Direct and indirect effects of training vouchers for the unemployed

Martin Huber

University of Fribourg, Switzerland

and Michael Lechner and Anthony Strittmatter

University of St Gallen, Switzerland

[Received September 2015. Final revision February 2017]

Summary. The 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 the chances of employment 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: Causal channels; Causal mechanisms; Direct effects; Indirect effects; Matching estimation; Mediation analysis; Training programmes; Voucher award

1. Introduction

In January 2003, the German Federal Employment Agency reformed the allocation of vocational training programmes, which are a cornerstone 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 between training providers. Before the reform, caseworkers determined placement into vocational training. Since 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, for example, the meta-analyses by Card et al. (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 effect through its redemption (i.e. participation in vocational training) there is
a direct effect of the voucher award (i.e. without participation). The latter may be driven by
motivational effects, preference shaping and salience or awareness regarding (the availability
of) ALMPs, which could affect labour market behaviour.

Such effects would be in line with Van den Berg et al. (2009), who documented that the sub-
jective individual expectation to be assigned to an ALMP affects job search behaviour. Further,
Crépon et al. (2014) reported that the mere notification of a planned training assignment has
effects on the unemployment exit probability. In a related endeavour, Crossley et al. (2014) doc-
umented that mere exposure to specific questions in a survey can alter subsequent behaviour.
Exploiting randomized assignment to survey modules within the ‘Longitudinal Internet stud-
ies for the social sciences’ panel survey, they found that households answering questions on
expenditures and needs in retirement significantly changed their non-housing saving rates.
They argued that the survey may have acted as a salience shock and comparable effects might
occur from being offered a training voucher, as this might induce individuals to become aware
of and to reflect on the ALMPs and their expected effects.

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 inten-
sity. The same effect occurs if awarding a voucher leads to high information costs of looking
for appropriate providers and courses. In contrast, if participation in ALMPs is perceived as
a burden or as ineffective in raising employability or if a voucher award increases the salience
or 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 effect 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 drives 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, for this, we consider the redem-
tion 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 redemp-
tion, 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. (We
refer to Pearl (2001) for a discussion of the differences between controlled and natural direct ef-
teffects.) However, causal mechanisms are not easily identified. Even if the vouchers were randomly
assigned, this would not imply randomness of the mediator (see Robins and Greenland (1992)).
(This issue has also been discussed in the context of randomized training programmes; see, for
instance, Ham and Lalonde (1996), who noted 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).)

To tackle the endogeneity of voucher award and redemption, a particular conditional
independence assumption is invoked for identification. It requires

(a) that the voucher award is independent of potential employment outcomes (under (non-)
award and (non-)redemption of the vouchers) conditional on observed covariates and
(b) 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 non-parametric mediation literature for
identifying controlled direct effects (see, for instance, Petersen et al. (2006) and VanderWeele (2009)). They also appear in the dynamic treatment effects literature on assessing sequences of treatments (see, for instance, Robins (1986, 1987a, b, c, 1989), Robins et al. (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 et al., 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 3 years after voucher receipt. It has a small positive effect thereafter, with an increased employment probability of approximately 2–3 percentage points throughout the fourth year after receiving the voucher. Thus, the initial negative lock-in effect of a voucher award (probably due to decreased job search intensity) is offset by higher probabilities of employment 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 to those of 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 3 years, suggesting that voucher award decreases job search intensity despite non-redemption. This points to potential losses in effectiveness of voucher award systems if individuals do not redeem their vouchers, as the chances of employment are lower than under non-award over the short run and under redemption over the long run. Therefore, voucher award and non-redemption appear 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 assess merely the total (gross) effect of the programme. Therefore, this research goes beyond Doerr et al. (2016) and Heinrich et al. (2010), who evaluated the effectiveness of various vocational training programmes in voucher systems but did not consider the direct effect of voucher award. It also differs from Doerr and Strittmatter (2016) and Rinne et al. (2013), who compared the effectiveness of vocational training via voucher and mandatory assignment regimes but did 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 that was 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. The on-line appendices A–F provide further details on data, estimation and results.

The programs that were used to analyse the data in the paper can be obtained from

http://wileyonline.library.com/journal/rss-datasets

2. Institutional background of voucher provision

Vocational training programmes constitute a cornerstone 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 course: classic vocational training, training in so-called practice firms and retraining. Examples are courses in information-technology-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 3 years with the goal of completing a vocational degree within the German apprenticeship system. They cover, for example, a full curriculum of vocational training for a care nurse for the elderly. 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) usually takes place on the job but can also involve classroom sessions. Between 2000 and 2002, average annual expenditures for vocational training exceeded €7 billion (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 1-day commuting zone. (For a training course lasting 6 or more h per day, commuting times of up to 2.5 h are reasonable. For a training course lasting less than 6 h per day, the reasonable commuting time is reduced to 2 h.) Third, the training vouchers are valid for a period ranging between 1 week and 3 months. Fourth, no sanctions or penalties are 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 the chances of employment. 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 outcomes 6 months after training, particularly for training programmes with long durations, remains difficult. There were probably regional differences in the way that predictions were formed. Doerr and Kruppe (2015) conducted a survey of caseworkers to analyse regional differences in voucher award intensities. They found that such differences can be (partly) explained by the preferences and sentiments of 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 have awarded a voucher for each course or only a single voucher, which certifies eligibility for the whole sequences.

Crépon et al. (2014) found a negative effect of a notification of possible training on exits from unemployment. Because notified unemployed people must search for an appropriate course on their own, this notification has many similarities to a voucher award. However, Crépon et al. (2014) analysed a French programme that differs in several dimensions from the German vocational training vouchers that we analyse. In particular, the notified unemployed workers must apply for funding for their training after receiving a notification. This differs from the German voucher system, which guarantees payment for 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 et al. (2014) investigated the hazard rate for leaving unemployment, whereas 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 $d \in \{1, 0\}$. (By defining the potential outcomes this way, we implicitly impose the stable unit treatment value assumption; see Rubin (1980).) 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

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

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 1 or 0, depending on whether an individual redeems a received voucher. Thus, the ATET provides the total 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. (It is worth noting that the mediation framework has some resemblance to the literature on instrumental variables (IVs), 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. IVs 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 effect).)

The extended notation allows further definitions of parameters, e.g. the average effect of voucher award and redemption versus no award and no redemption among voucher recipients:
\[ \theta = E(Y^{1,1} - Y^{0,0} | D = 1). \] (2)

The difference from the ATET is that, in model (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, whereas \( Y^{1,1} \) measures the potential outcome under voucher award and redemption (i.e. voucher and redemption statuses are ‘forced’ to be equal for this potential outcome). 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 \( \theta = \Delta \). Again, part or all of the effect might be due to redemption or to a direct award effect. In the next step, we disentangle these two components and consider the so-called controlled direct effect (see, for instance Pearl (2001)):

\[ \gamma = E(Y^{1,0} - Y^{0,0} | D = 1). \] (3)

(A related parameter is the so-called natural direct effect in the nomenclature of Pearl (2001) or the pure or total direct effect in the nomenclature of Robins and Greenland (1992) and Robins (2003), which is defined on potential mediator states rather than prescribed mediator values: \( E(Y^{1,M^1} - Y^{0,M^1} | D = 1) \) and \( E(Y^{1,M^0} - Y^{0,M^0} | D = 1) \). These two parameters and \( \gamma \) 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 et al. (2010) and Huber (2014), among many others.)

This is the effect 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

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

Here, the effect of redemption versus non-redemption is investigated conditionally on awarding a voucher. Note that \( \gamma \) and \( \delta \) sum to \( \theta \), 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

(a) with \( D \) and the potential outcomes and
(b) 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.** \( \{Y^{1,1}, Y^{1,0}, Y^{0,0}\} \perp\!\!\!\!\perp D | X = x \) 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 \( \theta, \gamma \) and \( \delta \) requires a further conditional independence assumption.
Assumption 2. \( \{Y^{1,1}, Y^{1,0}, Y^{0,0}\} \perp M | X = x, D = d \) for all \( d \in \{1, 0\} \) 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 et al. (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 Lechner (2009) and Lechner and Miquel (2010) 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 et al. (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

(a) the informative set of observed characteristics that are available (see the discussion in Section 5.1),
(b) the small time lag between \( D \) and \( M \) and
(c) the randomly assigned pseudostart dates for \( D \) and \( M \) among those with \( D = 0 \) and \( M = 0 \) or \( D = 1 \) and \( M = 0 \).

The aim of this procedure is to control for any differences in elapsed unemployment durations (that are likely to affect \( Y \)) across redeemers, non-redeemers and non-awardees respectively. See Section 4.2 for further details.

Assumption 3. \( \Pr(D = 1|X = x) < 1 < \Pr(M = 1|D = 1, X = x) < 1 \) 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 \( \Delta, \theta \) and \( \gamma \) (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, \( \theta \) (which involves \( Y^{1,1} \)), \( \gamma \) (which involves \( Y^{1,0} \)) and \( \delta \) (which involves both) are not identified.

Assumptions 1 and 2 together imply the following conditional independence restriction: \( \{Y^{1,1}, Y^{1,0}, Y^{0,0}\} \perp \{D, M\} | X = x \) for all \( x \) in the support of \( X \). Technically, the various combinations of \( D \) and \( M \) (despite their sequentiality) may be treated as distinct treatments when identifying \( \theta, \gamma \) and \( \delta \) 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

\[
E(Y^{1,M}|D = 1) = E(Y^{1}|D = 1) = E(Y|D = 1),
\]

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

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

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. The derivation of these results is standard (e.g. Heckman et al. (1998), section 3) and has been omitted.)

However, directly controlling for a possibly high dimensional vector $X$ when estimating $E_{X|D=1} \{ E(Y|D=0, X=x) \}$ and $E_{X|D=1} \{ E(Y|D=d, M=m, X=x) \}$ may lead to the curse of dimensionality. Rosenbaum and Rubin (1983) showed that we may instead condition on the treatment propensity scores, in our case $p(x) = \Pr(D = 1 | X = x)$ and $p_{dn}(x) = \Pr(D = d, M = m | X = x)$, which balance the distributions of $X$. Therefore, it holds that

$$E(Y_{0,M^0} | D = 1) = \frac{E}{p(x)} E \{ E(Y | D = 0, p(X) = p(x)) \},$$

$$E(Y_{d,m} | D = 1) = \frac{E}{p_{dn}(x)} E \{ E(Y | D = d, M = m, p_{dn}(X) = p_{dn}(x)) \}.$$

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 $p_{dn}(X)$. Specifically, $\Delta$ is estimated by

(a) matching to all voucher awardees comparison observations without voucher award that are similar in terms of estimates of $p(X)$ and

(b) taking the mean difference in outcomes between the two groups.

The estimation of $\theta$ 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 $\gamma$ and $\delta$.

### 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 by using the estimator of Lechner et al. (2011). Whereas the propensity scores $p_{x}(X)$ and $p_{dn}(X)$ are parametrically specified by probit models, the conditional expectations of the outcomes are unrestricted and thus non-parametric. The algorithm is more precise than nearest-neighbour matching because of 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) and Joffe et al. (2004)). Moreover, the regression adjustment should reduce small sample and asymptotic biases of matching. Huber et al. (2013) investigated the finite sample properties of this algorithm along with other matching-type estimators and found 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, i.e. 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. (If there is no comparison observation within the radius, then the nearest neighbour is matched.) Alternative radius sizes do not affect the results importantly (see Table D.1 of the on-line 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) showed 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 that is used. It performs well in a large-scale (empirically based) simulation study by Bodory et al. (2016), who investigated 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.

4.1. Data

Our analysis is based on administrative data provided by the Federal Employment Agency of Germany, namely the ‘Integrated employment biographies’ (IEB) database. (The IEB database is a rich administrative database and has been used in virtually all recent studies on German ALMPs (e.g. Biewen et al. (2014), Lechner et al. (2011), Lechner and Wunsch (2013) and Rinne et al. (2013)). The IEB is a merged data file containing individual records collected by four different administrative processes: the Institut für Arbeitsmarkt- und Berufsforschung (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). The IAB is the research department of the Federal Employment Agency.) The IEBs contain information on all individuals in Germany who received a voucher between 2003 and 2004, along with subsequent participation in vocational training programmes, i.e. 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 ALMPs (type and duration), as well as rich individual information (e.g. education, age, gender, marital status, profession and nationality) and regional (labour market) characteristics. Thus, we can control for a wealth of personal characteristics and detailed labour market histories (e.g. type of employment, industry, occupational status and 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 database and is a 3% random sample of individuals who experienced at least one transition from employment to non-employment (lasting at least 1 month) in 2003. (We account for the different sampling probabilities whenever necessary by using sampling weights. These probabilities differ only for unemployed people 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 3 months. Entering unemployment is defined as transitioning from (non-subsidized, non-marginal, non-seasonal) employment to registered non-employment for at least 1 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 who are eligible for early retirement schemes, we consider only people 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 before award and redemption of the voucher. The award and redemption decisions are dynamic processes. Caseworkers can award a voucher to an unemployed individual 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.

If the sample had been sufficiently large, an attractive approach would have been to use the dynamic evaluation framework (as suggested by Robins (1986), Lechner (2009) and 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) took the timing of programme starts explicitly into account. However, given our small sample size (we observe 8061 awardees with unredeemed vouchers), neither approach is feasible with our data at the desired level of flexibility. Furthermore, the sample sizes are too small to follow the approach that was suggested by Fredriksson and Johansson (2008) and applied by Sianesi (2004) to estimate differential effects by the elapsed duration of unemployment. 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 programme effects and differences due to shifted timings of voucher award and redemption.

As a compromise that trades off the issues that were 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 3 months. Non-redeemers do not start vocational training during the first 3 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 that were 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 had possibly already found a job before 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 pseudoaward dates to each individual in the control group. Thereby, we recover the distribution of the elapsed duration of unemployment at the time of voucher award from the treatment group (similarly to, for example, Lechner (1999) and Lechner and Smith (2007)). To ensure comparability of the treatment definitions of the awardees and non-awardees, we consider only individuals who are unemployed at their (pseudo)voucher award. Following similar arguments, the same approach is applied with respect to the mediator to create (pseudo)voucher redemption dates among those who did not redeem a voucher. (592 individuals with expired vouchers were dropped because of the definition of the pseudovoucher redemption dates.) This makes the groups of individuals with redeemed and expired vouchers comparable with respect to the duration of unemployment.

In Fig. 2 in Section 6 as well as Figs C.1–C.5 in the Web appendix C, we show the results controlling for the elapsed duration of unemployment until (pseudo)voucher award and
Table 1. Means and standardized biases of selected variables†

<table>
<thead>
<tr>
<th>Subsample means</th>
<th>Voucher awarded</th>
<th>Voucher redeemed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes, (1)</td>
<td>No, (2)</td>
</tr>
<tr>
<td>Voucher awarded</td>
<td>39.03</td>
<td>41.75</td>
</tr>
<tr>
<td>Voucher redeemed</td>
<td>0.43</td>
<td>0.35</td>
</tr>
<tr>
<td>Age</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>Children under 3 years</td>
<td>0.12</td>
<td>0.19</td>
</tr>
<tr>
<td>Health problems</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>Incapacities</td>
<td>0.04</td>
<td>0.07</td>
</tr>
<tr>
<td>No German citizenship</td>
<td>0.23</td>
<td>0.17</td>
</tr>
<tr>
<td>No schooling degree</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>University entry degree (Abitur)</td>
<td>0.29</td>
<td>0.35</td>
</tr>
<tr>
<td>Elementary occupation</td>
<td>0.25</td>
<td>0.16</td>
</tr>
<tr>
<td>Clerks</td>
<td>45.17</td>
<td>44.30</td>
</tr>
<tr>
<td>Half-months employed in last 2 years</td>
<td>1.59</td>
<td>2.19</td>
</tr>
<tr>
<td>Half-months out of labour force in last 2 years</td>
<td>91258</td>
<td>84199</td>
</tr>
<tr>
<td>Cumulative earnings in last 4 years</td>
<td>8.90</td>
<td>10.95</td>
</tr>
<tr>
<td>Months of remaining unemployment benefits</td>
<td>4.46</td>
<td>3.76</td>
</tr>
<tr>
<td>Elapsed unemployment duration</td>
<td>12.33</td>
<td>12.53</td>
</tr>
<tr>
<td>Observations</td>
<td>41138</td>
<td>51878</td>
</tr>
<tr>
<td>Sum of weighted observations</td>
<td>41138</td>
<td>559704</td>
</tr>
</tbody>
</table>

†See Rosenbaum and Rubin (1985) for a definition of the standardized difference. They considered an absolute standardized difference of more than 20 as being ‘large’. The full set of results is contained in Table A.1 in the Web appendix A. The dummy variable ‘health problems’ indicates disabilities and health problems which do not fully prevent the unemployed from working 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 incapability to work because of other reasons.

(pseudo)redemption. In Figs E.1–E.7 in the Web appendix E, we show the results without controlling for the elapsed duration of unemployment. 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 duration of unemployment suggest overall lower returns to training. These findings are in line with the presumption that not accounting for the elapsed duration of unemployment biases the results negatively (or positively for unemployment).

4.3. Descriptive statistics
The baseline sample includes 93,016 (or 600,842 weighted) observations (The IAB provided a data set that contains 230,842 (or 3,638,851 weighted) observations. This sample is representative
Fig. 1. (a) Mean employment and (b) mean unemployment (group means after using radius matching to adjust the covariate distributions of all groups to that in 'voucher award'): ––––, voucher award; –––––, voucher redemption; – – – – –, expired voucher; – – – – –, control group

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 3 months of unemployment, unemployment durations of at least 1 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 individuals a potential treatment within the first 12 months of their unemployment spell. Further, we drop individuals with marginal, seasonal or subsidized employment before their last unemployment spell. This leaves 124,696 observations. Another 31,680 observations were dropped because of the definition of the pseudovoucher award and redemption times. See the descriptive statistics of the initial and final sample in Table A.2 in the Web appendix A.) 41,138 observations include an awarded voucher in 2003 or 2004, whereas 51,878 do not. Of the former group, 33,077 individuals redeemed their voucher, whereas 8,061 did not. Table 1 reports the means of selected observed characteristics across groups defined in terms of treatment and mediator states (see the on-line Table A.1 for a more extensive set of variables): voucher awarded, no voucher awarded, voucher redeemed and voucher expired (note that the last two groups are subsamples of the first). Pairwise standardized mean differences (see Rosenbaum 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 duration of unemployment and remaining eligibility for unemployment benefits do we consider the measurements at the time of the (pseudo)voucher award.

The descriptive statistics in Table 1 reveal that voucher recipients (column (1)) and non-recipients (column (2)) differ importantly in several socio-economic characteristics, such as age, health, education and profession. In particular, those awarded vouchers are younger, healthier and better educated, and have higher paying jobs. However, elapsed time in duration of unemployment 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 (column (3)) with those whose voucher expires (column (4)), differences in socio-economic variables are small, with the important exception that the latter group is likely to be suffering from incapacities (and health problems in general), which may importantly drive non-redemption. Furthermore, although 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.

Figs 1(a) and 1(b) 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. (Further outcome variables are presented in the Web appendix A.) 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 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 is driven by merely obtaining a voucher and by actually redeeming it.

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 et al. (2014) and Lechner and Wunsch (2013), both based on German labour market data, suggest that so-called pretreatment outcomes (e.g. lagged employment and wages measured before 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 4 years before 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. Although Doerr et al. (2016) argued that a voucher award involves a similar selection process to that of assignment to ALMPs in general, they also noted that the decision is left to the discretion of the caseworker. Our data also contain information that was collected by the caseworker for use in counselling and assignment decisions, namely information on the jobseeker’s current and previous health status, proxy variables indicating whether an unemployed person lacks motivation (e.g. whether she or 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 that are associated with motivation and compliance in the counselling process (approximated by benefit withdrawal and programme dropout) should affect participation. Conditionally on the available covariates, exogenous variation is likely to come 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 courses (for details, see the discussion in Doerr et al. (2016)). This is a lengthy and inflexible administrative process which is unlikely to be correlated with the individual redemption decision or with the employment outcomes of voucher recipients.

Given that vouchers must be redeemed within 1 week–3 months, time varying (or dynamic) confounders of redemption due to important changes in control variables after voucher award but before 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 0. (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 influences 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 3 years; see Section 6.) Further, we report the means of the time varying covariates in Table A.3 of the Web appendix A. We find strong differences in the remaining eligibility for unemployment benefits and in the elapsed duration of unemployment. 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 previous 4 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 the Web appendix B for the full set of control variables).
Table 2. Selected average marginal effects from propensity score estimation†

<table>
<thead>
<tr>
<th></th>
<th>Award probability</th>
<th>Redemption probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Marginal effect (%)</td>
<td>Standard error, (1)</td>
</tr>
<tr>
<td>Individual characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>−0.034‡</td>
<td>0.0001</td>
</tr>
<tr>
<td>Older than 50 years</td>
<td>−9.86†</td>
<td>0.0031</td>
</tr>
<tr>
<td>Children under 3 years</td>
<td>1.14‡</td>
<td>0.0014</td>
</tr>
<tr>
<td>Health problems</td>
<td>−3.78‡</td>
<td>0.0028</td>
</tr>
<tr>
<td>Incapacities</td>
<td>−2.95‡</td>
<td>0.0016</td>
</tr>
<tr>
<td>No German citizenship</td>
<td>−1.65†</td>
<td>0.0022</td>
</tr>
<tr>
<td>No schooling degree</td>
<td>−2.80‡</td>
<td>0.0027</td>
</tr>
<tr>
<td>University entry degree (Abitur)</td>
<td>0.743‡</td>
<td>0.0021</td>
</tr>
<tr>
<td>Elementary occupation</td>
<td>0.299</td>
<td>0.0026</td>
</tr>
<tr>
<td>Craft, machine operators and related</td>
<td>0.743‡</td>
<td>0.0022</td>
</tr>
<tr>
<td>Clerks</td>
<td>3.98†</td>
<td>0.0022</td>
</tr>
<tr>
<td>Individual labour market history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Half-months employed in last 2 years</td>
<td>−0.037</td>
<td>0.0003</td>
</tr>
<tr>
<td>Half-months out of labour force in last 2 years</td>
<td>−0.071§</td>
<td>0.0004</td>
</tr>
<tr>
<td>Remaining unemployment insurance claim</td>
<td>0.150‡</td>
<td>0.00005</td>
</tr>
<tr>
<td>Cumulative half-months employment in last 4 years</td>
<td>0.034‡</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cumulative earnings in last 4 years</td>
<td>0.00001‡</td>
<td>2.1 × 10⁻⁸</td>
</tr>
<tr>
<td>Regional characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of employment in construction</td>
<td>5.17</td>
<td>0.0606</td>
</tr>
<tr>
<td>Share of vacant full-time jobs</td>
<td>0.512</td>
<td>0.0074</td>
</tr>
<tr>
<td>Population per km²</td>
<td>0.0004‡</td>
<td>63 × 10⁻⁷</td>
</tr>
<tr>
<td>Unemployment rate (%)</td>
<td>−0.037</td>
<td>0.0003</td>
</tr>
<tr>
<td>Unconditional probability (%)</td>
<td>6.85</td>
<td></td>
</tr>
<tr>
<td>Sample size (weighted)</td>
<td>93016 (600842)</td>
<td>41138 (41138)</td>
</tr>
</tbody>
</table>

†Probit model used. Heteroscedasticity robust standard errors are in parentheses. The complete set of variables is contained in Table B.1 in the Web appendix B. The dummy variable ‘health problems’ indicates disabilities and health problems which do not fully prevent the unemployed from working, 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 incapability to work because of other reasons.

‡Significant marginal effects at the 1% level.
§Significant marginal effects at the 5% level.
§§Significant marginal effect at the 10% level.

5.2. Empirical selection into award and redemption

Table 2 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 the Web appendix B for a full set of results). Largely these results confirm the pattern of univariate comparisons in Table 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 in the Web 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-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.

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 3-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 certifies eligibility for courses with only 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 choices of course. The maximum number of possible choices of course is 712. Fig. F.1 of the Web appendix F provides a histogram of the observed choices of course. This figure documents a large variation in the maximum number of possible choices of course per awardee. In Fig. F.2 of the Web appendix F, we report the time variation in the average number of maximum choices of course. 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 the Web appendix F, we regress the maximum number of choices of course on different regional characteristics and dummies for the month of the voucher award. We find strong variation in the maximum number of choices of course across states. We find that the maximum number of choices of course 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 the Web appendix F, we estimate the redemption probability with respect to the maximum number of choices of course. We find that this maximum number 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 rates of unemployment. 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 the Web 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-
Direct and Indirect Effects of Training Vouchers

redeemers a 22% probability of receiving a second voucher within 1 year. The control group includes individuals not receiving a voucher during the first 12 months of unemployment. Only 0.6% of individuals receive a voucher later. Individuals in the control group and non-redeemers have an approximately 1-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 do training participants and non-redeemers. These programmes are designed for unemployed workers with problems reintegrating into the labour market. They are a different target group from that for vocational training, which is supposed to have high chances of re-employment. Individuals in the control group and non-redeemers have a higher probability of receiving government grants than do training participants. Government grants are work subsidies. Typically these are start-up grants to become self-employed or subsidies paid to the employer for hiring unemployed workers.

6. Results

6.1. Main findings

The propensity score estimates that were presented in Table 2 serve as inputs into the matching algorithm. When performing matching, one should check for potential issues of

(a) 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

(b) 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 are shown in Figs B.1 and B.2 in the Web appendix B). Furthermore, the important covariates are well balanced (for details see Table B2 in the Web Appendix B).

Fig. 2 provides the estimates of the average employment and unemployment effects on voucher recipients, namely the (total) effects of voucher award versus non-award, Δ, and the effects of voucher award and redemption versus non-award, θ, voucher award without redemption versus non-award, γ, and voucher award with redemption versus voucher award without redemption, δ. Concerning employment, we consider only (non-marginal, non-subsidized) employment lasting at least 1 month. The curves 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 symbols superimposed on the curves (diamonds) indicate effects that are (pointwise) statistically significantly different from 0 at the 5% level.

The results in Fig. 2(a) suggest that awarding a voucher has a negative (total) employment effect among voucher recipients in the first 3 years, particularly in the initial months where the probability of employment decreases by as much as 10 percentage points. This dip suggests a lock-in effect that is likely to be due to reduced job search intensity in response to (anticipated) participation in vocational training. However, the negative effect fades over roughly 3 years, and the probability of employment increases by approximately 2–3 percentage points in the fourth year. The positive employment effect appears quite stable suggesting that the voucher award system successfully offsets the initial lock-in effect with higher placement success in later
Fig. 2. (a) Employment and (b) registered unemployment (separate effects for the first 48 months following the voucher award are estimated): \( \ast \), significant effects at the 5% level; \( \longrightarrow \), voucher award versus control group; \( \cdots \cdots \cdots \), voucher redemption versus control group; \( \cdots \cdots \), voucher redemption versus expired voucher; \( \cdot \cdot \cdot \cdot \cdot \cdot \), expired voucher versus control group.

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.

Fig. 2(b) 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 dropout from the labour market as shown by the effect on the 'out of labour force' labour
market state (see Fig. C.2 in the Web 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.

When investigating the causal mechanisms underlying the total employment effect (with essentially symmetric results for registered unemployment), it becomes apparent that it is predominately redemption (e.g. participation in or starting vocational training) driving the results. In fact, the estimated effect of voucher award and redemption vs no award, $\theta$, closely follows the overall effect of voucher award, albeit that 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, $\gamma$, is insignificant and close to 0 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 probability of employment in the first 3–3.5 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 et al. (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 that are 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 et al. (2014), who reported negative effects of notifications of planned training on exits from employment. They argued that 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 et al. (2009)).

We would expect this initial direct lock-in effect to be less severe than for the total effect (which includes the effect 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 $\gamma$ is considerably less negative than the estimates of $\Delta$ and in particular of $\theta$ (redemption vs non-award). Accordingly, the estimate of $\delta$ (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 2 years, the estimates of $\Delta$ and $\theta$ dominate those of $\gamma$, and the estimates of $\delta$ 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 the Web 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) effect of voucher award vs non-award, $\Delta$, and voucher award and redemption vs non-award, $\theta$, 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 with Fig. 2(a) the estimate of the direct effect $\gamma$ remains statistically significantly negative until the end of the evaluation window (implying that the adverse effect of not redeeming a voucher vs not receiving a voucher 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 that are shown in Fig. 2(a), 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 $\Delta$, $\theta$ and $\delta$
are moderately positive (between €30 and €70) and statistically significantly in the fourth year, whereas those of $\gamma$ approach 0.

6.2. Cumulative effects

In Table C.1 of the Web 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 with the control group during our observation period. We do not observe significant employment and earnings effects of redeeming versus not redeeming a voucher after 4 years. Redeeming a voucher increases or decreases the average probability over 4 years of being unemployed or out of the labour force respectively.

After 2 years, we observe more negative indirect than direct voucher effects. Over the long run, this relationship reverses but does not pay in terms of cumulated employment or earnings during our observation period. This suggests that the benefits of redeeming and not redeeming a voucher are, on average, similar over the 4-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 4-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 with the group who redeemed them (see Figs C.6–C.12 in the Web 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 Figs C.13–C.19 in the Web appendix C). These results are driven by full-time employment. This suggests that 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 chances of employment of voucher recipients. We also investigated the causal mechanism through which the overall effect materializes by using sequential conditional independence assumptions for identification. In particular, we considered the direct employment effect of voucher assignment (net of actual
Direct and Indirect Effects of Training Vouchers (versus 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 with the direct effect suggests that, conditionally 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 2–3 percentage points compared with 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 between training providers, may lead to a loss of effectiveness if individuals do not make use of the awards, because non-redemption entails lower chances of employment 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 ALMPs by using a voucher system.

Acknowledgements

Michael Lechner is also affiliated with the Centre for Economic Policy Research London, CESifo, Munich, the IAB, Nuremberg, and the Institute for the Study of Labor, Bonn. Anthony Strittmatter is also affiliated with the Albert-Ludwigs-Universität Freiburg. We benefited from comments by Annabelle Doerr, Bernd Fitzenberger, Thomas Kruppe and Gesine Stephan. Financial support from the Swiss National Science Foundation, grant 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 1155. We gratefully acknowledge data support from the IAB. Anthony Strittmatter acknowledges financial support from the IAB. The usual disclaimer applies.

References


**Supporting information**

Additional ‘supporting information’ may be found in the on-line version of this article:

‘Online appendix’.