estimates for the employment effects of training participation under the voucher system can be found in the report for the government by Schneider et al. 2007 and in the companion study by Rinne et al. 2011). Using a dynamic matching approach, Rinne et al. (2013) find positive effects of training participation on employment and earnings 1.5 years after the program start. However, the authors do not observe the award of vouchers but only actual program participation. Thus, the study does not evaluate the treatment “voucher award” but the treatment “training participation”. With the latter approach, first, individuals not redeeming a voucher are in the comparison group and, second, the treatment start and thus also the alignment of participants and comparison persons occurs in the month in which the treatment starts and not in the month in which the voucher is awarded. Evaluating the treatment “training participation” requires more stringent assumptions to identify a causal effect than those required for the treatment “voucher award”. Because actual participation in a training program is an outcome of the redemption decision by the unemployed after the voucher award, there may be endogenous sorting into training participation even if the conditional independence assumption holds for the treatment “voucher award”. For the treatment “training participation”, the researcher would need to account for the dynamic selection both for the voucher award and actual participation. Moreover, the fact that participants have already received a voucher some time ago when they actually start training may call into question the assumption that individuals cannot perfectly anticipate the start of the training spell, an assumption which typically would have to be invoked when applying a dynamic matching approach.

To the best of our knowledge, our study is the first to estimate the effect of *being awarded* a training voucher for the unemployed as an intention-to-treat

effect.³ From a policy perspective, it is the effect of the voucher award that is of prime interest, because this is the policy intervention. The caseworker decides upon the voucher award but cannot control the actual participation in training. This holds in particular because since 2003 caseworkers have not been permitted to sanction an unemployed individual for not redeeming a voucher. We apply a matching strategy which accounts for selection based on observable characteristics. To avoid the bias that is inevitable if a static evaluation approach is used in a dynamic setting (Frederiksson and Johansson, 2008), we follow Sianesi (2004) and estimate the effects of starting treatment now versus not starting treatment now for each month of elapsed unemployment. The alternative of not starting treatment now entails the possibility that treatment starts in the future. This evaluation approach aligns treated individuals and comparison individuals by the elapsed unemployment duration, and it only compares individuals who are still unemployed at the time of the treatment start. The approach is implemented using both inverse probability weighting (IPW) and ordinary least squares (OLS) regressions.

Our study uses unique rich administrative data provided by the Federal Employment Agency in Germany. We have information on *all* individuals who received training vouchers in 2003 or 2004 and on a 3% sample of all other unemployed. Our data allow us to follow individuals for more than seven years after the voucher award. The data include precise award dates and redemption dates for the vouchers. This information has not been previously available for evaluation studies. We merge the voucher data with individual data records from the Integrated Employment Biographies (IEB), which contain information on employment outcomes and a rich set of control variables, e.g., the complete employment and welfare history, various socioeconomic characteristics, information on health and disabilities, and regional labor market characteristics.

Our results imply that the award of a training voucher has strong negative

³There is a large literature estimating the effects of public sponsored training for the unemployed in Germany (see Biewen et al. 2014, Thomsen and Zeiss, 2006, Lechner, Miquel, and Wunsch, 2011, 2007, and Rinne et al. 2013). With the exception of the last study, the literature analyzes the time period before the introduction of the voucher system. The evidence on employment and earnings effects of vocational training is mixed; see Card, Kluve, and Weber (2010) for a recent review.

lock-in effects both regarding employment and earnings. Lock-in effects of training programs can be explained by a lower job search intensity during program participation, and training programs in Germany may even last more than two years. Four years after the voucher award significantly positive employment effects are found, which persist over the next three years until the end of our observation period. There are no positive long-term effects on earnings during the whole observation period. OLS and IPW lead to virtually the same results. A comparison of raw differences between the treatment and comparison group indicates a strong positive selection of voucher recipients with respect to observable characteristics that are expected to improve labor market chances.

Allowing for effect heterogeneity identifies important cases for which a voucher award is more effective. The employment and earnings effects are more positive for individuals without a vocational degree. A decomposition of the effect estimates reveals that programs leading to a vocational degree seem particularly effective and that those unemployed who do not redeem the voucher do much worse in the long-run than comparable individuals who are not awarded with a voucher. This suggests that any positive effect of a voucher award actually works through participation in training.

The remainder of the paper is organized as follows: The next section provides background information on training vouchers in the U.S. under WIA and training vouchers in Germany. This is followed by the data description. Section 4 discusses identification and estimation. We present our results on the average voucher effect and effect heterogeneity in Section 5. The final section concludes.

2 Background

First, as a reference point for our analysis, we review training vouchers in the U.S. under WIA and discuss some important literature for the U.S. case. Then, we provide further background information on training vouchers in Germany.

2.1 Training Vouchers in the U.S.

In the U.S., customers in the WIA program receive basic job search assistance, many receive intensive counseling or short training courses, and some are awarded an Individual Training Accounts (ITA). WIA customers can use the fixed budget provided by the government-funded ITA to pay for participation in training under the following restrictions. The content of the course must relate to an occupation in demand on the local labor market (which is defined by the local agency), and the training provider must be listed as an eligible provider. The choice of the content of the training is left to the customer. However, the customer typically has to undergo counseling, which involves an assessment of skills, research on the training programs and the labor market, and face-to-face discussions with the caseworker about the course to choose (McConnell et al. 2011, King and Barnow 2011). WIA customers in the U.S. receive guidance on how to use the voucher but ultimately make the decision regarding the content of the training. Thus, after a guided and mandatory decision process, the voucher recipient may decide, for example, to enroll in training to become an IT specialist instead of a care nurse.

There exist several studies on the ITA's and preceding voucher-like programs involving descriptive evidence, experimental evidence, or qualitative evaluations of the implementation (see Barnow, 2009, for an overview). Perez-Johnson et al. (2011) discuss an experiment that was conducted to study the relative effectiveness of different levels of counseling and control by the caseworkers in WIA. One extreme case would be a system in which caseworkers direct customers to a specific course through counseling, award an ITA corresponding directly to a customer's need, and have the right to reject a customer's choice. In a polar-opposite case, caseworkers can award all customers with the same fixed amount for the ITA and provide counseling upon request only. The majority of agencies use a system somewhere in between these two extremes. For the experiment, individuals who were to receive an ITA under the WIA at one of seven particular sites were randomly assigned to different treatments regarding the freedom of choice of the customer, the counseling requirements, and the award structure. With regard to long-term labor market outcomes, participants in the different treatments are

equally likely to be employed six to eight years after the experiment, but those with the greatest level of caseworker guidance (the structured choice group, see Perez-Johnson et al. 2011, Table ES.1) show the highest earnings. This suggests that the lack of guidance after the voucher award by the caseworker in Germany could play a role for earnings.

In a large scale econometric evaluation of WIA, Heinrich et al. (2013) provide separate estimation results for participants in the Adult programs (targeted to individuals with poor work histories) and participants in the Dislocated Worker programs (targeted to individuals who have been laid off).⁴ In their main analysis, they estimate the effects of participating in WIA (regardless of the services that are taken) as opposed to not entering WIA. They find large positive employment and earnings effects for the Adult program and find positive employment effects, though only small and insignificant earnings effects, for the Dislocated Worker program. Heinrich et al. (2013) also estimate the effects of receiving training through an ITA as opposed to receiving only the other services of the WIA (and possibly training not related to the WIA program), but suggest that the results be interpreted with caution. For the Adult program, the long-run earnings effects are large, and there are also positive long-run employment effects. The authors find no positive effects for the Dislocated Worker program in their observation period of four years. Heinrich et al. (2013) estimate the effect of participating in training assigned through an ITA and do not estimate the effect of being awarded an ITA. Although the study does not provide information about the share of individuals who receive an ITA and who do not actually participate in training, the authors say that “training slots may be a scarce resource” (Heinrich et al. 2013, page 4), i.e. WIA customers value the financial resources available for training, and so suggest that receiving an ITA will typically result in actual training participation. This is in contrast to Germany, where a considerable number of those receiving a voucher actually do not participate in training (22%).

⁴In a recent study, Heinrich and Mueser (2014) extend the analysis to the period of the great recession 2008 and 2009 in order to investigate as to whether the effects of training differs over the business cycle, e.g. because both the lock-in effect may be less pronounced and the selection of the training participants may be less negative in a recession.

2.2 Training Vouchers in Germany

Since 2003, participation in intensive training programs for the unemployed in Germany has been allocated by awarding training vouchers.⁵ During an unemployment spell, individuals repeatedly meet their caseworker for counseling. In the profiling process, the caseworker reviews their potential labor market opportunities. If there is a lack of necessary qualifications to find a job immediately, participation in a training course is considered necessary. The caseworker denotes the objective, content, and maximum duration of the course on the voucher. It is thus the task of the caseworker (potentially in discussion with the unemployed individual) to decide upon the training objective and the educational content of the course. The unemployed individual may choose the provider and the particular course offered. The provider must be an eligible training provider that is located within a one-day commuting zone subject to the restrictions denoted on the voucher.⁶ Eligible (certified) training providers are listed in an online tool provided by the employment agency, and providers may also advertise their courses, e.g., by placing handouts in the employment agency.⁷

The course the job seeker registers for must match all restrictions denoted on the voucher. Providers usually indicate on their webpage and in the online tool of the employment agency which objectives potentially denoted on the voucher their courses match with. A person seeking a training program in “financial accounting with SAP” in an urban area will have more choice than a person seeking for a rare training program like “CNC-programmer wood” in an area with hardly any forest industry. Having said this, caseworkers will broadly know which training

⁵At the time, the political debate evolved around the concern that vocational training was not effective and that this might have been related to the overly close relationships between local employment agencies and training providers. The First Modern Services on the Labor Market Act (the so-called *Hartz I Reform*) introduced a voucher system for the provision of training for the unemployed in January 2003. Its aim is to foster market mechanisms and transparency in the training market. For more details on the reform, see Schneider et al. (2007) and Hipp and Warner (2008).

⁶The one-day commuting zone is defined as a regional zone that can be reached by public transport in a reasonable amount of time. For a training course with six or more hours a day, commuting times of up to 2.5 hours are reasonable. For a training course with less than six hours a day, the reasonable commuting time is reduced to two hours.

⁷In 2003 and 2004, the Federal Employment Agency was in charge of the certification of the eligible training providers. Afterwards, the certification process was privatized.

programs are offered in the region and will usually not award vouchers which are impossible to redeem. To ensure that training providers offer courses that are in line with the regional labor demand, the local employment agencies publish regional and sector-specific demand once a year.⁸ The caseworker is not allowed to give advice as to the choice of provider in order not to intervene excessively into the decision sovereignty of the unemployed and to foster the competition among training providers for potential training participants. Training vouchers are valid for at most three months, so training has to start within this period. There are no sanctions if the awardee decides not to start a training program.⁹

The fact that the choice of the unemployed is restricted to eligible training providers who offer occupational skills in demand at the local labor market is similar to the WIA system in the U.S.. The award of a training voucher allows the recipient to choose among the set of certified training providers and also to decide not to redeem the voucher. However, there is less choice by the unemployed in Germany regarding the content of training as described above. The voucher recipient may state his preference (for example, to become an IT specialist) before the voucher award, but, ultimately, the caseworker decides upon the content of the training. But after the award of the voucher, there is less guidance and counseling than under the WIA. Furthermore, a difference between the German system and the WIA is that after redemption of the voucher in Germany training participation is mandatory, i.e. a drop out from a training program could result in a sanction.

Vocational training programs in Germany are used to improve the human capital and productivity of the participants. This allows adjustment of the skills of the unemployed to changing requirements of the labor market as well as facilitating increases in employability of the unemployed (e.g. in case of health problems). Participation in training prolongs the entitlement period for unemployment bene-

⁸This is similar to the WIA, stipulating that the local agency provides a list of occupations in demand at the local level.

⁹There are various potential reasons for the expiration of a training voucher (e.g., the individuals did not find an appropriate course or letting the voucher expire is part of an individual's strategy). In such cases, individuals have no legal claim to being awarded a second voucher and the award of a second voucher is up to the discretion of the caseworker. Our data show that 25% of those individuals who did not redeem the voucher receive a second voucher later on.

fits.¹⁰ Intensive training considered here mainly comprises two types of programs, namely *long-term training* and *degree courses*. Long-term training courses typically last several months to one year (in our sample, an average of five months) and usually involve full-time programs. Teaching takes place in class rooms or on the job in training firms. The course curriculum may also include internships. Typical examples of training schemes are courses on IT-based accounting or on customer orientation and marketing. With a typical duration of two to three years, degree courses (similar to the former retraining programs) last much longer and lead to a full new vocational degree within the German apprenticeship system. 77% of degree courses provide training for employment in the service sector. Typical examples are training as elderly care nurse or as office clerk. Another 20% of degree courses qualify the participants to be employed in the manufacturing sector or in the trades, e.g. as electrician or industrial mechanic. Other sectors are only marginally represented.

Although the Federal Employment Agency typically covers the full costs for at most two years, these programs may last for three years and other programs exist (e.g., those sponsored directly by the state government) that cover the additional costs.¹¹

There is a discussion as to whether training participation under the voucher system is particularly effective (see, for example, Barnow, 2000, 2009, Hipp and Warner, 2008, for a discussion of the potential advantages and disadvantages of using vouchers for the allocation of further training programs). First, voucher recipients have a choice of the training provider and the particular program, which should lead to efficient outcomes if they know their needs best. However, it may in contrast be the case that experienced caseworkers have a better understanding of the training providers that offer the best programs and the courses that are the most suitable for a particular unemployed individual. Furthermore, the choice by the unemployed individual may be driven by preferences unrelated to the effectiveness of the training program. Second, the increased course choices may

¹⁰The duration of unemployment benefits varies between 12 and 36 months depending on previous employment and age.

¹¹Recall that this differs from WIA where the individual has a fixed budget and may have to pay himself/herself for part of the training costs.

have a positive effect on the provider side. One would expect that competition for potential clients will have a positive effect on the selection of providers in addition to strengthening the efficiency on their part.¹²

In addition to the opportunity to take part in an intensive training program, training vouchers may influence future labor market opportunities directly. They may improve the self-responsibility of the training participants and the freedom not to redeem the voucher may change the attitude of the unemployed individual toward this service: The voucher may be perceived more like an *offer* and less like an *assignment*. This could create a positive attitude such that the unemployed individual may value the fact that a costly service is being offered to him or her. The individual may reciprocate by increasing the search effort or by participating wholeheartedly in the training program. But it may also be the case that some individuals who feel incapable of finding a suitable course end up being less motivated with regard to their career. Moreover, some unemployed with a training voucher may just enjoy their unemployment benefits for the three months after voucher award without being pushed to find a job during these months. Their employment chances may deteriorate because of reduced search effort right after the voucher award.

The so called “70% rule” was introduced at the same time as the voucher system. Based on the assessment of the employment chances after participation in a training program, the caseworkers are supposed to only award vouchers when it can be expected that at least 70% of the voucher recipients find a job within six months after training ends.¹³ Due to career concerns of both caseworkers and decision makers in local employment agencies, this rule is likely to result in cream-skimming effects, i.e., a positive selection of those awarded with a training

¹²The supply of training provider is not well studied in the literature due to data limitations (e.g. Hipp and Warner, 2008, do not have information on the level of competition among training providers). Anecdotal evidence suggests that the level of competition was low prior to the reform. One intention of implementing training vouchers was to introduce a market mechanism to increase the competition among training providers.

¹³The assessment was made by the caseworker without a formal prediction model or, in fact, without any formal guidelines. The actual job finding rate never reached this level. Since there were no consequences in case the rule was not satisfied, it was very hard to really enforce compliance with the rule. For this reason, the 70% rule was abolished after 2004. However, the recommendation to act according to the 70% rule exists until today.

voucher with regard to their predicted future employment outcomes irrespective of the actual effectiveness of participation in training.¹⁴

3 Data

This study is based on unique data provided by the Federal Employment Agency of Germany. These data contain information on *all* individuals in Germany who received a training voucher in 2003 or 2004. The data are generated from internal administrative data and include precise award and redemption dates for each voucher - information that previously has not been available for evaluation purposes.

For each voucher recipient, we merge the information on training vouchers to the individual's data record in the Integrated Employment Biographies (IEB).¹⁵ The data contain detailed daily information on employment subject to social security contributions, receipt of transfer payments during unemployment, job search, and participation in different active labor market programs as well as rich individual information.¹⁶ Thus, we are able to enrich the information from the voucher data with a large set of personal characteristics and a long labor market history for all voucher recipients.

Our comparison persons are from the same data base: A 3% random sample (based on twelve days of birth during the year) of those individuals in Germany who experience at least one switch from employment to non-employment (of at least one month) between 1999 and 2005 has been drawn. When constructing our sample of analysis, we apply the same selection rules for voucher recipients

¹⁴Hipp and Warner (2008) point towards likely cream-skimming effects, however, without providing corresponding econometric evidence.

¹⁵The IEB is a rich administrative data base that is the source of the subsamples of data used in all recent-year studies evaluating German ALMP. It is a merged data file containing individual data records collected in four different administrative processes: The IAB Employment History (*Beschäftigten-Historik*), the IAB Benefit Recipient History (*Leistungsempfänger-Historik*), the Data on Job Search originating from the Applicants Pool Database (*Bewerberangebot*), and the Participants-in-Measures Data (*Maßnahme-Teilnehmer-Gesamtdatenbank*).

¹⁶A more detailed description of the IEB in English can be found on the website of the Research Data Center of the Federal Employment Agency (<http://fdz.iab.de/en.aspx>). The version of the IEB we use in this project has been supplemented with some personal and regional information not available in the standard version.

and comparison persons. We account for the fact that we use a 100% sample of voucher recipients and a 3% sample of non-recipients using the inverse inclusion probabilities as weights.

We consider an inflow sample into unemployment consisting of individuals who became unemployed in 2003, after having been continuously employed for at least three months. Entering unemployment is defined as the transition from (non-subsidized, non-marginal) employment to non-employment of at least one month plus a subsequent (not necessarily immediate) contact with the employment agency, either through benefit receipt, program participation, or a job search spell.¹⁷ We only consider unemployed individuals who are eligible for unemployment benefits.¹⁸ This sample choice reflects the main target group for the training vouchers. To exclude individuals eligible for specific labor market programs targeted to youths and individuals eligible for early retirement schemes, we only consider persons aged between 25 and 54 years at the beginning of their unemployment spell.

Our inflow sample enters unemployment during the year 2003, a year with a high, but for Germany in the decade between 1995 and 2005 not atypically high, unemployment rate. The year 2003 falls into the prolonged recession period in Germany after the end of the new-economy-boom in the early 2000s, when the unemployment rate started to increase from 2001 onwards. In 2004, the economy starts to recover and the unemployment rate reaches its maximum in 2005, while participation in the longer programs is still continuing. In 2006 to 2007, when almost all participants have left their training program, the unemployment rate is falling and the economy is growing strongly. Thus, the economic situation is rather favorable for the participants who lost their jobs during a recession and finish the program in good times - but of course the same holds for the comparison group. However, the labor market conditions are by no means extraordinary. One year after the introduction of the voucher system in January 2003, further

¹⁷Subsidized employment refers to employment in the context of an ALMP. Marginal employment refers to employment of a few hours per week only; this is due to specific social security regulations in Germany.

¹⁸Note that, in particular, this condition excludes training programs for mothers returning to the labor market after longer employment interruptions.

labor market reforms follow which result in lower benefits in case the training participants remain unemployed after finishing the program. This means that while the participants are in the program the policy context changes to a less generous system, but this did not come as a surprise since the policy changes had been discussed since 2002. After 2005 the certification procedures for the training providers are changed and the 70% rule is transformed from being mandatory to being a recommendation. Therefore, we consider the sample analyzed to be specific for the institutional setting in 2003 and 2004, but we do not think that the economic context and the provision of training were atypical for the operation of the voucher system since the reform in 2003.

We aggregate the spell information in the original data into calendar months. We follow a person in the sample from the inflow into unemployment until the end of 2004 regarding a voucher award and until April 2012 regarding the employment outcome. We do not consider individuals who receive a training voucher after December 2004 because the next step of the labor market reforms also affecting training was implemented in January 2005. We analyze the first voucher awarded, and thus the treated in our sample have no previous experience with a voucher award. We distinguish two alternative outcome states: Non-subsidized, non-marginal employment (henceforth denoted as employment), and non-employment. As an alternative outcome variable, we use monthly earnings. Information from prior periods is exploited when constructing the covariates referring to the labor market history. The panel data set for the analysis is completed by adding personal, occupational, and regional information. Covariates on individual characteristics refer to the time of inflow into unemployment, whereas covariates on regional characteristics are updated each month.

The final sample includes 132,928 unweighted observations, which involve 45,287 individuals receiving a voucher and 87,641 individuals not receiving a voucher during the first twelve months of unemployment.¹⁹ There are 35,249

¹⁹The total number of inflows in ALMP programs which focus on skill formation amounts to 1,502,166 individuals in 2003 and 1,548,439 individuals in 2004. The vast majority of these involve short-term training programs (Trainingsmaßnahmen) involving fairly general skills and job search assistance. These short-term training programs last up to 3 months, whereas the training programs assigned via vouchers involve more intensive programs typically lasting much longer. 170,311 (184,039) persons receive a training voucher in 2003 (2004).

(10,038) individuals in our sample who (do not) redeem their vouchers. This amounts to a redemption rate of 78%.²⁰

Tables 1 to 4 report the mean values for the most important socioeconomic and labor market characteristics of the individuals in the evaluation sample. In the first two columns of each table, we display the mean value of the respective control variable in the treatment and in the comparison subsample. In columns six and seven, we distinguish between those who redeem the voucher and those who do not. Voucher recipients are on average more often middle-aged, single-parent and females than the individuals in the comparison group. They exhibit fewer health problems. Individuals who redeem the training voucher and thus participate in a training course are on average slightly older and healthier than individuals who do not redeem their voucher. In addition, the fraction of individuals with children living in the same household is somewhat higher, and the children are on average older than the children of individuals not redeeming a voucher.

Voucher recipients hold a higher schooling degree on average. Furthermore, they exhibit more successful employment histories during the seven years before the unemployment spell, and in particular, they had higher earnings. The share of individuals with stable employment and no participation in an active labor market program in the past is much higher in the treatment group, already suggesting a strong positive selection of the treated. We also have information about potential placement handicaps of the unemployed, e.g., indirect information about past psycho-social or drug problems, lack of motivation, sanctions from caseworkers, or past incapacities due to illness, pregnancy or child care. Those receiving a training voucher are less likely to exhibit problems of this type. The fraction of people with motivation deficits or past incapacities is even lower for individuals who redeem the voucher.

²⁰Those not redeeming their voucher would be part of the comparison group, if we used the sample design of Rinne et al. (2013).

4 Identification and Estimation

Our analysis relies on a dynamic selection-on-observables identification strategy, which is motivated by the richness of our administrative data. We also use an instrumental variables (IV) approach, but despite a highly significant first stage the IV treatment effects were only imprecisely estimated. The IV estimates do not contradict the results of the selection-on-observables approach discussed in the following and thus provide additional evidence in favor of this strategy. The IV results are available upon request.²¹

We consider voucher awards during the first twelve months of unemployment in the first unemployment spell between January 2003 and December 2004. Each unemployed individual is observed for at least 84 months. The indicator for a voucher award as an intention to treat is denoted by $D_{im} \in \{0, 1\}$ (with individuals $i = 1, \dots, N$ and $m = 1, \dots, 12$ indicating the elapsed unemployment duration at the time when the voucher is awarded in months). The outcome variable is denoted by Y_{imt} (where $t = 1, \dots, 88$ indicates the number of months since the award of the voucher). We consider employment and monthly earnings as outcome variables, and we estimate the effect of the voucher award (not the actual training participation). To avoid the bias that is inevitable if a static evaluation approach is used in a dynamic setting (Frederiksson and Johansson, 2008), we follow Sianesi (2004) and estimate the effect of treatment start versus no treatment start (treatment versus waiting) for each month of elapsed unemployment duration. The treatment is the award of a voucher, i.e., the intention to assign further training. In the results section, we report a weighted average of the twelve monthly dynamic treatment effects (see Appendix A for details). Our approach controls for pre-treatment outcome variables which are supposed to control for

²¹The IV-estimates were shown as a sensitivity analysis in an earlier version of this paper. We investigated the robustness of the main results by providing IV-estimates, which exploit the unexplained variation in policy styles across regional employment agencies. The IV effect estimates at the monthly frequency are quite imprecisely estimated, and often not significant. To gain precision, we also considered average effects by the year since the voucher award, which are much more precisely estimated. The difference between the yearly IV estimates and yearly estimates based on a selection-on-observables identification strategy is only significant in one out of eight years for employment and never significant for earnings. We take this as evidence in favor of our selection-on-observables identification strategy.

unobserved permanent characteristics that influence the post-treatment outcome. At the same time, we account for the dynamic selection of the eligible unemployed by controlling for elapsed unemployment duration.

The potential outcomes are indicated by Y_{imt}^d , where $d = 1$ under treatment and 0 otherwise. For each individual unemployed until month m , only the realized outcome $Y_{imt} = Y_{imt}^1 \cdot D_{im} + Y_{imt}^0 \cdot (1 - D_{im})$ is observed. Our goal is to estimate the expected difference between the outcomes Y_{imt}^1 and Y_{imt}^0 for treated individuals

$$\gamma_{mt} = E[Y_{imt}^1 | D_{im} = 1] - E[Y_{imt}^0 | D_{im} = 1].$$

$E[Y_{imt}^1 | D_{im} = 1]$ is identified from observed data. In contrast, $E[Y_{imt}^0 | D_{im} = 1]$ involves the expected counterfactual non-treatment outcome for treated individuals. To identify this parameter, we need to make further assumptions.

4.1 Identification Strategy

Assuming that there is only selection on observables, it is possible to control for all confounding variables that jointly influence the treatment probability and the potential non-treatment outcome, summarized by the vector of pre-treatment variables X_{im} . This is formalized by the following dynamic version of the conditional mean independence assumption.

Assumption 1 (*Strong Ignorability*).

- i) Dynamic mean independence assumption (DMIA):

$$E[Y_{imt}^0 | D_{im} = 1, X_{im} = x] = E[Y_{imt}^0 | D_{im} = 0, X_{im} = x] \text{ and}$$

- ii) Common support: $p(x) < 1$, where $p(x) = Pr(D_{im} = 1 | X_{im} = x)$

hold jointly for all $m = 1, \dots, 12$ and $t = 1, \dots, 88$.

The DMIA states that conditional on a given unemployment experience and a vector of observed covariates, the sequence of potential outcomes associated with not receiving the treatment in a particular month is mean independent of the treatment status in this month. In a dynamic context, not receiving the treatment in the current month entails the possibility of participation in later months. Our matching approach will produce valid estimates if we consider all the deter-

minants that jointly influence treatment status (i.e., voucher award) and potential outcomes. Conditional on these determinants, individuals are randomly allocated to receiving a voucher or not in a given month of elapsed unemployment duration, and the treated and non-treated have the same predictions of future treatment or employment chances. We argue in the following that these assumptions are plausible in light of a voucher assignment in Germany and the rich information in our data. Our approach controls for pre-treatment outcome variables which are supposed to control for unobserved permanent characteristics that influence the post-treatment outcome. At the same time, we account for the dynamic selection of the eligible unemployed by controlling for elapsed unemployment duration.²²

The literature (e.g., Heckman et al. 1999, and Mueser et al. 2007 with regard to US programs, and Biewen et al. 2014, and Lechner and Wunsch, 2013, with regard to German training programs) stresses the importance of conditioning flexibly on lagged employment and wages, benefit receipt history, basic personal characteristics and local labor market characteristics. These pieces of information are all available in our data, and we use them in a flexible way. The literature addresses the plausibility of the conditional independence assumption (CIA, which is the static counterpart of DMIA) with regard to directly assigning a training program, but we believe that the award of a voucher to be used for a training program involves a similar selection process, which is perhaps less demanding with regard to the CIA because the actual start of the program is not part of this selection. Although training participation was mandatory under the old system in Germany, some unemployed may have talked the caseworker into not assigning

²²Eligibility for treatment hinges on being unemployed. Furthermore, we only consider treatments during the first twelve months of an unemployment spell after a non-negligible employment spell. Our data involve precise information on the timing of the inflow into unemployment and on the elapsed duration of unemployment. A large part of the U.S. literature uses difference-in-differences estimation approaches to control for selection due to unobserved permanent characteristics (see e.g. Heckman et al. 1999, Mueser et al. 2007). This requires a common trends assumption, which seems questionable, especially regarding employment as an outcome variable, because eligibility for treatment requires being unemployed. Heckman and Smith (1999) discuss difference-in-differences estimation in the presence of Ashenfelter’s dip, i.e. a disproportionate decline of earnings shortly before treatment start. They recommend the use of a symmetric difference-in-differences estimation approach, thus assuming that the counterfactual dynamics for the treated after treatment start relative to the comparison group would be symmetric to the outcomes observed before treatment start. Such an assumption is not plausible in our case, because eligibility requires being unemployed before treatment start.

a program or who did not begin the program despite the requirement. Such cases are demanding for the CIA and do not have to be accounted for in our case. Our data allow us to control for the full labor market history of the previous seven years and on important local labor market characteristics. In their sensitivity analysis, Biewen et al. (2014) find that it is very important to exactly match on the elapsed unemployment duration in months. This is implemented in the present paper by the dynamic approach. Note that the award of a voucher is left to the discretion of the caseworker; thus, from the perspective of the unemployed, the receipt of a voucher cannot be perfectly anticipated. Moreover, the data involve pieces of information that are collected by the caseworker as a basis for his counseling activities and assignment decisions (see also Biewen et al. 2014). To be specific, we consider the following variables that reflect part of the caseworker’s information on the motivation, plans and labor market prospects of a particular unemployed individual: The caseworker’s assessment of the job-seeker’s potential health problems, information on his/her previous health problems (during the previous 6 years before the start of the current unemployment spell), a dummy variable indicating whether the unemployed person appeared to lack motivation (e.g., failed to attend regular meetings), dummies indicating whether the job-seeker dropped out of a program, whether benefits were withdrawn, and whether the person participated in a program providing psychosocial support, where all variables refer to the previous 3 years unless stated otherwise. In addition, we include the employment and welfare history as sequences of the previous 7 years before the start of the current unemployment spell and variables indicating whether the job-seeker is looking for a part-time job.

The common support assumption ii) requires that it is possible in large samples to identify for each treated observation some comparable non-treated comparison observations. We apply some simple support tests (see discussion in Lechner and Strittmatter, 2014). Table 6 shows descriptive statistics for the propensity score. The largest conditional treatment probability is 12.6%. Further, we show the density distribution of the propensity score for the treated and comparison group separately in Figure 10. These results show that the propensity distribu-

tion does not exhibit fat tails. There is no apparent violation of the common support assumption. Given Assumption 1,

$$E[Y_{imt}^0 | D_{im} = 1] = E \left[\frac{(1 - D_{im}) \cdot p(X_{im})}{Pr(D_{im} = 1) \cdot (1 - p(X_{im}))} \cdot Y_{imt} \right],$$

is identified from the observed data on $\{Y_{imt}, D_{im}, X_{im}\}$ (Hirano, Imbens, and Ridder, 2003). For estimation, we use inverse probability weighting (IPW) and ordinary least squares (OLS). For both approaches, we perform exact matching on the elapsed unemployment duration and the duration since the award of the voucher. Thus, we align treated individuals and controls by the elapsed unemployment duration, and we only compare individuals who are still unemployed at the time of the treatment start. Taking IPW as a benchmark, we specify our parametric OLS regressions to allow for sufficient flexibility.

4.2 Estimation Strategy

Asymptotic theory suggests that IPW has some efficiency advantage in comparison to matching estimators in large samples (Heckman, Ichimura, and Todd, 1997, Hirano, Imbens, and Ridder, 2003). Moreover, recent simulation studies support this result (Busso, DiNardo, and McCrary, 2014). Concerning the reweighting technique, we follow the suggestions of Busso, DiNardo, and McCrary (2014) and use weights that sum up to one as a small sample correction. The average effect for the treated is estimated by

$$\hat{\gamma}_{mt} = \sum_{i=1}^N \frac{D_{im}}{\sum_{i=1}^N D_{im}} \cdot Y_{imt} - \sum_{i=1}^N \frac{(1 - D_{im}) \cdot \frac{\hat{p}(X_{im})}{1 - \hat{p}(X_{im})}}{\sum_{i=1}^N \frac{(1 - D_{im}) \cdot \hat{p}(X_{im})}{1 - \hat{p}(X_{im})}} \cdot Y_{imt},$$

where $t = 1, \dots, 88$ indicates the time after treatment and $m = 1, \dots, 12$ indicates the elapsed unemployment duration until treatment. The propensity score $p(X_{im})$ is specified as a probit model. We perform different balancing tests to ensure that the treated and non-treated are well matched with respect to observable characteristics (see Appendix B for details).

Although IPW has some optimality properties, some critical issues may arise. First, the IPW estimators for the average treatment effect for the treated may exhibit fat tails when the treatment probability is close to one. However, the treatment probability in our application is far below one. Second, the implementation of the IPW estimator relies on the estimation of an appropriate specification for the treatment probability (we rely on probit estimates). To demonstrate that our results are robust and not driven by specific issues with one estimator, we contrast the IPW estimates with the estimates obtained by a very flexible OLS regression. Although the implicit parametric assumptions may not hold, OLS might provide a good estimate of the average treatment effects.²³ Because nearly all of the control variables in this study are binary (excluding the earnings history and regional characteristics), our model is very flexible. We find that OLS leads to qualitatively and quantitatively very similar results to those of IPW. As a further robustness check, we also provide effect estimates based on radius-matching, which is the matching method performing best on average in the recent Monte Carlo study by Huber, Lechner, and Wunsch (2013). Our detailed analysis of heterogeneous treatment effects will rely on the OLS estimates.

5 Results

We first discuss the OLS and IPW estimates of the average treatment effect for the treated. Then, we investigate the heterogeneity of the treatment effects across skill groups and across the type of training programs based on the OLS estimates. Finally, we decompose the effect estimates by whether the treated actually redeem the training voucher. Inference is based on a bootstrap clustered at the individual level, thus resampling all observations over time for an individual. Calculating all estimates based on the same resample allows us to test for differences between different estimators.

²³ Angrist and Pischke (2009) suggest that OLS results often do not differ substantially from results obtained by more demanding non-parametric or semi-parametric estimators in many cases. In particular, they emphasize that the OLS finds exactly the conditional expectation function in fully saturated models, thus providing the non-parametric estimates for such a case.

5.1 Average Treatment Effect for the Treated

This section discusses the estimated average effects of a voucher award on employment and earnings based on OLS and IPW. We provide graphical evidence on the descriptive average differences between the treated and the non-treated, and on the estimated average treatment effects for the treated. As explained above, we estimate separately the effect of treatment versus waiting for each of the first twelve months of elapsed unemployment durations. We only report the average over these twelve months (further month-specific results are available upon request). On the time axis, we depict the months since the voucher receipt, and on the vertical axis, the outcome variable is depicted. Diamonds indicate a significant effect for the corresponding month. In each figure, the results for the employment (earnings) outcome are placed to the top (bottom).

Figure 1 depicts the descriptive (unconditional) differences between the treated and nontreated (top line) together with the average treatment effects based on different estimators (OLS and IPW). The OLS and IPW results imply a long and pronounced lock-in effect (the lines for IPW and OLS results are very similar which makes it impossible to distinguish the two lines in Figure 1). It takes approximately 40 months until the negative effect reaches zero for the employment outcome and the effect on earnings remains slightly negative even much longer: The lock-in effect is much longer than what is typically found in studies for Germany (see, e.g., Biewen et al. 2014 or Rinne et al. 2013). However, these studies restrict their sample to participants in long-term training and do not consider the much longer degree courses, and the treatment start is defined by the actual start of the training program. At the end of our observation period of seven years after the award of the voucher, the OLS results imply a very small positive and significant treatment effect (approximately 1-2 percentage points - henceforth, ppoints) for employment. The effect for earnings remains negative even 88 months after the treatment. The results obtained from using IPW are basically the same as those obtained using OLS. This finding suggests that we use sufficient flexibility in our specification of the OLS regression.²⁴

²⁴As a further robustness check, we also provide effect estimates based on radius-matching. Figure

Figure 1 indicates that there are strong changes in the slopes of the treatment effect at approximately 12 to 14, 24 to 26, and 36 to 38 months. This finding can be explained by the fact that many programs have a duration of 12, 24 or 36 months and that the majority of treated individuals enter training within the first two months after receiving the voucher (Figure 3). Furthermore, the long lock-in effect is rationalized by the fact that about 20 percent of the treated remain in training almost to the end of the second year after treatment and about 10 percent almost to the end of the third year (Figure 3). Figure 4 displays the average employment and average earnings for treated individuals under treatment and under non-treatment (using the weights of the IPW estimation). Employment under non-treatment is higher than under treatment for the first 3 years after treatment. It takes 40 months after treatment until the employment effect becomes significantly positive, and it remains positive until month 88.

The descriptive difference in Figure 1 involves a shorter and less pronounced lock-in effect than that of the OLS estimates both for employment and earnings. As discussed in Section 3, the treated are clearly a positive selection of the unemployed with regard to their labor market chances. Their labor market history is better, with less unemployment experience and higher earnings in the past; they hold higher schooling degrees, suffer less from health problems and less sanctions and are less likely to have dropped out of programs. This positive selection corresponds to the requirement of awarding vouchers only to those unemployed individuals who are expected to have at least a 70% chance of entering employment soon after the program. The comparison group for the descriptive effect has average characteristics and will thus have a lower employment rate than the matched comparison group (see column 4 in Tables 1 to 4 for the average characteristics of the matched comparison group). Because the treated are unemployed individuals with relatively good labor market chances, many of them would have found a job in the short or medium run, if they had not been treated.

2 involves a comparison of the three estimators. We find that the estimated treatment effects based on radius-matching are also very close to the IPW and the OLS estimates. Since all three estimation approaches rely on a selection-on-observables identification strategy and, at the same time, they make different functional form assumptions for the implementation of the estimator, it is reassuring that all three estimators basically yield the same treatment effects.

In sum, the results so far imply that a voucher award leads to a strong and very long negative lock-in effect which are explained by the sizeable participation in programs with a very long duration. It takes four years after the voucher award to find small, significantly positive and persistent employment effects. There are no positive effects on earnings within the observation period. Different estimators (OLS and IPW) based on a selection on observables assumptions basically provide the same results. Raw employment differences indicate that with regard to observables, voucher recipients represent a strong positive selection with respect to both outcomes. Altogether, our findings are consistent with cream-skimming by the caseworkers. This seems undesirable because many of the voucher recipients would have found a job much sooner anyway, if they had not received a voucher, and the long-term effects are not positive at all or they are not sufficiently large to compensate for the lock-in period.

5.2 Heterogeneous Effects by Skill Level

The mostly negative average treatment effects reported so far may hide heterogeneous treatment effects, which for some subgroups may even be significantly positive. Now, we investigate the differences in effect estimates by skill level. We focus on the OLS results, and additionally, we refer to the descriptive differences. We first investigate effect heterogeneity by vocational degree.²⁵ One may be concerned that low-skilled individuals may not cope well with a voucher award. They may not find the best training provider, they may not redeem the voucher, or they may be more easily discouraged during participation. However, they may gain significantly by a major investment in their human capital and by obtaining a course certificate or even a vocational degree. Of the treated in our sample, 21% do not hold a vocational degree (low-skilled individuals). Of the treated, 11% are high-skilled, holding an academic degree. The majority of the treated hold a vocational degree (medium-skilled). The top line in Figure 5 depicts the effect of a voucher award for the group of those without a vocational

²⁵We have also looked into effect heterogeneity by gender. The effects of the voucher are quite similar for men and women. Women face a slightly less deep lock-in effect, and the effect estimates are slightly more positive at the end of the observation period.

degree. The lock-in effects last for approximately three to four years. During the fourth to seventh year after the award of the voucher, we find a significant positive employment effect of about 5-6 ppoints and a significant positive earnings effect of about 110 euro. In contrast, the effect for the high-skilled is strongly negative over the whole observation period, and there is only a small positive effect on the employment probability for the medium-skilled.

Can we say more as to why only low-skilled individuals benefit on average? A potential explanation would be that the low-skilled show a shorter lock-in effect because they redeem the vouchers at a lower rate. In our sample, this is not the case: 21.4% of those individuals who redeem the voucher hold no vocational degree, and the share is approximately the same (21.1%) among those who do not redeem the voucher. Furthermore, the average time spent in a training program (conditional on redeeming the voucher) is 14 months for the low-skilled and 8 for the high-skilled. Thus, shorter courses or early dropout do not explain a shorter lock-in period. Furthermore, from month 8 to month 22, the employment effects for the low-skilled are almost parallel to those of the medium-skilled, with a stronger lock-in effect in the levels for the medium-skilled. After month 24, the line for the low-skilled increases more rapidly. This is the time at which the participants in the longer courses complete their courses and search intensively for jobs. Note that low-skilled individuals participate more often in degree courses (44% as opposed to 22% among the medium-skilled), and participants in a degree course spend on average two years in their course. Hence, participants in degree courses (after a quick redemption of the voucher) re-enter the labor market with their new degree approximately 25 to 36 months after the voucher award, and Figure 5 indicates the strongest increase for the low-skilled during that time. These results suggest that the low-skilled voucher recipients eventually do better in finding a job compared to the medium-skilled. Substantiating this finding, Figure 6 displays the employment rates of the treated and matched comparison individuals by skill level. After 36 months, the treated low-skilled exhibit nearly the same employment rate as the treated with a higher skill level. In contrast, the matched low-skilled comparison individuals exhibit a much lower employment

rate than the matched comparison individuals for the two other skill groups.

The effect heterogeneity by skill group seems to be stronger under the voucher regime than under the assignment regime before 2003, and the voucher award is more effective for the low-skilled. This may be surprising, as there was concern that in particular the low-skilled may be overstrained by finding a suitable program. Rinne et al. (2011) and Biewen et al. (2014) find little evidence for effect heterogeneity by skill level for long-term training in the pre-reform period.²⁶ With regard to degree programs, there exists relatively little prior evidence, because to look beyond the lock-in effect of these very long programs, one needs an observation period of at least three or four years. A series of studies using data from the 1990s are an exception, as they have an extraordinarily long period to observe the labor market outcomes of up to eight years. These studies find positive employment effects for the long retraining program, which is closest to the degree courses investigated in this paper (see Fitzenberger and Völter, 2007, Fitzenberger, Osikominu, and Völter, 2008, Lechner, Miquel, and Wunsch, 2007). In line with our findings, Lechner, Miquel, and Wunsch (2011) estimate the largest positive effects for low-skilled women without a vocational degree. For the U.S., Heinrich et al. (2013) find more positive results for the WIA program for all services as well as for training in particular under the Adult program than for the Dislocated Worker program. Participants in the Adult program are more negatively selected than in the Dislocated Worker program. In line with our findings, the effectiveness of training declines with a better selection of the participants' group.

5.3 Heterogeneous Effects by Type of Training

In light of the above results, we now investigate the two types of training programs: Long-term training and degree courses.²⁷ We decompose the average treatment effect on the treated presented in section 5.1 into three parts: The

²⁶Biewen et al. (2014) report a slightly more positive effect of long-term training for low-skilled males who start their program during months 4 to 6 of the unemployment spell (see the online appendix of Biewen et al. 2014).

²⁷Here, we do not consider some very special programs.

contribution of those individuals who participate in long-term training, the contribution of those who participate in a degree course, and the contribution of those who do not redeem the voucher. Here, we focus on the two components of the decomposition associated with the type of training, and the next section discusses the decomposition by the redemption decision. We implement the decompositions by running OLS regressions based on the sample of those individuals considered (e.g. those who participate in long-term training) and those not being awarded with a voucher (i.e. the same comparison group as used above). These decompositions would yield causal estimates, if the observables controlled for at the time of the voucher award also accounted for the selection into voucher redemption. But dynamic selection after the voucher award is likely to occur: E.g., those who quickly find a job after the voucher award end up being in the group of those not redeeming the voucher. Defining treatment as participation in (a particular type of) training hence conditions on outcomes that occur after the voucher award. Thus, it is not possible to estimate a causal effect of receiving a voucher and redeeming it based on aligning treated and comparison individuals at the time of the voucher award. It is also not possible to estimate the causal effect of participating in a particular type of training this way, because participating in (a particular type of) training requires redeeming the voucher.²⁸ Therefore, the OLS estimates presented here just represent a statistical decomposition of the average effect of being awarded with a voucher by type of training and by the redemption decision. The results are nevertheless suggestive, because they show the channels through which the causal effect of the voucher award operates.

Tables 1 to 3 indicate that participants in degree courses are younger, more likely to be female and unemployed, with lower formal education, and earn lower wages in the recent past than participants in long-term training. Degree courses have typically a very long duration. For participants in degree courses, it is thus not surprising that we find long and very deep lock-in effects of more than 3 years, reducing the employment probability by nearly 36 ppoints and earnings by 600

²⁸To estimate the causal effect of e.g. receiving a voucher and redeeming it to participate in long-term training would require an approach which models all intermediate employment outcomes as used in Fitzenberger et al. (2010).

euros per month. However, after 44 months the employment gain is +8 ppoints, and earnings gains are relatively large with over 100 euros per month (Figure 7). These positive effects remain stable until the end of the observation period. Thus, degree courses involve high costs due to a long and deep lock-in period, but after three to four years they increase the long-run labor market chances in a considerable way. The vocational degrees obtained in these courses are comparable to degrees obtained within the German apprenticeship system. Thus, the amount of marketable human capital investment provided by these courses appears to be very high.²⁹ Considering long-term training programs, we find a pronounced lock-in period of approximately 12 months. This lock-in period is comparable to Rinne et al. (2013). However, after this pronounced lock-in period, the estimated effects on the employment probability are slightly positive but remain negative for monthly earnings for the whole observation period although the effect size is reduced over time. In contrast to our results, Rinne et al. (2013) find a positive employment effect of approximately 7 ppoints at the end of their observation period of 1.5 years after the program start. In Rinne et al. (2013), those who do not redeem a voucher are members of the comparison group and are likely to form good matches to control for selection. Furthermore, the alignment between the treated and comparison individuals in Rinne et al. refers to the start of participation in the training program, when a number of individuals who were comparable at the time of the voucher award (among them, some of those who did not redeem a voucher) may have found a job in the meantime and are thus excluded from the comparison group. This may induce an upward bias in the effect estimates. A comparison of the effect estimates for long-term training and degree courses with the descriptive differences between participants and the comparison group (not shown) indicates that the difference with the descriptive effect is stronger for long-term courses than for degree courses. This suggests that the effect of cream-skimming is stronger for long-term training. Correspondingly, a comparison of the characteristics of the comparison group to the treatment group of the degree courses and to the treatment group of long-term training

²⁹According to the European Qualifications Framework (EQF) degree courses with a duration of two years are classified in level 2 and three year courses are classified in level 4.

(the last two columns in Table 1 to 3) also suggests that the positive selection on observables is somewhat stronger for long-term training.

When discussing the results on effect heterogeneity by skill group, we have suggested that the positive employment effects for the low-skilled may result from those low-skilled who participate in degree courses. Table 2 confirms that a higher share of participants in degree courses is low skilled (36.3%) compared to long-term training (15.6%). Furthermore, degree courses generally exhibit more positive long-run effects than long-term training. Shedding further light on these findings, Figure 8 decomposes the voucher effect by skill level and by type of training. In degree courses, we find at least small positive employment effects after 40 months for all skill levels. We also find positive effects after 10 months for the low-skilled in long-term training, and the highest positive effect materializes for the low-skilled in degree courses. Positive earnings effects can be found for the low-skilled participating in both types of training and for the medium-skilled taking degree courses. Thus, participants in degree courses rather benefit than those in long-term training and the voucher is generally more effective for the low-skilled. In contrast, those medium- and high-skilled who participate in long-term training do not benefit on average from a voucher award.

5.4 Unredeemed Vouchers

The award of a voucher may have an effect by allowing the individual to participate in a training program, but it may also have an effect on the labor market outcomes themselves. Figure 9 shows the contribution to the overall average effect estimates by the redemption decision. These OLS estimates by redemption decision reflect the outcome differences in comparison to the corresponding comparison group of unemployed not being awarded with a voucher.³⁰ Note that these OLS estimates do not allow for a causal interpretation because the redemption decision itself is endogenous (see discussion above).

³⁰The estimates displayed in Figure 9 contrast the outcome by redemption decision to the outcome of comparison observations similar in the observable characteristics considered for the average treatment effects for the treated discussed in subsection 5.1. The average treatment effects for the treated are the weighted averages of the two contributions shown in Figure 9, where the weights are the sample shares by redemption decision.

Individuals who redeem their vouchers (this group comprises 78% of all the treated receiving a voucher) exhibit the same pattern as for the effect for all treated. However, both the positive and the negative effect estimates are slightly more pronounced. Individuals who do not redeem the voucher first show a small peak and then a sizeable negative effect over three to four years. The peak at the beginning may be related to those individuals who do not redeem a voucher because they happen to find a job quickly. Because not redeeming a voucher is not sanctioned, some unemployed with a training voucher may just enjoy their unemployment benefits for three months without being pushed to find a job (note that these are not the ones who find a job quickly). For these individuals, employment chances may have deteriorated over time because of reduced search effort right after the voucher award. Further potential reasons for the negative effects for those not redeeming the voucher shown in Figure 9 are the following: First, those who do not redeem the voucher may participate in other programs. Second, those who do not succeed in finding a training course may suffer from a loss in motivation. Even though we can not estimate the causal effects of the redemption decision, our descriptive decomposition in Figure 9 nevertheless suggests that the average long-run effects of actual training participation are better than the effects of a voucher award.

6 Conclusions

This paper estimates the long-run effect of the award of a training voucher on employment and earnings for the unemployed in Germany. We use rich administrative data on all training vouchers awarded in 2003 and 2004 and on participation in training programs after the redemption of the voucher. We estimate the average effect of a voucher award in a flexible way by OLS and by inverse probability weighting (IPW) as alternatives to control for selection on observables.

Our results imply that the award of a training voucher on average has strong and lasting negative lock-in effects. It takes four years after the voucher award to produce small but long-lasting, significantly positive employment effects. There

are no positive effects on earnings during the observation period. The two methods based on selection on observables assumptions (IPW and OLS) lead to nearly the same results. These OLS and IPW estimates basically do not differ significantly from IV estimates, which we obtain in an unreported sensitivity analysis exploiting the unexplained variation in differences in policy styles across regional employment offices.

A comparison to raw employment differences indicates that with regard to observables, voucher recipients represent a strong positive selection both regarding employment and earnings. The strong positive selection effects implied by our estimates are consistent with sizeable cream-skimming effects. Our evidence strongly questions the general effectiveness of awarding training vouchers during the years 2003 and 2004. Our results draw a more negative picture on vouchers than the large scale econometric evaluation of WIA by Heinrich et al. (2013). They find large positive employment and earnings effects of participating in WIA as opposed to not participating for the Adult program and positive employment effects but only small and insignificant earnings effects, for the Dislocated Worker program. When they estimate the effects of receiving training through an ITA as opposed to receiving only the other services of the WIA, they find again positive long-run earnings and employment effects for the Adult program but no positive effects for the Dislocated Worker program.

An investigation of effect heterogeneity by skill group and by type of training suggests a more positive picture for some subgroups and a more negative one for others: Individuals without a vocational degree are more successful in finding a job after training than higher skilled individuals and the voucher leads to considerable positive long-run effects. This suggests that targeting vouchers to job-seekers without a vocational degree would be beneficial on average, provided a similar group of job seekers in terms of other covariates is used. Both the finding that the voucher recipients represent a positive selection and the finding that job-seekers without a vocational degree are those benefitting from a voucher award suggest that the abolishment of the 70% rule (indicating that only those individuals should be awarded with a voucher who are predicted to have an em-

ployment probability of at least 70% after the program) was the right decision. Further research should investigate further the impact of the 70% rule making use of its abolishment in 2005. However, such an analysis is complicated by the changes in the certification of training providers in 2005.

A statistical decomposition by program type suggests that, despite strong lock-in effects, programs leading to a vocational degree work better than those that do not. The strongest positive effects are found for individuals without a vocational degree participating in degree courses. This provides suggestive evidence that the targeting could be improved by reducing those vouchers that are not suitable for programs leading to a vocational degree and to award the majority of vouchers suitable to obtain a vocational degree to those who do not have yet obtained a first vocational degree.

Finally, a statistical decomposition by the redemption decision suggests that those treated, who do not redeem the voucher, are worse off over almost the entire observation period in comparison to those comparable individuals who do not receive a voucher. 22% of those job-seekers being awarded with a voucher do not redeem it. Another 7% of individuals who have redeemed the voucher do not finish the program but drop out before having finished 80% of the program. Before the introduction of the voucher system the dropout rate in further training programs was 20% (Paul 2015). We conclude that under the voucher system less people drop out of a program they have already started but they rather do not redeem the voucher. Dropout is not necessarily harmful (Paul 2015), but it is less costly if those who would later drop out of the program without benefitting from it would not even start it. Thus, the possibility to not redeem the voucher may be beneficial and if this is the case it should not be restricted. Furthermore, our results provide suggestive evidence that some of those who do not redeem the voucher suffer from letting too much time pass during which they focus on unsuccessful program search or enjoy not being pressured by their caseworker before they eventually focus on job search again or enter a different program. It might be beneficial to introduce another meeting with the caseworker if the prospects of voucher redemption seem to be low. But such a meeting would

contradict the goal that after a voucher award the job seeker should be allowed to make independent choices and the meeting would be costly for those who cope well with the freedom of choice.

Overall, there are only small positive employment effects and no gains in earnings even four to seven years after the voucher award. Even though most recipients use the voucher to participate in training, they are hardly better off on average in the long-run than if they had not been awarded with a voucher. They suffer from a lock-in effect that seems to be particularly pronounced due to the strong positive selection of voucher recipients. There are two notable exceptions to these findings: Those without a vocational degree and those who participate in degree courses apparently show significant and economically sizeable long-run benefits of training. Our results suggest that a reduction of voucher awards for unemployed with good labor market prospects and a better targeting of training vouchers would be useful.

References

- ANGRIST, J., AND J. PISCHKE (2009): *Mostly Harmless Econometrics*. Princeton University Press.
- BARNOW, B. (2000): “Vouchers for Federal Training Programs,” in *Vouchers and the Provision of Public Services*, ed. by C. Steuerle, V. Doorn Ooms, G. Peterson, and R. Reischauer. Brookings Institution Press, Washington, D.C.
- BARNOW, B. (2009): “Vouchers in US Vocational Training Programs: An Overview of What We have Learned,” *Journal of Labor Market Research (Zeitschrift für ArbeitsmarktForschung)*, 42, 71–84.
- BAUMOL, W. (1982): “Contestable Markets: An Uprising in the Theory of Industry Structure,” *American Economic Review*, 72(1), 1–15.
- BIEWEN, M., B. FITZENBERGER, A. OSIKOMINU, AND M. PAUL (2014): “The Effectiveness of Public Sponsored Training Revisited: The Importance of Data and Methodological Choices,” *Journal of Labor Economics*, 32(4), 837–897.

- BUSSO, M., J. DINARDO, AND J. MCCRARY (2014): “New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators,” *Review of Economics and Statistics*, 96, 885–897.
- CARD, D., J. KLUVE, AND A. WEBER (2010): “Active Labour Market Policy Evaluations: A Meta-Analysis,” *The Economics Journal*, 120, F452–F477.
- FITZENBERGER, B., A. OSIKOMINU, AND M. PAUL (2010): “The Heterogeneous Effects of Training Incidence and Duration on Labor Market Transitions,” *IZA Discussion Paper*, 5269.
- FITZENBERGER, B., A. OSIKOMINU, AND R. VÖLTER (2008): “Get Training or Wait? Long Run Employment Effects of Training Programs for the Unemployed in West Germany,” *Annales d’Economie et de Statistique*, 91-92, 321–355.
- FITZENBERGER, B., AND R. VÖLTER (2007): “Long-run Effects of Training Programs for the Unemployed in East Germany,” *Labour Economics*, 14, 730–755.
- FREDERIKSSON, P., AND P. JOHANSSON (2008): “Dynamic Treatment Assignment – The Consequences for Evaluations Using Observational Studies,” *Journal of Business Economics and Statistics*, 26, 435–445.
- GÖRLITZ, K. (2010): “The Effect of Subsidizing Continuous Training Investments – Evidence from German Establishment Data,” *Labour Economics*, 17 (5), 789–798.
- HECKMAN, J., H. ICHIMURA, AND P. TODD (1997): “Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme,” *Review of Economic Studies*, 64(4), 605–654.
- HECKMAN, J., R. LALONDE, AND J. SMITH (1999): “The Economics and Econometrics of Active Labor Market Programs,” in *Handbook of Labor Economics*, Vol. 3, ed. by O. Ashenfelter, and D. Card, pp. 1865–2086. North Holland.

- HECKMAN, J., AND J. SMITH (1999): “The Pre-Program Earnings Dip and the Determinants of Participation in a Social Program: Implications for Simple Program Evaluation Strategies,” *Economic Journal*, 108, 313–348.
- HEINRICH, C., AND P. MUESER (2014): “Training Program Impacts and the Onset of the Great Recession,” *Unpublished Manuscript, University of Missouri*.
- HEINRICH, C., P. MUESER, K. TROSKE, K. JEON, AND D. KAHVECIOGLU (2013): “Do Public Employment and Training Programs Work?,” *IZA Journal of Labor Economics*, 2(6), 1–23.
- HIPP, L., AND M. E. WARNER (2008): “Market Forces for the Unemployed? Training Vouchers in Germany and the USA,” *Social Policy and Administration*, 42, 77–101.
- HIRANO, K., G. IMBENS, AND G. RIDDER (2003): “Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score,” *Econometrica*, 71(4), 1161–1189.
- HUBER, M., M. LECHNER, AND C. WUNSCH (2013): “The Performance of Estimators Based on the Propensity Score,” *Journal of Econometrics*, 175(1), 1–21.
- HUJER, R., S. THOMSEN, AND C. ZEISS (2006): “The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany,” *Allgemeines Statistisches Archiv*, 90, 299–322.
- KING, C. T., AND B. S. BARNOW (2011): “The Use of Market Mechanisms,” in *The Workforce Investment Act: Implementation Experiences and Evaluation Findings*, ed. by D. J. Besharov, and P. H., pp. 81–111. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
- LADD, H. F. (2002): “School Vouchers: A Critical View,” *Journal of Economic Perspectives*, 16(4), 3–24.

- LECHNER, M., R. MIQUEL, AND C. WUNSCH (2007): “The Curse and the Blessing of Training the Unemployed in a Changing Economy: The Case of East Germany after Unification,” *German Economic Review*, 8, 468–509.
- (2011): “Long- run Effects of Public Sector Sponsored Training,” *The Journal of the European Economic Association*, 9, 742–784.
- LECHNER, M., AND A. STRITTMATTER (2014): “Practical Procedures to Deal with Common Support Problems in Matching Estimation,” *SEPS Discussion Paper*, 2014-10.
- MCCONNELL, S., P. DECKER, AND I. PEREZ-JOHNSON (2011): “The Role of Counseling in Voucher Programs: Findings from the Individual Training Account Experiment,” *Unpublished Manuscript, Mathematica Policy Research, Princeton, NJ*.
- MUESER, P., K. TROSKE, AND A. GORISLAVSKY (2007): “Using State Administrative Data to Measure Program Performance,” *Review of Economics and Statistics*, 89, 761–783.
- PAUL, M. (2015): “Many Dropouts? Never Mind! - Employment Prospects of Dropouts from Training Programs,” *Annals of Economics and Statistics*, forthcoming.
- PEREZ-JOHNSON, I., Q. MOORE, AND R. SANTILLANO (2011): “Improving the Effectiveness of Individual Training Accounts: Long-Term Findings from an Experimental Evaluation of Three Service Delivery Models,” *Final Report, Mathematica Policy Research, Princeton, NJ*.
- POSNER, P., R. YETVIN, M. SCHNEIDERMAN, C. SPIRO, AND A. BARNETT (2000): “A Survey of Voucher Use: Variations and Common Elements,” in *Vouchers and the Provision of Public Services*, ed. by C. Steuerle, V. Doorn Ooms, G. Peterson, and R. Reischauer, pp. 503–539. Brookings Institution Press, Washington, D.C.

- RINNE, U., M. SCHNEIDER, AND A. UHLENDORFF (2011): “Do the skilled and prime-aged Unemployed benefit more from Training? Effect Heterogeneity of Public Training Programmes in Germany,” *Applied Economics*, 43, 3465–3494.
- RINNE, U., A. UHLENDORFF, AND Z. ZHAO (2013): “Vouchers and Caseworkers in Public Training Programs: Evidence from the Hartz Reform in Germany,” *Empirical Economics*, 45(3), 1089–1127.
- SCHNEIDER, H., K. BRENKE, B. JESKE, L. KAISER, U. RINNE, M. SCHNEIDER, J. STEINWEDE, AND A. UHLENDORFF (2007): “Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission, Bericht 2006,” *IZA Research Report*, 10.
- SIANESI, B. (2004): “An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s,” *The Review of Economics and Statistics*, 86, 133–155.
- SMITH, J., AND P. TODD (2005): “Does Matching Overcome LaLonde’s Critique of Nonexperimental Methods?,” *Journal of Econometrics*, 125(1-2), 305–353.

A Averaging across Starting Dates

Following a dynamic treatment evaluation approach (Sianesi, 2004, Frederiksson and Johansson, 2008), we estimate the effect of a voucher award versus waiting for each of the first twelve months of the unemployment period separately. In the first month, the treatment group includes only individuals who are awarded with a training voucher during the first month. Individuals who either receive a voucher later or never are in the comparison group. In the second month, we drop all individuals who have left the risk set in the first month, i.e., received a voucher or found employment in the first month. The treatment group in the second month consists of voucher recipients that are awarded with a voucher in their second month of the unemployment period. Everybody in the risk set who does not receive a voucher in the second month belongs to the comparison group. This procedure continues until month twelve. By using this dynamic approach, we end up with twelve different treatment effects for each of the twelve different times of elapsed unemployment duration. To communicate our results, we reduce the dimension of the results by reporting a weighted average of the twelve dynamic treatment effects in the following. The weights are calculated as the fraction of treated in the respective month of the total number of treated individuals

$$\hat{\gamma}_t = \frac{\sum_{m=1}^M \sum_{i=1}^N D_{im} \cdot \hat{\gamma}_{mt}}{\sum_{m=1}^M \sum_{i=1}^N D_{im}}.$$

Given that we observe the labor market outcomes of each individual for 48 months after treatment ($t = 1, \dots, 48$), we specify a separate model for each month after treatment ($m = 1, \dots, 12$). This induces flexibility in all parameters with respect to the duration since treatment.

B Matching Quality

We assess the matching quality by displaying the means of the matched comparison group for different control variables in Tables 1-3. Further, we report the standardized differences before and after matching. The standardized differences are defined as

$$SD = \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{0.5(\sigma_{X_1}^2 + \sigma_{X_2}^2)}} \cdot 100,$$

where $\bar{X}_d$ is the mean and $\sigma_{X_d}^2$ the variance in the respective treatment group $d \in \{0, 1\}$. Before matching, we observe standardized differences larger than 40. After matching, the standardized differences are always below one, suggesting a very good matching quality.

We also apply a second balancing test following Smith and Todd (2005). Therefore, we run the regression

$$x_k = \hat{\beta}_0 + \hat{\beta}_1 D_{im} + \hat{\beta}_2 \hat{p}(X_{im}) + \hat{\beta}_3 D_{im} \hat{p}(X_{im}) + \hat{\varepsilon}_{im},$$

where x_k indicates the specific control variable. We perform a joint F-test for the null hypothesis that $\hat{\beta}_1$ and $\hat{\beta}_3$ equal zero. In Table 5, we report the summarized results of the test for each of the twelve treatment times. Overall, we run 1,272 regressions, of which the test indicates a rejection of the null hypothesis in only 77 cases. We take the results of the assessment as an indication that the propensity score is well balanced and acceptable for the performance of the IPW estimations. Because we control directly for X_{im} in the OLS regressions, it is not necessary to assume that the propensity score is balanced for these estimators.

Figures and Tables

Figure 1: Effect of a voucher award on employment and earnings averaged over elapsed unemployment durations until treatment.

(a) Effects on employment

(b) Effects on monthly earnings (in Euros)

Note: Diamonds indicate significant effects. Note that the OLS and IPW effect estimates almost coincide.

Figure 2: Robustness Check: Comparison of OLS, IPW, and Radius-Matching

(a) Effects on employment

(b) Effects on monthly earnings (in Euros)

Note: Point estimates are presented without standard errors. Effect estimates using OLS and IPW as in Figure 1.

Figure 3: Fraction of individuals in training after the award of a voucher.

Figure 4: Comparison of average employment and average earnings between treatment and matched comparison group averaged over elapsed unemployment durations until treatment.

(a) Employment Probability

(b) Monthly earnings (in Euros)

Figure 5: Heterogeneous effects on employment and earnings by skill group (OLS) averaged over elapsed unemployment durations until treatment.

(a) Effects on employment

(b) Effects on monthly earnings (in Euros)

Note: Diamonds indicate significant effects.

Figure 6: Comparison of average employment of treated and matched comparison group by skill group averaged over elapsed unemployment durations until treatment.

(a) Employment Probability Treated Group

(b) Employment Probability Matched Comparison Group

Note: Diamonds indicate significant effects.

Figure 7: Decomposition of the Average Effects of Treatment on the Treated on employment and earnings with regard to the type of training (OLS) averaged over elapsed unemployment durations until treatment

(a) Effects on employment

(b) Effects on monthly earnings (in Euros)

Note: Diamonds indicate significant effects.

Figure 8: Decomposition of the Average Effects of Treatment on the Treated on employment and earnings with regard to the type of training and the vocational degree (OLS) averaged over elapsed unemployment durations until treatment.

(a) Effects on employment

(b) Effects on monthly earnings (in Euros)

Note: Diamonds indicate significant effects.

Figure 9: Decomposition of the Average Effects of Treatment on the Treated on employment and earnings with regard to the redemption decision (OLS) averaged over elapsed unemployment durations until treatment.

(a) Effects on employment

(b) Effects on monthly earnings (in Euros)

Note: Diamonds indicate significant effects.

Figure 10: Density plot of the propensity score for treatment and comparison group by month of the elapsed unemployment duration at voucher award.

(a) Elapsed unemployment duration 1 month.

(b) Elapsed unemployment duration 2 months.

(c) Elapsed unemployment duration 3 month.

(d) Elapsed unemployment duration 4 months.

(e) Elapsed unemployment duration 5 month.

(f) Elapsed unemployment duration 6 months.

< Figure continues on next page >

Figure 10 continued:

(g) Elapsed unemployment duration 7 months.

(h) Elapsed unemployment duration 8 months.

(i) Elapsed unemployment duration 9 month.

(j) Elapsed unemployment duration 10 months.

(k) Elapsed unemployment duration 11 month.

(l) Elapsed unemployment duration 12 months.

Table 1: Means and Standardized Differences (SD) for Personal Characteristics

	Treatment group	Comparison group	SD before matching	Matched comparison group	SD after matching	Voucher redeemed	Voucher expired	Degree courses	Long-term courses
Female	.4539	.4245	6.84	.4531	.19	.4529	.4566	.5044	.4215
Age									
25-29 years	0.148	0.162	3.940	0.149	0.360	0.146	0.153	0.219	0.121
30-34 years	0.186	0.181	2.010	0.186	0.140	0.186	0.186	0.248	0.169
35-39 years	0.232	0.208	5.790	0.233	0.170	0.234	0.227	0.249	0.228
45-49 years	0.147	0.150	2.220	0.146	0.190	0.147	0.148	0.077	0.169
50-54 years	0.074	0.107	11.600	0.074	0.170	0.073	0.076	0.017	0.092
Nationality									
Germany	0.944	0.932	5.110	0.944	0.130	0.945	0.940	0.930	0.951
Outside EU	0.0004	0.0004	1.590	0.0004	-	0.0004	0.001	0.001	0.0003
Missing	0.031	0.042	5.950	0.031	0.110	0.030	0.034	0.037	0.0269
Marital status									
Single	0.327	0.330	3.350	0.327	0.220	0.322	0.346	0.282	.341
Single parent	0.074	0.060	5.390	0.074	0.090	0.075	0.072	0.105	.063
Married	0.484	0.490	2.740	0.484	0.120	0.490	0.463	0.470	.499
Missing	0.071	0.071	3.570	0.071	0.220	0.070	0.075	0.092	.053
Child	0.380	0.372	2.600	0.380	0.130	0.386	0.361	0.444	.368
Age of youngest child									
One year	0.011	0.012	2.020	0.011	0.130	0.011	0.011	0.014	0.010
Between 1 and 3 years	0.034	0.032	2.180	0.035	0.120	0.035	0.033	0.041	0.034
Between 3 and 6 years	0.066	0.063	1.980	0.066	0.110	0.067	0.062	0.087	0.061
Between 6 and 10 years	0.084	0.077	2.960	0.084	0.050	0.085	0.082	0.106	0.077
Older than 14 years	0.099	0.107	2.610	0.098	0.130	0.100	0.095	0.097	0.102
Missing	0.622	0.630	2.670	0.622	0.130	0.616	0.640	0.557	0.634
Disabled	0.021	0.025	2.530	0.021	0.120	0.020	0.026	0.007	0.025
Health									
Health problems	0.097	0.118	6.780	0.097	0.160	0.095	0.107	0.081	0.100
Health problems before unemployment	0.043	0.052	4.070	0.043	0.090	0.042	0.047	0.034	0.0443
N	45,287	87,641				35,249	10,038	9,017	22,199

Omitted categories:

Age: 40-44 years

Nationality: Member EU

Marital status: Common law marriage

Age of youngest child: Between 10 and 14 years

Table 2: Means and Standardized Differences (SD) for Education, Occupation, and Sector

	Treatment group	Comparison group	SD before matching	Matched comparison group	SD after matching	Voucher redeemed	Voucher expired	Degree courses	Long-term courses
Education									
No schooling degree	0.040	0.065	11.420	0.039	0.100	0.040	0.039	0.045	0.037
University entry degree	0.218	0.184	8.720	0.220	0.300	0.220	0.212	0.157	0.259
Missing	0.012	0.014	2.300	0.012	0.100	0.012	0.014	0.016	0.011
Vocational training									
No vocational degree	0.213	0.232	8.390	0.212	0.240	0.214	0.211	0.357	0.156
Academic degree	0.106	0.094	3.940	0.107	0.350	0.107	0.101	0.049	0.140
Missing	0.010	0.012	2.140	0.010	0.120	0.010	0.011	0.013	0.009
Classification of occupation									
Agriculture, energy, water and miners	0.270	0.267	2.040	0.270	0.070	0.271	0.267	0.252	0.275
Construction	0.050	0.036	6.700	0.050	0.090	0.050	0.049	0.025	0.066
Trade and retail	0.137	0.163	7.220	0.137	0.140	0.137	0.137	0.197	0.117
Communication and information service	0.110	0.116	1.920	0.110	0.070	0.109	0.115	0.130	0.104
Part-time work									
Full-time	0.816	0.803	7.740	0.817	0.210	0.818	0.811	0.788	0.842
Missing	0.061	0.071	4.230	0.061	0.230	0.059	0.066	0.068	0.052
Part-time work desired									
Not desired	0.842	0.838	3.900	0.843	0.200	0.844	0.838	0.837	0.861
Missing	0.069	0.070	3.640	0.069	0.230	0.068	0.074	0.089	0.052
Type of work									
White-collar	0.477	0.390	17.590	0.478	0.200	0.475	0.486	0.337	0.535
Missing	0.108	0.107	5.800	0.108	0.090	0.110	0.102	0.138	0.093
Azubi									
	0.028	0.019	13.130	0.027	0.230	0.029	0.024	0.046	0.011
Sector									
Agriculture	0.009	0.014	5.290	0.009	0.080	0.009	0.008	0.012	0.009
Mining	0.002	0.002	1.270	0.002	0.080	0.002	0.001	0.001	0.002
Utilities	0.002	0.002	1.310	0.002	0.130	0.002	0.002	0.001	0.002
Construction	0.068	0.097	10.550	0.068	0.140	0.069	0.066	0.059	0.074
Trade	0.150	0.133	4.800	0.150	0.110	0.148	0.157	0.139	0.152
Hotels and restaurants	0.028	0.037	5.200	0.028	0.090	0.027	0.032	0.038	0.023
Traffic, transportation	0.054	0.055	1.640	0.054	0.120	0.054	0.053	0.064	0.052
Financial services	0.020	0.013	5.310	0.020	0.130	0.020	0.020	0.014	0.022
Renting	0.011	0.009	1.870	0.011	0.050	0.011	0.010	0.006	0.013
Data processing	0.141	0.123	5.280	0.141	0.160	0.140	0.144	0.089	0.165
Public sector, education	0.057	0.060	4.270	0.057	0.160	0.056	0.059	0.061	0.056
Health and social service	0.075	0.071	14.490	0.075	0.180	0.076	0.070	0.140	0.042
Other services	0.040	0.042	2.050	0.040	0.120	0.041	0.039	0.050	0.039
Temporary employment	0.133	0.168	12.540	0.134	0.270	0.132	0.136	0.143	0.129
N	45,287	87,641				35,249	10,038	9,017	22,199

Omitted categories:

Education: Schooling degree without Abitur

Vocational training: Vocational degree

Classification of occupation: Other

Part-time work: Part-time

Part-time work desired: Desired

Type of work: Blue-collar

Sector: Production

Table 3: Means and Standardized Differences (SD) for Employment/Unemployment/ALMP History

	Treatment group	Comparison group	SD before matching	Matched comparison group	SD after matching	Voucher redeemed	Voucher expired	Degree courses	Long-term courses
Noticeable problems									
Problem group	0.018	0.025	4.620	0.018	0.110	0.018	0.018	0.016	0.020
Sanction	0.011	0.029	12.950	0.011	0.100	0.011	0.012	0.014	0.009
Lack of motivation	0.011	0.029	12.950	0.011	0.100	0.011	0.012	0.014	0.009
Incapacity	0.041	0.078	15.780	0.041	0.140	0.040	0.044	0.044	0.040
Dropout	0.014	0.052	21.740	0.014	0.170	0.013	0.015	0.017	0.011
Employment history (last 7 years), Sequences (1 for employed, 0 for unemployed)									
Mostly employed in last period (i.e., 1111000, 1101000, 1000101)									
Mostly unemployed (i.e., 1000010)	0.167	0.218	12.920	0.168	0.300	0.166	0.171	0.226	0.147
3 years employed, close (i.e., 1111010)	0.125	0.102	7.020	0.125	0.100	0.124	0.126	0.127	0.122
3 years employed, far (i.e., 1100111)	0.026	0.051	13.230	0.026	0.110	0.026	0.028	0.025	0.026
3 years unemployed, close (i.e., 1000011)	0.012	0.023	8.280	0.012	0.080	0.012	0.011	0.012	0.013
3 years unemployed, far (i.e., 1101000)	0.098	0.089	3.020	0.098	0.170	0.098	0.096	0.108	0.095
Mixed employment (i.e., 1101101)	0.048	0.060	5.170	0.049	0.110	0.048	0.048	0.053	0.047
Mostly unemployed in last period (i.e., 0111000, 0101000, 0000101)									
Mostly employed (i.e., 0101101)	0.015	0.028	9.330	0.015	0.070	0.015	0.015	0.013	0.015
3 years employed, close (i.e., 0111001)	0.004	0.006	2.580	0.004	0.050	0.004	0.005	0.005	0.003
3 years employed, far (i.e., 0100111)	0.001	0.003	5.250	0.001		0.001	0.001	0.001	0.001
Program history (last 3 years), sequences									
Often in programs	0.012	0.030	12.760	0.012	0.160	0.012	0.012	0.013	0.011
No programs	0.909	0.792	33.100	0.908	0.220	0.909	0.908	0.903	0.908
History of wages while employed (measured as average daily wages)									
Real wage (t-1)	67.408	60.027	24.090	67.453	0.170	67.262	67.950	57.934	71.447
Real wage (t-2)	61.130	49.635	32.190	61.192	0.190	60.972	61.704	50.270	65.475
Real wage (t-3)	55.068	45.383	25.090	55.001	0.240	55.037	55.195	43.804	59.580
Real wage (t-4)	50.139	43.875	16.010	50.007	0.330	50.086	50.365	39.364	54.467
Real wage (t-5)	45.670	40.534	13.030	45.575	0.240	45.597	45.967	35.044	49.816
Real wage (t-6)	42.595	37.379	13.430	42.501	0.240	42.537	42.836	32.275	46.681
Real wage (t-7)	40.970	35.730	13.540	40.865	0.260	40.801	41.608	30.425	44.945
N	45,287	87,641				35,249	10,038	9,017	22,199

Omitted categories:

Mostly employed in last period: Mostly Employed

Mostly unemployed in last period: 3 years unemployed (far) and mixed employment

History of programs (last 3 years): Seldom in programs

Table 4: Means and Standardized Differences (SD) for Regional Characteristics

	Treatment group	Comparison group	SMD before matching	Matched comparison group	SMD after matching	Voucher redeemed	Voucher expired	Degree courses	Long-term courses
Unemployment and population									
Unemployment rate	12.284	12.694	7.867	12.302	0.355	12.314	12.185	12.914	12.449
Share of male unemployed	0.564	0.562	7.034	0.564	0.204	0.564	0.567	0.561	0.565
Share of German unemployed	0.859	0.867	8.851	0.859	0.316	0.861	0.854	0.872	0.858
Share of vacant fulltime jobs	0.793	0.790	4.864	0.793	0.192	0.793	0.794	0.789	0.793
Population per km^2	620.732	606.853	3.096	621.593	0.145	590.511	727.340	549.349	661.950
Industries									
Management of forests and agriculture	0.012	0.013	11.602	0.012	0.317	0.012	0.011	0.013	0.012
Fishing	0.000	0.000	6.307	0.000	0.177	0.000	0.000	0.000	0.000
Mining	0.005	0.005	3.966	0.005	0.177	0.005	0.005	0.005	0.005
Energy and water supply	0.010	0.010	2.814	0.010	0.229	0.010	0.010	0.010	0.010
Construction	0.064	0.066	9.267	0.064	0.259	0.065	0.063	0.067	0.064
Trade	0.149	0.150	3.726	0.149	0.122	0.150	0.149	0.149	0.150
Hotels and restaurants	0.028	0.028	2.551	0.028	0.118	0.028	0.028	0.029	0.028
Transport and communications	0.056	0.057	7.893	0.056	0.340	0.056	0.056	0.057	0.056
Bank and insurance business	0.038	0.037	4.893	0.038	0.187	0.038	0.038	0.037	0.038
Real estate activities	0.118	0.117	4.180	0.118	0.227	0.117	0.120	0.115	0.120
Public administration and defense	0.065	0.067	8.478	0.065	0.185	0.065	0.066	0.067	0.066
Education	0.041	0.042	7.591	0.041	0.358	0.041	0.041	0.042	0.041
Healthcare and social sector	0.118	0.117	3.235	0.118	0.129	0.117	0.118	0.118	0.118
Services	0.047	0.047	3.838	0.047	0.234	0.047	0.048	0.047	0.048
Production at the household level	0.001	0.001	10.010	0.001	0.414	0.001	0.001	0.001	0.001
Extraterritorial organizations and bodies	0.001	0.001	2.677	0.001	0.279	0.001	0.001	0.001	0.001
Other	0.000	0.000	6.030	0.000	0.116	0.000	0.000	0.000	0.000
N	45,287	87,641				35,249	10,038	9,017	22,199

Omitted categories:

Industries: Manufacturing industry

Table 5: Balancing Test (Smith and Todd, 2005).

Elapsed Unempl. Duration (in months)	Weighted Obs	Treated	Number of Parameters	# sign.
1	2,350,543	7,439	106	6
2	2,232,379	4,008	105	10
3	2,050,877	4,206	106	10
4	1,892,323	4,173	106	7
5	1,766,314	4,094	106	3
6	1,667,045	3,926	106	6
7	1,575,298	3,673	104	9
8	1,500,706	3,439	106	7
9	1,430,171	3,141	106	7
10	1,364,425	2,883	106	6
11	1,306,412	2,422	105	3
12	1,247,830	1,883	105	3
			1,272	77

Note: For each control variable and period of the elapsed unemployment duration we apply a separate balancing test, following (Smith and Todd, 2005). The last column indicates for how many control variables the balancing test is rejected, i.e. for how many control variables we find statistical evidence for an unbalance. See detailed description in Appendix B.

Table 6: Descriptive statistic of the propensity score.

Elapsed Unempl. Duration (in months)	Mean	Std. Dev.	Min	Max
1	0.004	0.005	1.14E-05	0.126
2	0.002	0.001	2.10E-05	0.011
3	0.002	0.001	1.63E-05	0.011
4	0.002	0.002	2.04E-06	0.014
5	0.003	0.002	2.30E-06	0.016
6	0.003	0.002	3.81E-06	0.016
7	0.003	0.002	9.01E-06	0.015
8	0.003	0.002	2.43E-05	0.019
9	0.003	0.002	1.52E-05	0.018
10	0.002	0.002	6.88E-06	0.019
11	0.002	0.002	7.56E-06	0.022
12	0.002	0.001	4.40E-06	0.019
All	0.003	0.002	2.04E-06	0.126

Note: This table shows the descriptive statistics of the propensity score for each month of the elapsed unemployment duration separately.