

Direct and indirect effects of training vouchers for the unemployed

Martin Huber⁺, Michael Lechner^x, and Anthony Strittmatter^{x*}

⁺University of Fribourg, ^xUniversity of St. Gallen, SEW

Abstract: This paper evaluates the effects of awarding vouchers for vocational training on the employment outcomes of unemployed voucher recipients in Germany, as well as the potential mechanism through which they operate. This study assesses the direct effects of voucher assignment net of actual redemption, which may be driven by preference shaping and learning about possible human capital investments or simply by the costs of information gathering. Using a formal mediation analysis framework based on sequential conditional independence assumptions and semiparametric matching estimators, our results suggest that the negative short-term and positive long-term employment effects of receiving a voucher are mainly driven by actual training participation. However, the direct effect of just obtaining a voucher is negative over the short-run as well. This result points to potential losses in the effectiveness of such training provision systems if individuals decide not to redeem vouchers, as employment chances are lower than under non-award over the short-run and under redemption over the long-run, which makes non-redemption the least attractive option.

Keywords: Mediation analysis, voucher award, training programmes, direct effects, indirect effects, causal mechanisms, causal channels, matching estimation

JEL classification: J64, J68, C21, C31.

Addresses for correspondence: Martin Huber (Martin.Huber@unifr.ch, unifr.ch/appecon/en), Department of Economics, University of Fribourg, Bd. de Pérolles 90, CH-1700 Fribourg; Michael Lechner (Michael.Lechner@unisg.ch, www.michael-lechner.eu), Anthony Strittmatter (Anthony.Strittmatter@unisg.ch, www.anthonystrittmatter.com), Swiss Institute for Empirical Economic Research (SEW), University of St. Gallen, Varnbuelstrasse 14, CH-9000 St. Gallen.

* Michael Lechner is also affiliated with CEPR, London, CESifo, Munich, IAB, Nuremberg, and IZA, Bonn. Anthony Strittmatter is also affiliated with the Albert-Ludwigs-University Freiburg. We benefitted from comments by Annabelle Doerr, Bernd Fitzenberger, Thomas Kruppe, and Gesine Stephan. Financial support from the Swiss National Science Foundation, grant number SNF 100018_137769 for the project “Identifying causal mechanisms in empirical economics” is gratefully acknowledged. This study is also part of the project “Regional Allocation Intensities, Effectiveness and Reform Effects of Training Vouchers in Active Labor Market Policies”, IAB project number 1155. We gratefully acknowledge data support from the IAB. Anthony Strittmatter acknowledges financial support from the IAB. The usual disclaimer applies.

1 Introduction

In January 2003, the German Federal Employment Agency reformed the allocation of vocational training programmes, which are a corner stone of Germany's active labour market policies (ALMPs). An assignment system based on vouchers replaced the direct assignment of unemployed individuals to vocational training by caseworkers. The vouchers certify eligibility for fully funded vocational training. The aims of the reform were to increase the involvement of training participants in the training decision and to increase competition among training providers. Before the reform, caseworkers determined placement into vocational training. After the reform, unemployed workers receive vouchers and select, with a few restrictions, both training providers and courses. The redemption decision is made solely by the awardee. Caseworkers are not allowed to influence the decision of awardees.

This paper contributes to the literature estimating the effects of publicly sponsored training programmes. This is a mature literature and many high-level studies exist on this topic, see, e.g., the meta analyses by Card, Kluve, and Weber (2010, 2015). Our contribution to this literature is to investigate the labour market effects of vocational training, as well as the potential mechanisms through which these effects might operate. Specifically, we investigate whether, in addition to the voucher's impact through its redemption (i.e., participation in vocational training), there exists a direct effect of the voucher award (i.e., without participation). The latter may be driven by motivational effects, preference shaping, and salience/awareness regarding (the availability of) ALMPs, which could affect labour market behaviour.

Such effects would be in line with Van den Berg, Bergemann, and Caliendo (2009), who document that the subjective individual expectation to be assigned to an ALMP affects the job search behaviour. Further, Crépon, Ferracci, Jolivet, and Van den Berg (2014) report that the mere notification of a planned training assignment has effects on the unemployment

exit probability. In a related endeavour, Crossley, de Bresser, Delaney, and Winter (2014) document that mere exposure to specific questions in a survey can alter subsequent behaviour. Exploiting randomized assignment to survey modules within the LISS panel survey, they find that households answering questions on expenditures and needs in retirement significantly changed their non-housing saving rates. The authors argue that the survey may have acted as a salience shock and comparable impacts might occur from being offered a training voucher, as this might induce individuals to become aware of and reflect upon the ALMPs and their expected impacts.

For instance, the award of a voucher could increase awareness of and preferences for possibilities to build up human capital and therefore immediately reduce job search intensity. The same effect occurs if awarding a voucher leads to high information costs of looking for appropriate providers and courses. On the other hand, if participation in ALMPs is perceived as a burden or as ineffective in raising employability or if a voucher award increases the salience/awareness of potential obligations to participate in future (unattractive) ALMPs, then an immediate increase in job search intensity might be expected. Therefore, it appears to be an interesting and open issue whether the impact of receiving a voucher is solely rooted in its actual use or whether a direct effect whose direction is a priori ambiguous exists as well. In particular, this allows us to determine whether it is the quality of the training providers that drive the voucher effect (through voucher redemption), or whether other dimensions are also important.

We use a formal mediation framework (see, for instance, the seminal paper by Baron and Kenny, 1986) to identify these specific causal mechanisms and, to this end, consider the redemption of a voucher as a mediator, i.e., an intermediate outcome on the causal path from the voucher award to the individual labour market outcomes. In addition to the effect of voucher redemption, we are particularly interested in the so-called controlled direct effect

(see, for instance, Pearl, 2001), i.e., the employment effect of a voucher award in the absence of actual redemption.¹ However, causal mechanisms are not easily identified. Even if the vouchers were randomly assigned, this would not imply the randomness of the mediator (see Robins and Greenland, 1992).²

To tackle the endogeneity of voucher award and redemption, a particular conditional independence assumption is invoked for identification. It requires (i) that the voucher award is independent of potential employment outcomes (under (non-)award and (non-)redemption of the vouchers) conditional on observed covariates and (ii) that voucher redemption is independent of the potential outcomes conditional on the covariates and voucher award. These assumptions are related to those invoked in the nonparametric mediation literature for identifying controlled direct effects (see, for instance, Petersen, Sinisi, and van der Laan, 2006, and VanderWeele, 2009). They also appear in the dynamic treatment effects literature on assessing sequences of treatments (see, for instance, Robins, 1986, 1989, Robins, Hernan, and Brumback, 2000, Lechner, 2009, and Lechner and Miquel, 2010), as well as in the multiple treatment effects framework (see Imbens, 2000, and Lechner, 2001). For estimation, we use semiparametric radius matching with bias adjustment (Lechner, Miquel, and Wunsch, 2011) based on the propensity score capturing the probability of (joint) voucher award and redemption conditional on covariates.

The results suggest that among voucher recipients, a voucher award has a negative average (total) employment effect in the first three years after voucher receipt. It has a small positive effect thereafter, with an increased employment probability of approximately 2 to 3 percentage points throughout the fourth year after receiving the voucher. Thus, the initial

¹ We refer to Pearl (2001) for a discussion of the differences between controlled and natural direct effects.

² This issue has also been discussed in the context of randomized training programmes; see, for instance, Ham and Lalonde (1996). They notice that even under the randomization of training, conditioning on a mediator such as employment introduces selection bias when assessing the effects on wages or duration outcomes (e.g. the length of a post-treatment employment spell).

negative lock-in effect of a voucher award (likely due to decreased job search intensity) is offset by higher employment probabilities in later periods. Concerning the causal mechanisms, voucher redemption (and thus, actual participation in vocational training) has similar, slightly more pronounced, negative short-term and positive long-term effects as voucher award. Therefore, voucher redemption primarily drives the total effect on voucher recipients.

In contrast, the direct effect on voucher recipients, i.e., the differences in mean potential outcomes between voucher award and non-redemption and non-award (and non-redemption), is small and insignificant during most of the fourth year. Therefore, over the long-run, mere voucher receipt does not affect employment (e.g., through a change in preferences). Nevertheless, a negative direct effect appears over the first three years, suggesting that voucher award decreases job search intensity despite non-redemption. This points to potential losses in effectiveness of the voucher award systems if individuals do not redeem their vouchers, as employment chances are lower than under non-award over the short-run and under redemption over the long-run. Therefore, voucher award and non-redemption appears to be the least attractive option. This finding is important when designing voucher award systems.

The main contribution of this study is to disentangle the causal mechanisms of an ALMP based on voucher awards rather than to merely assess the total (gross) effect of the programme. Therefore, this research goes beyond Doerr et al. (2016) and Heinrich, Mueser, Troske, Jeon, and Kahvecioglu (2010), who evaluate the effectiveness of various vocational training programmes in voucher systems but do not consider the direct effect of voucher award. It also differs from Doerr and Strittmatter (2016) and Rinne, Uhendorff, and Zhao (2013), who compare the effectiveness of vocational training via voucher and mandatory assignment regimes but do not separate award and redemption effects.

The remainder of this paper is organized as follows. Section 2 discusses the institutional background of voucher awards in Germany for ALMPs. Section 3 presents the econometric framework, namely, the definitions of the effects of interest, the identifying assumptions, and the estimator used. Section 4 introduces the data. In Section 5, we provide descriptive statistics and discuss the plausibility of the identifying assumptions. Section 6 presents the estimation results. Section 7 concludes. Appendices A-F provide further details on data, estimation, and results. All appendices are available online on the web pages of this journal.

2 Institutional background of voucher provision

Vocational training programmes constitute a corner stone of ALMPs in Germany. Their main objective is to adjust the skills of unemployed individuals to changing requirements of the labour market and/or to changing individual conditions. Essentially, there are three types of vocational training courses: classic vocational training, training in so-called practice firms, and retraining. Examples are courses in IT-based accounting or customer orientation and sales. Training in practice firms aims at simulating (real) work environments. Retraining courses have longer durations of up to three years with the goal of completing a vocational degree within the German apprenticeship system. They cover, for example, a full curriculum of a vocational training for an elderly care nurse. Vocational training is organized either in classrooms or on the job. The curriculum may also include internships. The exact course format depends on the type of training. Practice firm training typically takes place in artificial ‘training’ companies. Retraining for occupations within the dual-apprenticeship system (e.g., service managers) takes usually place on-the-job, but can also involve classroom sessions. Between 2000 and 2002, average annual expenditures for vocational training exceeded seven billion euros.³

³ Source: Labour Market Reports, Federal Employment Agency of Germany.

In January 2003, a voucher-based allocation system for the provision of vocational training was introduced. It aims at promoting the responsibility of training participants and introducing market mechanisms among training providers. Potential training participants receive vocational training vouchers, which allow them to choose training providers and courses. As explained in Doerr and Strittmatter (2014), several rules apply. First, the voucher specifies the objective, content, and maximum duration of the course. Second, it can only be redeemed within a one-day commuting zone.⁴ Third, the training vouchers are valid for a period ranging between one week and three months. Fourth, there are no sanctions or penalties imposed on the recipient (such as reduced unemployment benefits) for non-redemption.

The voucher award in the period considered was based on a statistical selection rule. Caseworkers were to award vouchers to unemployed workers who had at least a 70% probability of finding new employment within 6 months after finishing a training programme. Because our data are extracted from administrative records, we observe the individual characteristics and the regional labour market conditions that caseworkers used to predict employment chances. In addition, caseworkers had the opportunity to use information from mandatory counselling interviews and sometimes had access to test results from medical or psychological services (later, we control for motivation and health problems by means of proxy variables). Nevertheless, to predict employment outcome six months after training, particularly for training programmes with long durations, remains difficult. There were likely regional differences in the way predictions were formed. Doerr and Kruppe (2015) conduct a survey of caseworkers to analyse regional differences in voucher award intensities. They find that such differences can be (partly) explained by the preferences and sentiments of

⁴ For a training course lasting six or more hours per day, commuting times of up to 2.5 hours are reasonable. For a training course lasting less than six hours per day, the reasonable commuting time is reduced to two hours.

caseworkers and managers at local employment agencies regarding the use of training vouchers.

Caseworkers were not allowed to sanction unemployed workers who did not redeem vouchers. However, the unemployed worker had to provide a reasonable explanation for non-redemption. This might have caused a mental burden for the unemployed and may be one reason why the direct voucher effect on leaving the labour force is positive. Nevertheless, non-redeemers could receive a second voucher if they remained unemployed, but they did not have any legal claim. The award of a second voucher is based on the same selection rules as for the first voucher, but the final award decision is left to the discretion of the caseworkers. We do not analyse second vouchers because of this more involved dynamic selection procedure. In our data, we observe that 11% of redeemers and 22% of non-redeemers received a second voucher at a later time. For training sequences, caseworkers may awarded a voucher for each course or only a single voucher, which certifies eligibility for the whole sequences.

Crépon, Ferracci, Jolivet, and Van den Berg (2014) find a negative effect of a notification of possible training on unemployment exits. Because notified unemployed have to search for an appropriate course on their own, this notification has many similarities to a voucher award. However, Crépon, Ferracci, Jolivet, and Van den Berg (2014) analyse a French programme that differs in several dimensions from the German vocational training vouchers we analyse. In particular, the notified unemployed workers have to apply for funding for their training after receiving a notification. This differs from the German voucher system, which guarantees the payment of the training and extends the duration of unemployment benefit payments. The French system involves compulsory counselling with the caseworker every 6 months. Caseworkers might align the notified unemployed person to a training provider, which is not allowed in the German vocational training system. Finally, Crépon, Ferracci, Jolivet, and Van den Berg (2014) investigate the hazard rate for leaving

unemployment, while we focus directly on different employment, unemployment, and earnings outcomes.

3 Econometric framework

3.1 Potential outcomes and causal effects

Let D denote a binary indicator for voucher award, the so-called treatment variable, and Y the labour market outcome of interest, e.g. employment. Furthermore, let M be a binary indicator for voucher redemption (which implies participation in vocational training), which is supposedly the major mediator through which D affects Y . To define the effects of interest, we use the potential outcome framework (e.g., Rubin, 1974). Here Y^d denotes the potential outcome as a function of voucher award $d \in \{1,0\}$.⁵ The average treatment effect on the treated (ATET) of a voucher award is given by $\Delta = E(Y^1 - Y^0 | D=1)$. To investigate the distinct causal mechanisms, $Y^{d,m}$ denotes the potential outcome as a function of both voucher award and redemption, $d, m \in \{1,0\}$. Note that the two ways of denoting potential outcomes are linked: $Y^d = Y^{d,M^d}$, where M^d is the potential redemption state under voucher award $D = d$. Therefore, the ATET may be expressed as:

$$\Delta = E(Y^{1,M^1} - Y^{0,M^0} | D=1). \quad (1)$$

In our application, $M^0 = 0$ for everyone because vouchers cannot be redeemed if not awarded, so $\Delta = E(Y^{1,M^1} - Y^{0,0} | D=1)$. In contrast, M^1 might be either one or zero, depending on whether an individual redeems a received voucher. Thus, the ATET provides the total

⁵ By defining the potential outcomes this way, we implicitly impose the Stable Unit Treatment Value Assumption (SUTVA); see Rubin (1980).

effect of an award, which may operate indirectly through actual redemption (given that M^d changes with the value of d for at least some individuals) or directly without redemption.⁶

The extended notation allows further definitions of parameters, e.g., the average effect of voucher award and redemption vs. no award and no redemption among voucher recipients:

$$\theta = E(Y^{1,1} - Y^{0,0} \mid D = 1). \quad (2)$$

The difference to the ATET is that in (2), the redemption status is ‘forced’ to correspond to the voucher award status. This means that Y^{1,M^1} measures the potential outcome under voucher award and a mixture of redemption and non-redemption, while $Y^{1,1}$ measures the potential outcome under voucher award and redemption (i.e., voucher and redemption status are ‘forced’ to be equal for this potential outcome). Note that only in the special case of perfect compliance, i.e., everyone’s redemption decision corresponds to the voucher award (i.e., $M^d = d$ for $d \in \{1,0\}$), is θ equal to Δ . Again, part or all of the impact might be due to redemption or to a direct award effect. In the next step, we disentangle the latter two components and consider the so-called controlled direct effect (see, for instance, Pearl, 2001):⁷

$$\gamma = E(Y^{1,0} - Y^{0,0} \mid D = 1). \quad (3)$$

⁶ It is worth noting that the mediation framework has some resemblance to the literature on instrumental variables (IV), as an instrument affects the outcome via an endogenous intermediate variable, which is usually the actual treatment of interest. However, an important distinction is that in the IV context, a direct effect of the instrument on the outcome is ruled out through the exclusion restriction. IV may therefore be regarded as a special case of a mediation framework in which direct effects are excluded by assumption. Also note that the ‘intention to treat’ effect in the IV literature also corresponds to the total causal effect of the instrument in the mediation literature and to the IV’s indirect effect (as the total effect corresponds to the indirect effect in the absence of a direct one).

⁷ A related parameter is the so-called natural direct effect in the nomenclature of Pearl (2001) or the pure/total direct effect in the nomenclature of Robins and Greenland (1992) and Robins (2003), which is defined upon potential mediator states rather than prescribed mediator values: $E(Y^{1,M^1} - Y^{0,M^1} \mid D = 1)$, $E(Y^{1,M^0} - Y^{0,M^0} \mid D = 1)$. The latter two parameters and γ are equivalent only in the cases where there are no interaction effects between D and M on the outcome Y (such that the effect of M does not depend on D , and vice versa). Identification and estimation of natural direct effects have been considered in Pearl (2001), Robins (2003), Flores and Flores-Lagunes (2009), Imai, Keele, and Yamamoto (2010), and Huber (2014), among many others.

This is the impact of training voucher award among voucher recipients net of actual redemption, i.e., under prescribed non-redemption for everyone. Finally, the effect of redemption is identified by

$$\delta = E(Y^{1,1} - Y^{1,0} | D = 1). \quad (4)$$

Here, the effect of redemption vs. non-redemption is investigated conditional on awarding a voucher. Note that γ and δ sum up to θ , which can be seen by adding and subtracting $Y^{1,0}$ in the expectation of expression (2).

3.2 Identifying assumptions

To identify the effects of interest, we impose (sequential) conditional independence of the potential outcomes on the one hand and voucher award and redemption on the other hand (Assumptions 1 and 2 below). This requires that we observe all factors that are jointly related (i) with D and the potential outcomes and (ii) with M and the potential outcomes. We henceforth denote the vector of observed covariates by X . Furthermore, a particular common support restriction is needed (Assumption 3 below), implying that suitable comparisons in terms of X exist across various combinations of D and M .

Assumption 1: $\{Y^{1,1}, Y^{1,0}, Y^{0,0}\} \perp\!\!\!\perp x$ for all x in the support of X .

Assumption 1 states that the potential outcomes are jointly independent of a voucher award conditional on X . It rules out unobserved confounders that affect both the award and the outcome after controlling for the covariates. It is sufficient for identifying the ATET (in combination with the first part of Assumption 3 below). In contrast, the identification of θ , γ , and δ requires a further conditional independence assumption.

Assumption 2: $\{Y^{1,1}, Y^{1,0}, Y^{0,0}\} \perp\!\!\!\perp x, D = d$ for $d \in \{1, 0\}$ and all x in the support of X .

Under Assumption 2, redemption is independent of the potential outcomes conditional on the covariates and voucher award, which rules out unobserved confounders of the mediator and the outcome.

Assumptions 1 and 2 are closely related to conditions (4) and (5) in Petersen, Sinisi, and van der Laan (2006) for the identification of the controlled direct effect. They are also related to conditions (1) and (2) in VanderWeele (2009) for identifying the controlled direct effect, and conditions (a) and (b) of the Weak Dynamic Conditional Independence Assumption in Lechner (2009) and Lechner and Miquel (2010) evaluating dynamic treatments. The difference is, however, that the latter papers allow for different sets of covariates to control for confounding of D and M (where the covariates of M may be affected by D), whereas we (similarly to Petersen, Sinisi, and van der Laan, 2006) assume the same X for D and M . Further below we argue that this appears reasonable in our application (where D and M represent voucher award and redemption, respectively). The main reasons are (i) the informative set of observed characteristics available (see the discussion in Section 5.1), (ii) the small time lag between D and M , and (iii) the randomly assigned pseudo start dates for D and M among those with $D=0$ and $M=0$ or $D=1$ and $M=0$, respectively. The aim of the latter procedure is to control for any differences in elapsed unemployment durations (that likely affect Y) across redeemers, non-redeemers, and non-awardees, respectively. See Section 4.2 for further details.

Assumption 3: $\Pr(D = 1 | X = x) < 1$ and $0 < \Pr(M = 1 | D = 1, X = x) < 1$ for all x in the support of X .

The first part of Assumption 3 requires that no combination of covariates perfectly predicts a voucher award; otherwise, no comparable observations (in terms of conditioning variables X) without an award (and thus, without redemption) exist, implying that Δ , θ , and γ (which involve $Y^{0,0}$) cannot be identified. The second part requires that conditional on a voucher

award, no combination of X perfectly predicts redemption or non-redemption; otherwise, θ (which involves $Y^{1,1}$), γ (which involves $Y^{1,0}$), and δ (which involves both) are not identified.

Note that Assumptions 1 and 2 together imply the following conditional independence restriction: $\{Y^{1,1}, Y^{1,0}, Y^{0,0}\} \perp\!\!\!\perp X = x$ for all x in the support of X . Technically, the various combinations of D and M (despite their sequentially) may be treated as distinct treatments when identifying θ , γ , and δ by conditioning on X . Therefore, we can analyse the effects of the various treatment-mediator combinations in a standard multiple treatment effect framework, as outlined in Imbens (2000) and Lechner (2001). It follows that:

$$\begin{aligned} E(Y^{1,M^1} | D = 1) &= E(Y^1 | D = 1) = E(Y | D = 1), \\ E(Y^{0,M^0} | D = 1) &= E(Y^{0,0} | D = 1) = E(Y^0 | D = 1) = E_{X|D=1} [E(Y | D = 0, X = x)], \\ E(Y^{d,m} | D = 1) &= E_{X|D=1} [E(Y | D = d, M = m, X = x)], \end{aligned}$$

with $E_{A|B=b} [C]$ denoting the expectation of C taken over the distribution of A conditional on $B=b$.

The second and third lines are implied by Assumption 1 and Assumptions 1 and 2, respectively.⁸

However, directly controlling for a possibly high dimensional vector X when estimating

$$E_{X|D=1} [E(Y | D = 0, X = x)] \text{ and } E_{X|D=1} [E(Y | D = d, M = m, X = x)]$$

may lead to the curse of dimensionality. Rosenbaum and Rubin (1983) show that one may instead condition on the treatment propensity scores, in our case $p(x) = \Pr(D = 1 | X = x)$ and $p_{dm}(x) = \Pr(D = d, M = m | X = x)$, respectively, which balance the distributions of X . Therefore, it holds that

⁸ The derivation of these results is standard (e.g., Heckman, Ichimura, and Todd, 1998, Section 3) and omitted.

$$E(Y^{0,M^0} | D = 1) = \underset{p(X)|D=1}{E} \{E[Y | D = 0, p(X) = p(x)]\},$$

$$E(Y^{d,m} | D = 1) = \underset{p_{dm}(X)|D=1}{E} \{E[Y | D = d, M = m, p_{dm}(X) = p_{dm}(x)]\}.$$

This has the practical advantage that the vector of covariates consists of a single variable. This circumvents the curse of dimensionality if the propensity scores are well approximated by parametric probability models. The effects of interest are obtained by matching on estimates of $p(X)$ and $p_{dm}(X)$. Specifically, Δ is estimated by (i) matching to all voucher awardees comparison observations without voucher award that are similar in terms of estimates of $p(X)$ and (ii) taking the mean difference in outcomes between the two groups. The estimation of θ is based on two matching steps: First, to all awardees, one matches redeemers that are comparable in terms of estimates of $p_{11}(X)$. Second, to all awardees, one matches non-redeemers that are comparable in terms of estimates of $p_{00}(X)$. Taking the mean difference in outcomes between matched redeemers and non-redeemers yields the effect of interest. Analogous approaches are used for the estimation of γ and δ .

3.3 Estimation

Estimation of the various effects of interest (see Section 3.1) is based on radius matching on the propensity score with bias adjustment using the estimator of Lechner, Miquel, and Wunsch (2011). While the propensity scores, $p_x(X)$ and $p_{dm}(X)$, are parametrically specified by probit models, the conditional expectations of the outcomes are unrestricted and thus nonparametric. The algorithm is more precise than nearest-neighbour matching due to the idea of radius matching (e.g., Dehejia and Wahba, 2002). Furthermore, the procedure uses the initial matching weights for a (weighted) regression adjustment for bias reduction in a second step (see Abadie and Imbens, 2011). Therefore, the estimator satisfies a so-called double robustness property, implying that it is consistent if either the propensity score or the regression model is correctly specified (e.g., Rubin, 1979; Joffe et al., 2004). Moreover, the regression

adjustment should reduce small sample and asymptotic biases of matching. Huber, Lechner, and Wunsch (2013) investigate the finite sample properties of this algorithm along with other matching type estimators and find it to be very competitive.

We match on the linear index of the probit specification of the propensity score and use a data-driven approach to the choice of the radius size. That is, we set the latter to 90% of the 0.9th quantile of the distance between matched treated and control observations occurring in standard nearest-neighbour matching.⁹ Alternative radius sizes do not affect the results importantly (see Table D.1 of Appendix D). Inference is based on bootstrapping the respective effect 999 times and using the standard deviation of the bootstrapped effects as an estimate of the standard error of the t-statistic. Abadie and Imbens (2008) show that bootstrap-based standard errors may be invalid for matching based on a fixed number of comparison observations. However, our matching algorithm is smoother than the latter approach because it (by the nature of radius matching) uses a variable number of comparisons that are distance-weighted within the radius and, moreover, applies the regression adjustment. Therefore, the bootstrap is likely to be a valid inference procedure for the radius matching estimator used. It performs well in a large-scale (empirically based) simulation study by Bodory, Camponovo, Huber, and Lechner (2016), who investigate the performance of several variance estimators in the context of propensity score based matching estimation.

4 Empirical implementation

This section describes the data and the selection of our estimation sample.

⁹ If there is no comparison observation within the radius, then the nearest neighbour is matched.

4.1 Data

Our analysis is based on administrative data provided by the Federal Employment Agency of Germany, namely, the Integrated Employment Biographies (IEB).¹⁰ The latter contain information on all individuals in Germany who received a voucher between 2003 and 2004, along with subsequent participation in vocational training programmes. That is, the precise award and redemption dates for each voucher as well as the start and end dates of vocational trainings are observed. Furthermore, the data include detailed daily information on employment subject to social security contributions, receipt of transfer payments during unemployment, job search, and participation in various active labour market programmes (type, duration), as well as rich individual information (e.g., education, age, gender, marital status, profession, and nationality) and regional (labour market) characteristics. Thus, we are able to control for a wealth of personal characteristics and detailed labour market histories (e.g., type of employment, industry, occupational status, earnings) for all individuals receiving a voucher and thus capture the key confounders in such settings, as identified by Lechner and Wunsch (2013). Furthermore, we make use of a control sample of unemployed individuals without voucher awards during the years 2003 and 2004. The sample also originates from the IEB and is a three percent random sample of individuals who experience at least one transition from employment to non-employment (lasting at least one month) in 2003.¹¹

¹⁰ The IEB is a rich administrative database and has been used in virtually all recent studies on German ALMPs (e.g., Biewen, Fitzenberger, Osikominu, and Paul, 2014, Lechner, Miquel, and Wunsch, 2011, Lechner and Wunsch, 2013, Rinne, Uhlendorff, and Zhao, 2013). The IEB is a merged data file containing individual records collected by 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). IAB (Institut für Arbeitsmarkt- und Berufsforschung) is the German abbreviation for the research department of the Federal Employment Agency.

¹¹ We account for the different sampling probabilities whenever necessary using sampling weights. Note that these probabilities differ only for unemployed persons obtaining a voucher versus not obtaining a voucher. They are identical within the two groups.

4.2 Sample definition

The evaluation sample is an inflow sample into unemployment. It consists of individuals who became unemployed in 2003 after having been continuously employed for at least three months. Entering unemployment is defined as transitioning from (non-subsidised, non-marginal, non-seasonal) employment to registered non-employment for at least one month. We focus on individuals who are eligible for unemployment benefits at the time of inflow into unemployment. Thus, this sample focuses on the main target groups of these programmes. To exclude specific ALMPs targeting youths and individuals eligible for early retirement schemes, we consider only persons aged between 25 and 54 years at the beginning of the unemployment spell.

One concern regarding the award and redemption definition is the timing with respect to the elapsed unemployment duration prior to award and redemption of the voucher. The award and redemption decisions are dynamic processes. Caseworkers can award a voucher to an unemployed on any day after the start of the unemployment spell, as long as the caseworkers' client has not found employment. Awardees can redeem vouchers on any day of their validity, unless the awardee started a new job. Therefore, several issues must be taken into account.

Had the sample been large enough, an attractive approach would have been to use the dynamic evaluation framework (as suggested by Robins, 1986, Lechner, 2009, Miquel and Lechner, 2010), which would allow us to account both for the timing of the voucher award and the subsequent redemption. Alternatively, Abbring and van den Berg (2003, 2004) and Heckman and Navarro (2007) take the timing of program starts explicitly into account. However, given our small sample size (we observe 8,061 awardees with unredeemed vouchers), neither approach is feasible with our data at the desired level of flexibility. Furthermore, sample sizes are too small to follow the approach suggested by Fredriksson and Johansson (2008) and applied by Sianesi (2004) to estimate differential effects by the elapsed

unemployment duration. It is difficult to interpret these effects, because at any given duration, a substantial fraction of non-awardees or non-redeemers change their status shortly thereafter. Hence the estimated effects are mixtures of the true program effects and differences due to shifted timings of voucher award and redemption.

As a compromise that trades off the issues just raised and is feasible for our data, we consider a classical static evaluation model and use the following definitions. Awardees are those unemployed who receive their first voucher in the first 12 months of unemployment. Control group members are not awarded with a voucher during this time period. Redeemers are those awardees who redeem their voucher within the maximum validity of three months. Non-redeemers do not start vocational training during the first three months after the voucher award. Using these definitions within a static evaluation approach provides a more obvious - though not necessarily cleaner - definition of awardees and redeemers than the approaches advocated by Fredriksson and Johansson (2008) and Sianesi (2004).

Potentially this approach could lead to a higher share of individuals with better labour market characteristics among control group and non-redeemers than among the awardees and redeemers, because individuals in the control group and non-redeemers possibly found already a job prior to their potential award or redemption times (e.g., Fredriksson and Johansson, 2008). This would bias the results negatively. To check the sensitivity of our results, we randomly assign pseudo award dates to each individual in the control group. Thereby, we recover the distribution of the elapsed unemployment duration at the time of voucher award from the treatment group (similar to, e.g., Lechner, 1999, Lechner and Smith, 2007). To ensure comparability of the treatment definitions of the awardees and non-awardees, we only consider individuals who are unemployed at their (pseudo) voucher award. Following similar arguments, the same approach is applied with respect to the mediator in order to create (pseudo) voucher redemption dates among those who did not redeem a

voucher.¹² This makes the groups of individuals with redeemed and expired vouchers comparable with respect to the duration of unemployment.

In Figures 6.1 and 6.2 as well as Figures C.1-C.5 in Internet Appendix C, we show the results controlling for the elapsed unemployment duration until (pseudo) voucher award and (pseudo) redemption. In Figures E.1-E.7 in Internet Appendix E, we show the results without controlling for the elapsed unemployment duration. The results are not strongly affected by the omission of this variable. Most findings are qualitatively similar. Nevertheless, the results without controlling for the elapsed unemployment duration suggest overall lower returns to training. These findings are in line with the presumption that not accounting for the elapsed unemployment duration biases the results negatively (positively for unemployment).

4.3 Descriptive statistics

The baseline sample includes 93,016 (or 600,842 weighted) observations.¹³ 41,138 observations include an awarded voucher in 2003 or 2004, whereas 51,878 do not. Of the former group, 33,077 individuals redeem their voucher, whereas 8,061 do not. Table 4.1 reports the means of selected observed characteristics across groups defined in terms of treatment and mediator states (see Table A.1 for a more extensive set of variables): voucher awarded, no voucher awarded, voucher redeemed, voucher expired (note that the last two groups are subsamples of the first one). Pairwise standardized mean differences (see Ros-

¹² Note that 592 individuals with expired vouchers are dropped because of the definition of the pseudo voucher redemption dates.

¹³ The IAB provided a data set that contains 230,842 (or 3,638,851 weighted) observations. This sample is representative of the inflow into unemployed in 2003 and 2004 subject to the following sample restrictions: previous employment of at least 3 months, some contact with the employment agency within the first three months of unemployment, unemployment durations of at least one month, eligible for unemployment benefits, and aged between 25 and 54 years. We do not consider treatments after 2004 because in January 2005, a substantial labour market reform took place in Germany (the so-called Hartz IV reform). Thus, we restrict our sample to individuals who became unemployed in 2003. This enables us to consider for all unemployed persons a potential treatment within the first twelve months of their unemployment spell. Further, we drop individuals with marginal, seasonal, or subsidised employment before their last unemployment spell. This leaves 124,696 observations. Another 31,680 observations are dropped because of the definition of the pseudo voucher award and redemption times. See the descriptive statistics of the initial and final sample in Table A.2 in Appendix A.

enbaum and Rubin, 1985) are also shown as measures of covariate balance. Information on individual characteristics refers to the time of inflow into unemployment. Only for elapsed unemployment duration and remaining eligibility for unemployment benefits do we consider the measurements at the time of the (pseudo) voucher award.

The descriptive statistics in Table 4.1 reveal that voucher recipients (1) and non-recipients (2) differ importantly in several socio-economic characteristics, such as age, health, education, and profession. In particular, those awarded vouchers are younger, healthier, better educated, and have higher paying jobs. However, elapsed time in unemployment duration is higher for recipients, and accordingly, the remaining eligibility for unemployment benefits is lower. Regional differences are generally less pronounced.

When comparing samples of unemployed individuals who redeem vouchers (3) to those whose voucher expires (4), differences in socio-economic variables are small, with the important exception that the latter group is likely suffering from incapacities (and health problems in general), which may importantly drive non-redemption. Furthermore, while the employment histories are quite comparable, non-redeemers have higher elapsed unemployment durations and thus lower eligibility for unemployment benefits at the time of the voucher award than redeemers.

Figures 4.1 and 4.2 show the evolution of employment and registered unemployment over time after using radius matching (as outlined above) to adjust the covariate distributions of all groups to the respective distributions of voucher recipients.¹⁴ Over a horizon of 4 years (48 months) after voucher award, employment rates reach approximately 60% and registered unemployment falls below 20% for all groups. When comparing development across different groups, differences arise as a function of time. Over the short-run, the groups that partly or fully redeem vouchers appear to experience so-called lock-in effects, i.e., they take up fewer

¹⁴ Further outcome variables are presented in Appendix A.

jobs than non-recipients or non-redeemers. Over the longer run, this effect disappears, and the former group experiences higher employment than the group not participating in vocational training. The econometric analysis below will reveal how much of these differences are driven by merely obtaining a voucher and by actually redeeming it.

Table 4.1: Means and standardized biases of selected variables

	Voucher		... awarded		... redeemed		Comparisons of groups			
	yes	no	yes	no	yes	no	(1) - (2)	(1) - (3)	(1) - (4)	(3) - (4)
	(1)	(2)	(3)	(4)	Subsample means		Standardized differences			
Individual characteristics										
Age	39.03	41.75	39.01	39.11	31.44	0.27	1.10	1.38		
Children under 3 years	0.43	0.35	0.43	0.41	15.06	0.58	2.39	2.97		
Health problems	0.02	0.06	0.02	0.03	20.08	2.08	7.35	9.38		
Incapacities	0.12	0.19	0.11	0.17	19.53	3.70	13.75	17.43		
No German citizenship	0.07	0.10	0.07	0.08	10.35	0.60	2.40	3.00		
No schooling degree	0.04	0.07	0.04	0.04	13.74	0.33	1.34	1.67		
University entry degree (Abitur)	0.23	0.17	0.24	0.23	17.21	0.40	1.66	2.06		
Elementary occupation	0.07	0.10	0.07	0.07	11.15	0.21	0.87	1.09		
Craft, machine operators & related	0.29	0.35	0.29	0.28	13.60	0.37	1.52	1.89		
Clerks	0.25	0.16	0.25	0.25	23.08	0.02	0.10	0.12		
Individual labour market history										
Half months employ. in last 2 years	45.17	44.30	45.19	45.12	12.33	0.21	0.83	1.04		
Half months OLF in last 2 years	1.59	2.19	1.59	1.63	11.74	0.17	0.67	0.84		
Cumulative earnings in last 4 years	91'258	84'199	91'126	91'799	14.71	0.27	1.10	1.37		
Months of remaining UE benefits	8.90	10.95	9.14	7.92	31.44	4.09	16.42	20.54		
Elapsed unemployment duration	4.46	3.76	4.19	5.56	21.04	8.08	30.56	38.81		
Regional characteristics										
Share of employment in construction	0.06	0.06	0.06	0.06	7.17	2.64	11.20	13.84		
Share of vacant full-time jobs	0.78	0.78	0.78	0.77	3.94	2.84	10.63	13.46		
Population per km ²	965	868	919	1156	5.72	2.69	10.02	12.69		
Unemployment rate (in %)	12.33	12.53	12.33	12.36	3.69	0.13	0.52	0.65		
Observations	41'138	51'878	33'077	8'061						
Sum of weighted observations	41'138	559'704	33'077	8'061						

Note: See Rosenbaum and Rubin (1985) for a definition of the standardized difference. They consider an absolute standardized difference of more than 20 as being 'large'. The full set of results is contained in Table A.1 in Appendix A. The dummy variable 'health problems' indicates disabilities and health problems which do not fully prevent the unemployed to work, but may reduce the number of possible working hours or the number of possible jobs. The dummy variable 'incapacities' indicates pregnancies, necessarily of medical rehabilitation, or other the incapability of work because of other reasons.

Figure 4.1: Mean employment

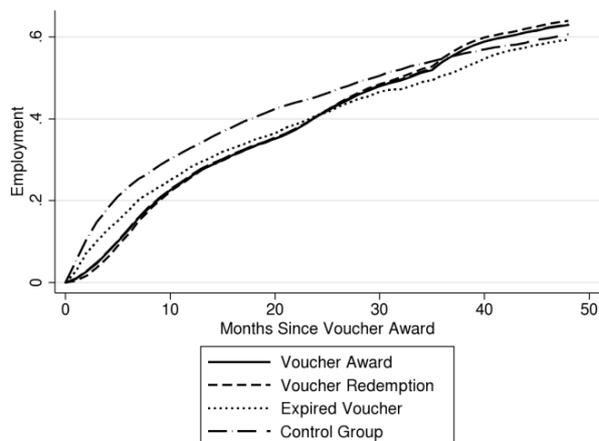
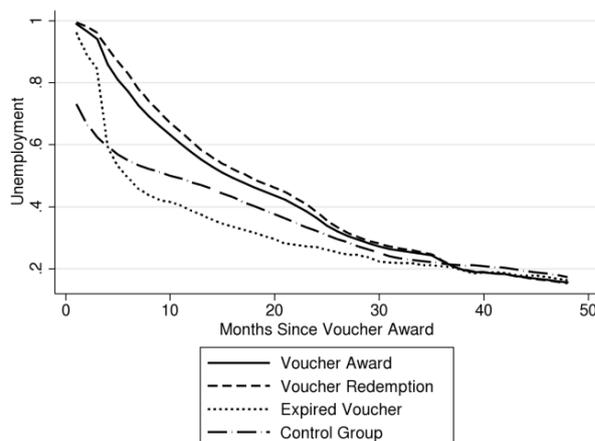


Figure 4.2: Mean unemployment



Note: Group means after using radius matching to adjust the covariate distributions of all groups to that in 'Voucher Award'.

5 Selection processes

5.1 Variables

Our identification strategy requires observing all variables that jointly affect the voucher award and the outcome and/or voucher redemption and the outcome in a relevant way. It is therefore essential to understand which factors affect both voucher award and redemption.

Concerning voucher awards, the analyses of Biewen, Fitzenberger, Osikominu, and Paul (2014) and Lechner and Wunsch (2013), both based on German labour market data, suggest that so-called pre-treatment outcomes (e.g., lagged employment and wages measured prior to the intervention or treatment of interest), benefit receipt histories, socio-economic factors, and local labour market characteristics are important confounders. This information is available in our data. In particular, the individuals' labour market histories are observed up to four years prior to unemployment, and regional factors can be controlled for at the level of the local employment agency district. Furthermore, we observe a range of socio-economic characteristics, such as gender, age, education, profession, marital status, and having children.

While Doerr et al. (2016) argue that a voucher award involves a similar selection process as assignment to ALMPs in general, they also note that the decision is left to the discretion of the caseworker. Our data also contains information collected by the caseworker for use in counselling and assignment decisions, namely, information on the job-seeker's current and previous health status, proxy variables indicating whether an unemployed person lacks motivation (e.g., whether she/he dropped out of a past programme or benefits were withdrawn), and former sanctions.

Concerning actual redemption, Kruppe's (2009) analysis of redemption behaviour suggests that individuals with poor labour market prospects are less likely to redeem their vouchers. We therefore suspect that previous labour market history, socio-economic characteristics such as education and age, and local labour market conditions importantly influence an unemployed individual's decision to participate in vocational training, as they also influence the (personal assessment of the) expected benefits. Furthermore, physical and mental health and personality traits associated with motivation and compliance in the counselling process (approximated by benefit withdrawal and programme drop-out) should affect participation. Conditional on the covariates available, exogenous variation is likely coming from the temporal and regional availability of particular courses (see also, Section 5.3). One factor creating this variation is that the Federal Employment Agency must certify the respective courses (for details, see the discussion in Doerr et al., 2016). This is a lengthy and inflexible administrative process unlikely to be correlated with the individual redemption decision or with the employment outcomes of voucher recipients.

Given that vouchers have to be redeemed within one week to three months, time-varying (or dynamic) confounders of redemption due to important changes in control variables after voucher award but prior to redemption should not be an issue. To verify this argument, we use radius matching to estimate the effects of voucher assignment on a range of

covariates measured on the redemption date, which were all close to and not statistically different from zero.¹⁵ Further, we report the means of the time varying covariates in Table A.3 of Appendix A. We find strong differences in the remaining eligibility for unemployment benefits and in the elapsed unemployment duration. All other time-varying variables show only small differences between the two reference times. We therefore control for the same set of covariates at the same point in time to tackle selection into both voucher award and redemption: gender, age, family background, health and incapacities, nationality, school and vocational education, occupation, complete employment and welfare history over the last four years, past programme and sanction experience, timing and region of unemployment, and economic indicators at the level of the local employment agency (see Table B.1 in Appendix B for the full set of control variables).

5.2 Empirical selection into award and redemption

Table 5.1 provides the probit estimates of two propensity score models for selection into treatment and mediator states for a subset of variables (see Table B.1 in Appendix B for a full set of results). Largely, these results confirm the pattern of univariate comparisons in Table 4.1. Again, it appears that the group receiving vouchers has better overall labour market prospects than the control group, with regional characteristics playing only a limited role.

Individual characteristics have a smaller influence on the redemption decision than on the award decision (Table B.1 Appendix B). Individuals with small children, incapacities, health problems, or lower motivation redeem their vouchers less frequently. However, the timing of unemployment and regional characteristics seem to play important roles in the redemption decision. Vouchers are redeemed less often in regions with a larger share of non-

¹⁵ One might nevertheless be worried about changes in unobserved characteristics. One particular concern is that unemployed individuals receive or anticipate a job offer which we do not observe in our data and that influence the decision not to participate in a programme. This would entail positive bias in the direct effect of voucher award, particularly over the short-run. Note, however, that we find a statistically significant negative direct effect in the first three years; see Section 6.

German unemployed workers and higher population densities. The redemption probability is higher in regions with high employment shares in the production sector, many male unemployed workers, and many vacant full-time jobs.

Table 5.1: Selected average marginal effects from propensity score estimation

	Award Probability		Redemption Probability	
	Marg. Eff. (in %)	Std. Error	Marg. Eff. (in %)	Std. Error
	(1)	(2)	(3)	(4)
Individual characteristics				
Age	-.034***	(.0001)	.068*	(.0004)
Older than 50 years	-9.86***	(.0031)	-4.90***	(.0163)
Children under 3 years	1.14***	(.0014)	.187	(.0048)
Health problems	-3.78***	(.0028)	-4.48***	(.0122)
Incapacities	-2.95***	(.0016)	-6.27***	(.0057)
No German citizenship	-1.65***	(.0022)	-.274	(.0078)
No schooling degree	-2.80***	(.0027)	-.406	(.0106)
University entry degree (Abitur)	.743***	(.0021)	.812	(.0059)
Elementary occupation	.299	(.0026)	.385	(.0100)
Craft, machine operators & related	.743***	(.0022)	.768	(.0080)
Clerks	3.98***	(.0022)	1.173*	(.0072)
Individual labour market history				
Half months empl. in last 2 years	-.037	(.0003)	.179	(.0012)
Half months OLF in last 2 years	-.071**	(.0004)	-.033	(.0013)
Remaining unempl. insurance claim	.150***	(.00005)	.131***	(.0002)
Cum. half months empl. in last 4 y.	.034***	(.0001)	.074***	(.0002)
Cumulative earnings in last 4 years	.00001***	($2.1 \cdot 10^{-8}$)	-.00003***	($6.9 \cdot 10^{-8}$)
Regional characteristics				
Share of empl. in construction	5.17	(.0606)	-4.89	(.1987)
Share of vacant full-time jobs	.512	(.0074)	13.2***	(.0218)
Population per km ²	.0004***	($6.3 \cdot 10^{-7}$)	-.001***	($1.9 \cdot 10^{-6}$)
Unemployment rate (in %)	-.037	(.0003)	.131	(.0009)
Unconditional probability...	6.85%		80.4%	
Sample size (weighted)	93'016 (600'842)		41'138 (41'138)	

Note: Asterisks indicate significant marginal effects at the 1% (***), 5% (**), and 10% (*) level, respectively. Probit model used. Heteroscedasticity robust standard errors are in parentheses. The complete set of variables is contained in Table B.1 in Appendix B. The dummy variable 'health problems' indicates disabilities and health problems which do not fully prevent the unemployed to work, but may reduce the number of possible working hours or the number of possible jobs. The dummy variable 'incapacities' indicates pregnancies, necessarily of medical rehabilitation, or other the incapability of work because of other reasons.

5.3 Regional and temporal availability of vocational training courses

To gain some idea of regional differences in the provision of training courses, we calculate the number of different courses per region (which we observe in the population of training participants). A course is assigned to a specific region if at least one participant from that region joins in a specific month. Other courses may be open to voucher recipients but not observed in our data. Subsequently, we calculate the maximum number of course choices for each awardee as the sum of the number of course choices per region in the three-month period following the award. Although vouchers are valid for a period ranging from 1 week to 3 months, 89% of all vouchers are valid for 3 months, with an average validity of 2.8 months. However, the individual voucher only certifies eligibility for courses with a certain objective,

content and maximum duration. Unfortunately, we cannot match vouchers to specific courses because we have only crude course information. Therefore, this measure approximates an upper limit on the real number of choices. On average, each awardee has a maximum choice set of 112 different courses. Only for 51 of 41,138 vouchers awarded do we observe zero possible course choices. The maximum number of possible course choices is 712. Figure F.1 of Appendix F provides a histogram of the observed course choices. This figure documents a large variation in the maximum number of possible course choices per awardee. In Figure F.2 of Appendix F, we report the time variation in the average number of maximum course choices. We find strong seasonal effects. Most courses take place either in August/September or in February/March. These dates are related to the typical school cycle in Germany. The school year begins in August or September (depending on the state). Fewer courses start before Christmas or summer holidays.

In Table F.1 of Appendix F, we regress the maximum number of course choices on different regional characteristics and dummies for the month of the voucher award. We find strong variation in the maximum number of course choices across states. We find the maximum course choice is larger in regions with higher population densities, higher regional unemployment rates, many vacant full-time jobs, and relatively large shares of female and non-German unemployed workers. These findings are robust to controlling for the month that the voucher is awarded. In Table F.2 of Appendix F, we estimate the redemption probability with respect to the maximum number of course choices. We find that the maximum number of course choices has a positive association with the redemption decision. Vouchers are redeemed less often in regions with high population densities, but more often in regions with high unemployment rates. This supports our hypothesis that the regional and temporal availability of training courses is correlated with the redemption decision.

5.4 Alternative treatments

In Table A.4 of Appendix A, we provide additional descriptive evidence of the amount of other training obtained by the control group, the awardees, the redeemers, and the non-redeemers. Second vouchers may be provided for consecutive vocational training programmes or because the first voucher expired (e.g., because no appropriate course was found). Individuals are not entitled to a second voucher. The award of a second voucher is entirely up to the discretion of the caseworker. Our data show that redeemers have an 11% probability and non-redeemers a 22% probability of receiving a second voucher within one year. The control group includes individuals not receiving a voucher during the first twelve months of unemployment. Only 0.6% of individuals receive a voucher later. Individuals in the control group and non-redeemers have an approximately one percentage point higher probability of participating in an alternative training programme (other than vocational training) than individuals who redeem their first voucher. Usually, these training programmes provide direction for the period of unemployment, information about future employment possibilities, or application training. Individuals in the control group have a higher probability of participating in job creation schemes than training participants and non-redeemers. These programmes are designed for unemployed workers with problems re-integrating into the labour market. They are a different target group than that for vocational training, which is supposed to have high re-employment chances. Individuals in the control group and non-redeemers have a higher probability of receiving government grants than training participants. Government grant are work subsidies. Typically, these are start-up grants to become self-employed or subsidies paid to the employer for hiring an unemployed workers.

6 Results

6.1 Main findings

The propensity score estimates presented in Table 5.1 serve as inputs into the matching algorithm. When performing matching, one should check for potential issues of (i) insufficient support in the propensity scores across treatment states that may result in incomparable matches as well as large matching weights of some non-treated observations with specific propensity scores and (ii) imbalances in covariates after matching (due to inappropriate propensity score specifications). In our application, insufficient support is not a problem, as seen from the distributions of the propensity scores of the different groups (details in Figures B.1 and B.2 in Appendix B). Furthermore, the important covariates are well balanced (for details, see Table B.2 in Appendix B).

Figures 6.1 and 6.2 provide the estimates of the average employment and unemployment effects on voucher recipients, namely, the (total) impacts of voucher award vs. non-award (Δ), and the effects of voucher award and redemption vs. non-award (θ), voucher award without redemption vs. non-award (γ), and voucher award with redemption vs. voucher award without redemption (δ). Concerning employment, we consider only (non-marginal, non-subsidised) employment lasting at least one month. The lines reflect the effect magnitudes on the probability of being employed or unemployed in a particular month after receiving a voucher over 4 years (48 months). The superimposed symbols on the lines (diamonds) indicate effects that are (pointwise) statistically significantly different from zero at the 5%-level.

The results in Figure 6.1 suggest that awarding a voucher has a negative (total) employment impact among voucher recipients in the first three years, particularly in the initial months, where the employment probability decreases by as much as 10 percentage points. This dip suggests a lock-in effect likely due to reduced job search intensity in response to (an-

anticipated) participation in a vocational training. However, the negative effect fade over roughly three years, and the employment probability increases by approximately 2 to 3 percentage points in the fourth year. The positive employment effect appears quite stable, suggesting that that the voucher award system successfully offsets the initial lock-in effect with higher placement success in later periods. These results are qualitatively and quantitatively similar to the findings of Doerr et al. (2016), even though they rely on a different empirical approach.

Figure 6.2 shows that the time patterns in unemployment are reversed (as expected), but the effects are initially larger. This is because over the short-run the award of a voucher reduces drop-out from the labour market, as shown by the effect on the ‘out-of-labour force’ labour market state (see Figure C.2 in Appendix C for details). Similar in magnitude (but with the opposite sign) to the employment effect, registered unemployment is somewhat reduced over the long-run.

Figure 6.1: Employment

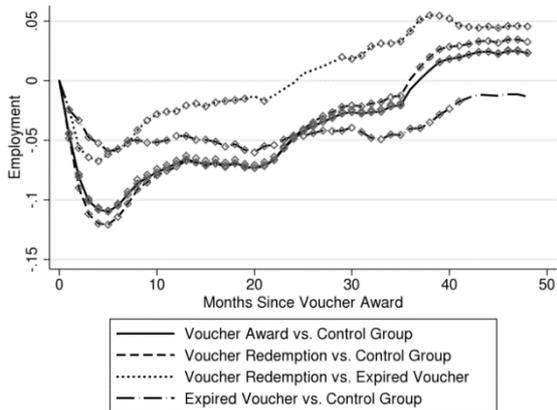
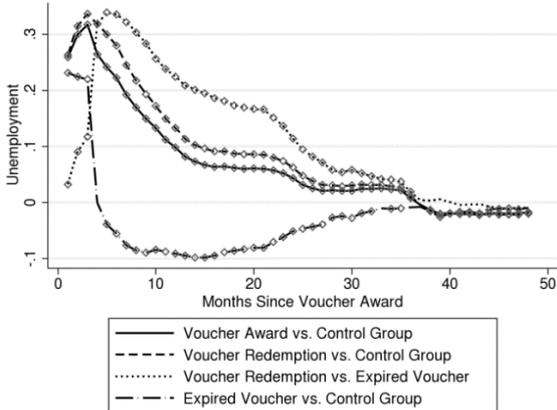


Figure 6.2: Registered unemployment



Note: Separate effects for the first 48 months following the voucher award are estimated. Diamonds indicate significant effects at the 5%-level.

When investigating the causal mechanisms underlying the total employment effect (with essentially symmetric results for registered unemployment), it becomes apparent that it is predominantly redemption (e.g., participation in/starting vocational training) driving the results.

In fact, the estimated effect of voucher award and redemption vs. no award (θ) closely follows the overall impact of voucher award, albeit it is somewhat more negative in earlier periods and more positive in later periods. In contrast, the direct effect of a voucher award without redemption (γ) is insignificant and close to zero over most of the fourth year. This suggests that over the long-run, voucher assignment alone does not affect, for instance, preferences for human capital investments in a way that influences employment success.

We find a negative direct effect over the short-run: Even without redemption, a voucher award decreases the employment probability in the first three to 3.5 years. Therefore, it appears that non-redeemers reduce job search activities. This may be rooted in the learning and decision process about the supply of vocational training. The possibility of awareness effects with respect to inconspicuous survey questions has been documented in Crossley, de Bresser, Delaney, and Winter (2014). Comparable effects may occur in the labour market. Individuals may initially reduce their job search intensity in response to a voucher award and consider the programmes available instead. Some of them may not be satisfied with the available options and decide not to redeem the voucher. Instead, they try to find employment again. This possible channel would be in line with the results of Crépon, Ferracci, Jolivet, and Van den Berg (2014), who report negative effects of notifications of planned training upon employment exits. They argue notification of planned training causes an ‘attraction effect’ that reduces search efforts, rather than a ‘threat effect’ that intensifies job search (e.g., Van den Berg, Bergemann, and Caliendo, 2009).

We would expect this initial, *direct* lock-in effect to be less severe than for the total effect (which includes the impact of actual redemption leading to training participation), as individuals should be available for intensive job search sooner by foregoing redemption. Indeed, we find that in the initial periods, the estimate of γ is considerably less negative than the estimates of Δ and in particular of θ (redemption vs. non-award). Accordingly, the estimate of

δ (redemption vs. award without redemption) is initially negative (as $\delta = \theta - \gamma$) and significantly so. In later periods, however, redemption pays off for the group of voucher recipients: After roughly two years, the estimates of Δ and θ dominate those of γ , and the estimates of δ are statistically significant and non-negligible (up to 5 percentage points) in later periods.

We considered several further outcome variables (detailed results are presented in Appendix C). They contain a measure of employment stability, i.e., being employed for at least 6 months, for which the outcome evaluation window starts only in month 7 after the voucher award. The estimates of the (total) impact of voucher award vs. non-award (Δ) and voucher award and redemption vs. non-award (θ) on stable employment are qualitatively similar to those on employment, albeit significantly positive at a later point in time and of a somewhat smaller magnitude. In contrast to Figure 6.1, the estimate of the direct effect (γ) remains statistically significantly negative until the end of the evaluation window (implying that the adverse effect of not redeeming a voucher vs. not receiving one is more severe for stable employment), even though it shows an upward tendency.

Furthermore, we investigated the effects on full-time employment. Again, the results are qualitatively similar to the employment effects shown in Figure 6.1, including an insignificant direct effect in the fourth year after a voucher award. Similar conclusions are drawn concerning the effects on monthly earnings: After an initial lock-in phase, the estimates of Δ , θ , and δ are moderately positive (between 30 and 70 euros) and statistically significantly in the fourth year, while those of γ approach zero.

6.2 Cumulative effects

In Table C.1 of Appendix C, we report the cumulative effects of the different treatments 2 and 4 years after the award of a voucher. The cumulative employment and earnings effects

of awarding, redeeming, and not redeeming a voucher never become positive compared to the control group during our observation period. We do not observe significant employment and earnings effects of redeeming vs. not redeeming a voucher after four years. Redeeming a voucher increases (decreases) the average probability over four years of being unemployed (out of labour force).

After two years, we observe more negative indirect than direct voucher effects. Over the long-run, this relation reverts but does not pay in terms of cumulated employment or earnings during our observation period. This suggests the benefits of redeeming and not redeeming a voucher are, on average, similar over the four-year period. However, individuals profit more from not redeeming a voucher over the short-run, whereas training is more beneficial over the long-run. Suggestively, the benefits of actual participation are greater than not redeeming a voucher beyond the four-year time horizon.

6.3 Effect heterogeneity

In this section, we investigate effect heterogeneity for redeemers and non-redeemers. In general, the results remain qualitatively similar across groups. However, the negative direct voucher effects on employment are less severe for individuals who do not redeem the vouchers compared to group who redeemed them (see Figures C.6-C.12 in Appendix C). These differences in the direct voucher effects on employment are mainly driven by part-time and stable employment. The direct voucher effects on earnings are slightly more negative for non-redeemers than for redeemers. The indirect voucher effects on employment and earnings are, over the long-run, more beneficial to non-redeemers than to redeemers (see Figures C.13-C.19 in Appendix C). These results are driven by full-time employment. This suggests the self-selected redemption of vouchers reduces the long-term effectiveness of training participation. However, the negative lock-in effect of training participation would be steeper for non-redeemers than for redeemers.

If the long-run positive employment benefits exceed the initial period of deterioration, these results could point to a behavioural bias. Awardees who could ultimately benefit most from training participation let their voucher expire. They might have incorrect information or inaccurate expectations about the future returns to vocational training. Alternatively, they merely have higher preferences for short-term employment and earnings opportunities. However, these possibilities cannot be verified with our data.

7 Conclusion

Using rich administrative labour market data from Germany, we evaluated the effectiveness of awarding vouchers for vocational training programmes to unemployed individuals. We found an overall negative short-run but a positive longer-run effect on the employment chances of voucher recipients. We also investigated the causal mechanism through which the overall effect materializes using sequential conditional independence assumptions for identification. In particular, we considered the direct employment effect of voucher assignment (net of actual redemption), which may be driven, for instance, by decreased job search intensity during the assessment of training options or increased salience/awareness of (and changed preferences for) human capital investments in general.

The direct effect was small and statistically insignificant over the longer run but negative over the short-run (albeit less so than the overall impact), pointing to decreased search intensity shortly after voucher assignment (despite non-redemption). In contrast, the effect of actual voucher redemption (vs. non-award and non-redemption) closely follows the overall effect, although it is somewhat more negative in earlier periods and more positive in later periods. Comparing the latter to the direct effect suggests that, conditional on voucher assignment, redemption (and thus, actual programme participation) entails a more severe negative (lock-in) effect on voucher recipient than non-redemption, which is intuitive because individuals not redeeming vouchers are available for the labour market sooner. Over the

longer run, however, redemption pays off by increasing the employment probability by approximately two to three percentage points compared to a non-award in the fourth (and last observed) year after voucher assignment.

From a policy perspective, these results suggest that the introduction of a voucher award system, which was embraced to promote responsibility for training among participants and competition among training providers, may lead to a loss in effectiveness if individuals do not make use of the awards, because non-redemption entails lower employment chances than both redemption over the long-run and non-assignment over the short-run. Therefore, non-redemption appears to be the least attractive option. These findings are relevant to delivering active labour market policies using a voucher system.

References

- Abadie, A., and Imbens, G.W. (2008): On the Failure of the Bootstrap for Matching Estimators. *Econometrica*, 76, 1537–1557.
- Abadie, A., and Imbens, G.W. (2011): Bias-Corrected Matching Estimators for Average Treatment Effects, 29, *Journal of Business & Economic Statistics*, 1–11.
- Abbring, J.H. and van den Berg, G.J. (2003): The Non-Parametric Identification of Treatment Effects in Duration Models", *Econometrica*, 71, 1491-1517.
- Abbring, J.H. and van den Berg, G.J. (2004): Analyzing the Effect of Dynamically Assigned Treatments using Duration Models, Binary Treatment Models, and Panel Data Models, *Empirical Economics*, 29, 5-20.
- Baron, R.M., and Kenny, D.A. (1986): The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Biewen, M., Fitzenberger, B., Osikominu, A., and Paul, M. (2014): The Effectiveness of Public Sponsored Training Revisited: The Importance of Data and Methodological Choices. *Journal of Labor Economics*, 32, 837–897.
- Bodory, H., Camponovo, L., Huber, M., and Lechner, M. (2016): The finite sample performance of inference methods for propensity score matching and weighting estimators. University of St. Gallen, Dept. of Economics Discussion Paper no. 2016-04.
- Card D., Kluve, J., and Weber, A. (2010): Active labour market policy evaluations: A meta-analysis. *The Economic Journal*, 120, F452-F477.

- Card D., Kluve, J., and Weber, A. (2015): What works? A meta analysis of recent active labor market program evaluations. IZA Discussion Paper, 9236.
- Crépon, B., M. Ferraci, G. Jolivet, and G.J. Van den Berg (2014): Information shocks and the empirical evaluation of training programs during unemployment spells. Working paper.
- Crossley, T. F., de Bresser, J., Delaney, L., and Winter, J. (2014): Can survey participation alter household saving behavior? Forthcoming in the *Economic Journal*.
- Dehejia, R.H., and Wahba, S. (2002): Propensity-score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, 84, 151–161.
- Doerr, A., Fitzenberger, B., Kruppe, T., Paul, M., and Strittmatter, A. (2016): Employment and Earnings Effects of Awarding Training Vouchers in Germany. Forthcoming in *Industrial and Labor Relations Review*.
- Doerr, A. and Kruppe, T. (2015): Training Vouchers, Local Employment Agencies, and Policy Styles. *Journal of Labour Market Research*, 48, 41-56.
- Doerr, A., and Strittmatter, A. (2014): Assignment Mechanisms, Selection Criteria, and the Effectiveness of Training Programmes. SEPS Discussion Paper, 1421.
- Flores, C.A., and Flores-Lagunes, A. (2009): Identification and Estimation of Causal Mechanisms and Net Effects of a Treatment under Unconfoundedness. IZA Discussion Paper No. 4237.
- Fredriksson, P. and Johannsen, P. (2008): Dynamic Treatment Assignment - The Consequences for Evaluations using Observational Data, *Journal of Business and Economic and Statistics*, 26, 435-445
- Ham, J. C., and Lalonde, R. J. (1996): The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training. *Econometrica*, 64, 175–205.
- Heckman, J., Ichimura, H., and Todd, P. (1998): Matching as an Econometric Evaluation Estimator. *The Review of Economic Studies*, 65, 261-294.
- Heckman, J. and Navarro, S. (2007): Dynamic Discrete Choice and Dynamic Treatment Effects, *Journal of Econometrics*, 136, 341-396.
- Heckman, J., Pinto, R., and Savelyev, P. (2013): Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, 103, 2052–2086.
- Heinrich, C., Mueser, P., Troske, K., Jeon, K., and Kahvecioglu, D. (2010): New Estimates of Public Employment and Training Program Net Impacts: A Nonexperimental Evaluation of the Workforce Investment Act Program,” Working Paper No. 1003, Department of Economics, University of Missouri.
- Huber, M. (2014): Identifying causal mechanisms (primarily) based on inverse probability weighting. *Journal of Applied Econometrics*, 29, 920-943.
- Huber, M., M. Lechner, M., and Mellace, G. (2014): Why do tougher caseworkers increase employment? The role of programme assignment as a causal mechanism. Forthcoming in *Review of Economics and Statistics*.
- Huber, M., Lechner, M., and Wunsch, C. (2013): The performance of estimators based on the propensity score. *Journal of Econometrics*, 175, 1–21.
- Imai, K., Keele, L., and Yamamoto, T. (2010): Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, 25, 51–71.

- Imbens, G. W. (2000): The Role of the Propensity Score in Estimating Dose-response Functions. *Biometrika*, 87, 706 – 710.
- Joffe, M.M., Ten Have, T.R., Feldman, H., and Kimmell, S.E. (2004): Model Selection, Confounder Control, and Marginal Structural Models: Review and New Applications. *The American Statistician*, 58, 272–279.
- Kruppe, T. (2009): Bildungsgutscheine in der aktiven Arbeitsmarktpolitik. *Sozialer Fortschritt*, 58 (1), 9-19.
- Lechner, M. (1999): Earnings and Employment Effects of Continuous Off-the-job Training in East Germany after Unification", *Journal of Business Economics and Statistics*, 17, 74-90.
- Lechner, M. (2001): Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption. In: Lechner, M., and Pfeiffer, F., (eds.) *Econometric Evaluation of Labour Market Policies*, 43–58, Heidelberg, Physica/Springer.
- Lechner, M. (2009): Sequential Causal Models for the Evaluation of Labor Market Program. *Journal of Business & Economic Statistics*, 27, 71–83.
- Lechner, M., and Miquel, R. (2010): Identification of the effects of dynamic treatments by sequential conditional independence assumptions. *Empirical Economics*, 39, 111–137.
- Lechner, M., Miquel, R., and Wunsch, C. (2011): Long-Run Effects of Public Sector Sponsored Training In West Germany. *Journal of the European Economic Association*, 9, 742–784.
- Lechner, M., and J. Smith (2007): What is the Value Added by Caseworkers?, *Labour Economics*, 14, 135-151.
- Lechner, M., and Wunsch, C. (2013): Sensitivity of Matching-Based Program Evaluations to the Availability of Control Variables. *Labour Economics*, 21, 111–121.
- Miquel, R. and Lechner, M. (2009): Identification of Effects of Dynamic Treatments by Sequential Conditional Independence Assumptions, *Empirical Economics*, 39, 111-137.
- Pearl, J. (2001): Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, San Francisco: Morgan Kaufman, 411–420.
- Petersen, M.L., Sinisi, S.E., and van der Laan, M.J. (2006): Estimation of Direct Causal Effects. *Epidemiology*, 17, 276–284.
- Rinne, U., Uhlendorff, A., and Zhao, Z. (2013): Vouchers and Caseworkers in Training Programs for the Unemployed, *Empirical Economics*, 45, 1089–1127.
- Robins, J.M. (1986): A new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survivor effect. *Mathematical Modelling*, 7, 1393–1512, with 1987 Errata to “A new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survivor effect”. *Computers & Mathematics with Applications*, 14, 917–921; 1987 Addendum to “A new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survivor effect”, *Computers & Mathematics with Applications*, 14, 923–945; and 1987 Errata to “Addendum to ‘A new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survivor effect’”, *Computers & Mathematics with Applications*, 18, 477.

- Robins, J.M. (1989): The analysis of randomized and nonrandomized aids treatment trials using a new approach to causal inference in longitudinal studies. In: Sechrest, L., Freeman, H., and Mulley, A. (eds.) Health service research methodology: a focus on AIDS, 113–159, Washington, D.C., Public Health Service, National Center for Health Services Research.
- Robins, J.M., Hernan, M.A., and Brumback, B. (2000): Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11, 550–560.
- Robins, J.M. (2003): Semantics of causal DAG models and the identification of direct and indirect effects. In *Highly Structured Stochastic Systems*, P.J. Green, N.L. Hjort, and S. Richardson (eds.), Oxford: Oxford University Press, 70–81.
- Robins, J.M., and Greenland, S. (1992): Identifiability and Exchangeability for Direct and Indirect Effects. *Epidemiology*, 3, 143–155.
- Rosenbaum, P., and Rubin, D.B. (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70, 41–55.
- Rosenbaum, P., and Rubin, D. B. (1985): Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. *The American Statistician*, 39, 33 – 38.
- Rubin, D.B. (1974): Estimating Causal Effects of Treatments in Randomized and Nonrandomized studies. *Journal of Educational Psychology*, 66, 688 – 701.
- Rubin, D.B. (1979): Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies. *Journal of the American Statistical Association*, 74, 318–328.
- Rubin, D.B. (1980): Discussion of ‘Randomization Analysis of Experimental Data in the Fisher Randomization Test’ by D. Basu. *Journal American Statistical Association*, 75, 591–593.
- Sianesi, B. (2004): An Evaluation of the Swedish System of Active Labour Market Programmes in the 1990s", *Review of Economics and Statistics*, 86, 133-155.
- Van den Berg, G.J., Bergemann, A., and Caliendo, M. (2009): The effect of active labor market programs on not-yet treated unemployed individuals. *Journal of the European Economic Association*, 7, 696-616.
- Van derWeele, T.J. (2009): Marginal Structural Models for the Estimation of Direct and Indirect Effects. *Epidemiology*, 20, 18–26.