Vocational Training for Female Job Returners - Effects on Employment, Earnings and Job Quality*

Annabelle Doerr

University of Basel

January 30, 2021

Abstract

This paper studies how training vouchers increase the employment prospects of women with interrupted employment histories. Using the population of female job returners who receive a training voucher to participate in training programs and a randomly selected control group from German administrative data, I analyze the effectiveness of training on various labor market outcomes. The results suggest that receiving a training voucher translates into substantial gains in employment and earnings and increases job quality and stability. Analyzing the heterogeneity effects reveals that the effectiveness of training increases with the provided human capital. Several robustness checks support a causal interpretation of the results.

JEL-Classification: J68, H43, C21

Keywords: Job Return, Active Labor Market Policies, Treatment Effects, Program Evaluation, Training Voucher, Return to work

*I gratefully acknowledge very helpful comments from Conny Wunsch, David Card, Bernd Fitzenberger, Anthony Strittmatter and Thomas Kruppe as well as two anonymous referees. I want to thank for all the comments of participants at various seminars and conferences. This study is part of the project "Regional Allocation Intensities, Effectiveness and Reform Effects of Training Vouchers in Active Labor Market Policies", IAB project number 1155. I gratefully acknowledge financial and material support from the IAB. Furthermore, I gratefully acknowledge financial support of the Akademie der Wissenschaften Leopoldina for funding my research visit at UC Berkeley. The usual caveats apply.

1 Introduction

A growing literature shows that family-related employment interruptions negatively affect the subsequent working careers of women (see, for example, Drange and Rege, 2013, Evertsson, 2016, Beblo et al., 2009) and that interruption periods contribute to the gender gap in wages and income (see, for example, Angelov et al., 2016, Goldin, 2014). These adverse consequences are traditionally explained by human capital theory, which suggests a reduction of human capital accumulation or even a depreciation and atrophy of job-related skills when the interruption periods are long (Mincer and Polachek, 1974). Empirical studies also document a lower occupational mobility following interruption periods (see, for example, Evertsson et al., 2016, Evertsson and Grunow, 2012) and a reduction in the probability of receiving job-related training (Puhani and Sonderhof, 2011). Goldin (2014) provides a different explanation for the negative effects from employment interruptions. Independent of the level of human capital, interruption penalties in wages differ by the production and compensation structures across occupations and industries.

Another strand of the literature studies policy instruments influencing the length of employment interruptions after giving birth. This literature focuses on the effects of parental leave and job protection legislation on return behavior and female employment outcomes (most recently, e.g., Kluve and Schmitz, 2018, Schönberg and Ludsteck, 2014, Lalive et al., 2014, Lalive and Zweimüller, 2009) and discusses how to implement legislation to make long employment interruptions less attractive. In a recent study focused on Germany, Kluve and Schmitz (2018) document that a reform of the parental benefit system that introduced very generous but temporally limited benefit payments that could be divided between parents changed the interruption behavior of women and positively affected employment rates and job quality after return. However, women career patterns remain characterized by interruption periods and episodes of part-time work that increase the gender wage gap over the employment life cycle (Schrenker and Zucco, 2020). Thus, the re-integration of women into the labor market after employment interruptions is of high economic importance, not only in light of recent skill shortages and the increasing demand for skilled labor but also because of its impact on the gender wage gap and the negative long-term consequences for women facing a remarkably higher risk of poverty at retirement if their working life is characterized by long interruptions.

In this paper, I investigate the impact of publicly funded vocational training for women who intend to re-enter the labor market after family-related interruptions. Thereby, I contribute to the literature by estimating the effectiveness of a policy instrument that could alleviate gender penalties in wages and employment when the interruption period has already occurred. Women can return to the labor market in different ways. I focus on the subset of re-entrants that register at a local employment agency in Germany to use the services provided by the public employment service (PES). Re-entrants with a family-related interruption period of at least 12 months are defined as "job returners", and are eligible for placement services and active labor market policies (ALMP). Thereby, German legislation recognizes the relevance of a successful re-integration of returning women.

Knowledge is limited about the effectiveness of instruments targeted to women returning to the labor market. Bergemann and van den Berg (2014) evaluate an adult education program for young, low-skilled mothers with small children in Sweden. This program purposes to provide general skills to low-educated individuals to improve their employability. The findings suggest positive effects on wage and employment rates; however, enrollment is low. The authors explain the low enrollment rate with non-pecuniary factors and a strong preference of young mothers to stay home with the infant. In recent years, the German Federal Employment Agency initiated a modular program targeting women considering a return. The modules consist of informative events, short training measures and intensive counseling. Diener et al. (2013) report significant positive effects on the probability of working full-time for the women who finished all the modules.

In line with the existing evidence, the training programs considered here are expected to improve the labor market prospects of job returners. These programs provide occupation-specific knowledge and job-related skills to participants, and, in long programs, participants can obtain a (new) vocational degree. Thus, training participation is a considerable investment in human capital, and the certificates are an important signaling device for potential employers. Furthermore, the training is an opportunity to meet women in similar situations and to build networks that are found to positively affect labor market opportunities (e.g., Calvo-Armengol and Jackson, 2004). Especially after long interruption periods, these programs may provide a first chance to make contact with potential coworkers. Program participation may also serve as an exercise in managing work and family life with respect to time constraints and childcare facilities.

In Germany, publicly funded vocational training is assigned through training vouchers. A challenge in this study is that the assignment of these vouchers is likely to be endogenous. The responsible caseworkers in the agency decide whether to award the voucher and are likely to base their assignment decision on the women's job-finding probability after training participation. To identify the causal effects of being awarded a training voucher, I apply a matching strategy to account for selection based on observable characteristics. I implement the analysis using the population of female job returners who receive a training voucher from 2003-2005 and a random sample of non-treated job returners from German administrative data. These data contain daily records of the pre- and post-interruption periods as well as a large set of characteristics including the employment histories of the women before their labor market exit. Conditional on these determinants, I argue that the job returners are randomly awarded a training voucher. Many robustness checks support the identifying assumptions and justify a causal interpretation of the findings.

The results reveal that training programs are effective in improving the labor market outcomes of women after family-related withdrawal from the labor force. Awarding a training voucher translates into a 12 percentage points (ppoints) higher probability of employment and higher monthly earnings of 210 Euros/250 US dollar (USD). Additionally, job quality and stability increase. Job returners who receive training vouchers have a significant higher probability of working full-time and receiving at least their previous earnings. Moreover, the probability of experiencing unstable marginal employment decreases permanently, and treated job returners are more often employed in stable jobs.

Analyzing effect heterogeneity in terms of qualification of the job returners and their participation in different program types, reveals interesting insights. I find a training voucher to be most effective for job returners without a vocational degree. But it also contributes to a successful re-integration of medium-skilled women with a vocational degree. For the highest skilled job returners with academic degree, awarding a training voucher is ineffective. The results strongly implicate that the effects of training vouchers work through the human capital channel. A decomposition of the overall effect by the redemption decision reveals a zero effect for unredeemed training vouchers. Conditional on voucher redemption, the effectiveness of training increases with the intensity of human capital provided in the different program types.

This paper is related to several fields in the literature. It is the first study to estimate the impacts of vocational training for female job returners on their subsequent labor market outcomes. Thus, this study relates to the literature evaluating active labor market programs focusing on women (for an overview, see, Bergemann and van den Berg, 2008) and especially to the rare literature that focus on programs targeted to women after family-related employment interruptions. In contrast to the studies by Bergemann and van den Berg (2014) and Diener et al. (2013), I focus on a sample of women that has a manifested return idea. Moreover, none of the programs evaluated in these studies are comparable to vocational training programs in terms of content or intensity.

This paper also contributes to the large body of literature concerning the effectiveness of vocational training for unemployed individuals in Germany, other EU countries (e.g., Doerr et al., 2017, Lechner et al., 2011) and the US (e.g., Andersson et al., 2013, Heinrich et al., 2013). A detailed overview of the effectiveness of these programs is provided by McCall et al. (2016). Nevertheless, by definition, these studies exclude the population of job returners because they restrict their evaluation samples to individuals entering unemployment directly out of employment. Observations of job returners that enter unemployment out of inactivity are not considered. Doerr et al. (2017) study the effectiveness of training vouchers in the same period of time for the regular unemployed. The authors report significant positive effects on the employment probability and wages that are much lower in size than the effects presented here. They find no significant differences by gender.

More broadly, this study also relates to the literature that focuses on the effects of parental leave and benefit policies on interruption durations and female labor market outcomes. Kluve and Schmitz (2018), Schönberg and Ludsteck (2014), Lalive et al. (2014), Dahl et al. (2016) for example, investigate such policies in European countries and, e.g., Rossin-Slater et al. (2013), Baker et al. (2008) do so in the US and Canada. In contrast, this paper focuses on how an active labor market policy, namely, the opportunity to receive vocational training, affects the employment situation of women after family-related employment interruptions that last beyond the usual maternal leave period.

The remainder of this study is structured as follows. The background information and a description of the institutional setting follow in Section 2. The data description and sample definitions are presented in Section 3. I discuss the empirical strategy and show descriptive statistics in Section 4. Finally, I present and discuss the empirical results with regard to the different outcomes and effect heterogeneity. The final section concludes.

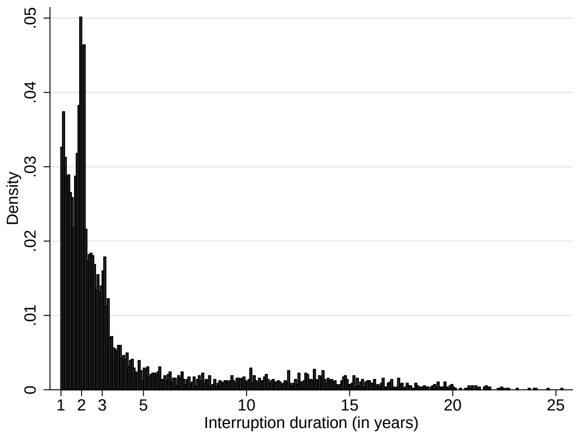
2 Institutional background

2.1 Definition of a job returner

Women have several options when they re-enter the labor market after a family-related withdrawal from the labor force in Germany. Re-entrants who interrupted a regular employment relationship could, in principal, return to their former employer if the interruption had a maximum duration of three years.[1] Re-entrants for which this is not an option could either search for a job by themselves or use the services of the PES. In this

[1]Maternity allowance is paid in Germany within a period of 6 weeks before and 8 weeks after giving birth. In that time, mothers are not allowed to work. During the time considered in this study, the income-independent childcare payment (*Bundeserziehungsgeldgesetz*) was paid up to two years after giving birth and was independent of the mothers' employment state. This instrument was replaced by the *Elterngeld* in 2007, which is paid for 12 months only, is income dependent, and can be split between couples.

Figure 1: Distribution of interruption duration



Note: Each bar represents one month. The abscissa is labeled in years. The minimum interruption duration according to the formal definition for job returners is 12 months. The maximum duration observed in the data is 304 months (25 years).

study, I focus exclusively on women belonging to the latter category. The PES categorizes these re-entrants as job returners (in German: *Berufsrückkehrer*).[2] A job returner (according to §20 German Social Code (SGB) III) is defined as women (or men) who have interrupted employment, unemployment or an apprenticeship for at least 12 months to care for their children (younger than 15 years) or other family members and who show the intention to become employed in the near future.[3] Figure 1 displays the distribution of the interruption duration of job returners analyzed in this study. The distribution is clearly skewed to the right. Approximately 65% of job returners re-enter the labor market within a period of three years.

To use the services offered by the PES, re-entrants have to register at a local employment agency for one of three types of registration: (1) seeking advice, (2) searching for employment, or (3) unemployed. These registration types are the same for all individuals that register at the PES and are not tailored to the group of re-entrants. The status

[2] The wording "job returner" is somewhat imprecise since the women do not necessarily return to a job. However, to be defined as a job returner, they must show a serious intention to start working within the near future. Furthermore, the PES uses this terminology. Therefore, I use "re-entrants to the labor market" and "job returners" interchangeably throughout the text.

[3] Since 98% of job returners are women, I focus on female job returners only. No information is available about the reason for the family-related interruption. In a survey conducted by Diener et al. 2013, 78% of surveyed women gave the upbringing of their own children as the reason for the interruption, 8% had to care for relatives, and some women gave both reasons. Other reasons that were mentioned were marriage and participation in a family business.

definition narrows from registration state (1) to (3) and the numbers of available services and benefits increase. Whereas registration (1), seeking advice (in German: *ratsuchend*), allows the re-entrant to use information and counseling services only, individuals that also want to use the placement services must at least register as (2), searching for employment (in German: *arbeitsuchend*).[4] Finally, registration (3), unemployed (in German: *arbeitslos*), is the requirement to obtain unemployment benefits and for unemployed individuals only. It opens the possibility of participating in ALMP programs and contributing to the pension system, but it also imposes obligations. Such registered individuals must actively search for employment for at least 15 hours a week, attend regular meetings with a caseworker and be available if caseworkers present a suitable job offer irrespective of whether they obtain unemployment benefits. An overview of the available services corresponding to each registration state is provided in Table A.1 in Appendix A.

The initial registration of re-entrants likely depends on their eligibility to receive unemployment benefits. They could be eligible to receive benefits if they were regularly employed directly before the interruption because entitlements obtained from pre-interruption employment persist in the case of family-related interruptions until the child turns three years old. If a job returner is eligible to receive unemployment benefits, she must register as being unemployed to receive them. Job returners who are not eligible for unemployment benefits may choose between the three registration states. Two-thirds of the job returners investigated in this paper registered as being unemployed upon their return. The other job returners registered as searching for employment. No job returner was registered as seeking advice only. Ebach and Franzke (2014) investigate the motivations behind the registration behavior of job returners. The authors find that job returners switch between states and report that the registration in the pension system and the opportunity to participate in ALMP programs are the driving forces of an unemployment registration even when registrants do not receive benefits. Because the registration status

[4]Individuals registered in states (1) or (2) could be employed or unemployed. Registration (1), seeking advice, allows individuals with a more diffuse job search intention to register and receive an initial counselling service only. Registration (2), searching for employment, would be optimal for the unemployed who do not receive benefits or the employed searching for a new job because they wish to switch firms or are at risk of losing their jobs.

determines the services that can be accessed and is crucial to the identification strategy, I discuss it in further detail in Section 4.2.

2.2 Vocational training

Vocational training is an important part of ALMP in Germany and the most important instrument among the policies focussing on qualification. During the years 2003-2005 the German government spent overall more than 10 billion Euros for vocational training. These programs are intended to increase the human capital of the participants by providing, maintaining and updating occupation-specific skills. Additionally, the obtained certificates may serve as an important signaling device for potential employers. Therefore, such training appears to be a particularly suitable instrument for increasing the re-employability of job returners.

The provision of vocational training by local employment offices is organized through a voucher system.[5] Once the re-entering women register at a local employment office and receive consultation, the caseworker may decide that vocational training is the appropriate program for improving her employability. If the woman agrees and is not yet registered as unemployed, she may switch her registration state to become eligible to participate in ALMP programs. Then, the woman is awarded a voucher enabling her to participate in a training program. The caseworker notes the educational objective, maximum program duration and validity on the voucher.[6] The job returner may choose a program offered by a certified training provider subject to the restrictions noted on the voucher. Although the award of a training voucher is at the discretion of the caseworker, its redemption cannot be controlled. The voucher redemption rate during the observation period is 90%.

Training programs differ with respect to content and duration. Following Lechner et al. (2011), I distinguish among practice firm training, classical vocational training and retraining. Practice firm training programs are rather short and simulate a work envi-

[5] The voucher system was introduced in January 2003. Previously, the provision of vocational training functioned through a direct assignment system by caseworkers. The reform of the provision of vocational training and the institutional changes are discussed in Doerr et al. (2017).

[6] In addition, information regarding the funding and commuting zone and various other information is included. For a detailed description of the voucher system, see Doerr et al. (2017) or Rinne et al. (2013).

Table 1: Vocational training programs for job returners

Program type	Average planned duration	Share	Description	Examples
Practice firm training	164 days	13%	Courses that took place in practice firms to simulate a work environment.	Training for commercial software, office clerks, data processing
Short training	127 days	37%	Provision of occupation specific skills (duration $\leq$ 6 months).	Training courses as medical assistants, office clerks, draftsman, hairdresser, lawyer
Long training	285 days	23%	Provision of occupation specific skills (duration $>$ 6 months).	Training as tax accountant, elderly care nurse, office clerks, physical therapist
Retraining	833 days	26%	Courses to obtain a first/new vocational degree.	Apprenticeship as elderly care nurse, physical therapist, hotel and catering assistant
Others	-	1%	e.g. courses for career improvement	

Note: I use the categorization of programs proposed by Lechner et al. (2011) and analyze the information on the training voucher with regard to the content of the training. The presented examples refer to training goals that are often denoted on the training voucher. The category "others" contains different types of training programs with very few participants.

ronment in a practice firm. The educational aims and contents of these programs are to update the skills of job returners mainly in commercial software programs and MS Office applications. Classical training programs provide occupation-specific skills and mainly occur in classrooms. I differentiate between short (maximum duration of 6 months) and long (duration over 6 months) training programs. Typical examples of classical vocational training programs for job returners include programs for office clerks, craftsmen, and medical assistants to adjust knowledge and skills according to recent developments. Retraining programs have a long duration of up to three years.[7] These programs lead to a (new) vocational degree within the German apprenticeship system and cover the full curriculum of vocational training for occupations such as an elderly care nurse or physical therapist. These programs may be particularly relevant for re-entrants with former occupations for which there is low demand at return.

[7] Only the first two years are funded by local employment agencies. Other programs sponsored by the state government may cover the additional costs of the third year.

3 Data description and definitions

3.1 Data and sample description

This analysis is based on administrative data on training vouchers provided by the German Federal Employment Agency. The sample contains information of all individuals in Germany who received a training voucher between 2003 and 2005 and includes the exact award and redemption dates for each voucher. To enrich this voucher data, it is merged with the individual data records from the Integrated Employment Biographies (IEB) for each voucher recipient.[8] The sample of control persons originates from the IEB and is drawn as a 3% random sample of individuals who did not receive a training voucher between 2003 and 2005. I account for the fact that I use a 100% sample of voucher recipients and a 3% random control sample with different sampling probabilities in all calculations.

I restrict the treated and control samples to women who register at local employment offices and are categorized as job returners. The sample does not include women who return to their former employers or search for employment alone. Additionally, women who are self-employed or civil servants are not included in the data. Omitting these subgroups is not problematic since I am interested in estimating the effects of receiving a training voucher. Because eligibility for a voucher depends on registration at local employment agencies, those who do not register are not relevant for the effect estimates of interest in this study.

I consider registrations between January 2003 and December 2004. The reason is that I focus on voucher awards to job returners within the first year after their registration, and my observation period of the voucher data covers only the period until December 2005. I do not restrict the sample based on the employment status before the interruption. This

[8]The IEB is a rich administrative database and source of the subsamples of the data used in all recent studies evaluating the programs of German ALMP (e.g., Biewen et al., 2014, Lechner, Miquel, and Wunsch, 2011, Lechner and Wunsch, 2013, among others). The IEB is a merged data file containing individual data records collected in four 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 an abbreviation for the research department of the German Federal Employment Agency.

method is justified by the formal definition of a job returner that also includes women who were interrupted from unemployment or apprenticeship. Nevertheless, I restrict the sample to women who are 25-49 years at the time of registration.

This study focus on women deciding to return to work in a period of high, but for Germany in the decade between 1995 and 2005, not particularly high unemployment rate. The period under consideration can be seen as the prolonged recession period after the end of the new-economy boom in the early 2000s. The unemployment rate reached its maximum in 2005. As an answer to the increasing unemployment rate, the German labor market was intensively reformed in this time period, particularly in 2005. This situation is not problematic for the empirical analysis because the comparison groups were equally affected by the Hartz reforms. Depending on the length of the training program, most of the job returners that received a training voucher might have finished training in 2004 or 2005 when the economy already started to recover, or, for those who participated in the longer programs between 2006 and 2007, they might have finished when the unemployment rate was falling and the economy was strongly growing. Therefore, the labor market situation was not exceptionally but rather good when participants finished training. The same holds for the comparison group.

3.2 Treatment and outcome definitions

By using these data, I can not only observe the exact start date of the training program but also the exact date when the voucher was awarded, thus, when the assignment into the program occurred. The treatment of interest in this study is the award of a training voucher within the first year after registration at the local employment agency. Using this treatment definition, I estimate the intention-to-treat effect, i.e., the effect of being awarded a voucher (not the effect of participating in a training program). I follow all individuals over a period of more than six years (76 months) which allows me to estimate short-, medium-, and long-term effects.

I use various outcomes to measure the effectiveness of awarding training vouchers to job returners. The two standard outcome measures in the literature are non-subsidized non-marginal employment (henceforth, employment) and monthly earnings. I count a job returner as employed if I have data for non-subsidized and non-marginal employment of at least 31 days. Earnings are calculated as real gross earnings per month. When considering job returners, I find it particularly interesting to determine whether vocational training alters job quality and employment stability. Therefore, I consider the probabilities of being employed full-time, earning at least 100% of the previous earnings (average of inflation-adjusted earnings in the year prior to the interruption), and being marginally employed as additional outcome measures of job quality.[9] I use the probability of finding a job that lasts at least 36 months to evaluate the effects on employment stability.

The award of a training voucher represents the intention to invest in human capital that involves direct and indirect costs. The direct costs of the programs are not observed in the data. Nevertheless, I measure the indirect costs from earning and employment losses during the participation period. Following Lechner et al. (2011), I use the accumulated employment and earnings as additional outcome measures to assess the effectiveness of vocational training for the job returners in terms of net benefits over the long term.

4 Empirical strategy

4.1 Evaluation framework

The definition of the treatment and control groups depends on the evaluation framework, i.e., whether a dynamic or a static concept is used. The static evaluation framework is simple. Its implementation follows in a comparison between those who received a training voucher and those who did not over a certain period (here, 12 months). Hence, this concept ignores the timing of the treatment. Frederiksson and Johansson (2008) argue that doing so may lead to an underestimation of the results because such a constructed control group may consist of individuals that do not receive treatment because they

[9]Marginal employment (*so-called geringfügige Beschäftigung*) is an employment relationship with maximum monthly earnings of 450 Euros for which no social security contributions or taxes are paid. This mostly represents jobs with a low number of hours and low qualification requirements.

already found a job, thereby representing a positive selection of the non-treated with better labor market characteristics than those in the treatment group.[10] However, this argument is not applicable to job returners. They register at local employment agencies to receive counselling and advice for returning to work (often without receiving benefits). Those who do not receive a voucher do not necessarily represent a positive selection.

The dynamic evaluation framework allows for more flexibility but demands a large amount of data. Nevertheless, it captures the timing dimension of the treatment, which may be important. Sianesi (2004) propose to estimate treatment effects conditional on the time elapsed since unemployment, here since return. By implementing this concept, I would estimate the effect of receiving a training voucher in the current month versus not receiving it in the current month but possibly receiving it later. One major concern of dynamic evaluation approaches is raised by Lechner et al. (2011). They argue that the composition of the control sample changes for each month elapsed since return, a fact that hinders the interpretation of the results.

The decision to implement a dynamic or static evaluation approach is mainly driven by data limitations (e.g., small sample sizes), and there are arguments in favor of and against each approach. The numbers of observations in the different evaluation samples are presented in Table 2. Because job returners constitute a relatively small group (compared to the number of regular unemployed individuals), I am faced with restrictions regarding the number of observations.[11] Therefore, I use a static evaluation approach and compare the job returners who received a training voucher within the first 12 months of return to those who did not. To reduce the possibility of bias introduced by a static approach, I randomly assign pseudo treatment start dates to each individual in the control group to partially capture the timing dimension. Thus, I recover the conditional distribution of the time elapsed since return at the treatment start from the treatment group (similar to, e.g.,

[10] This argument derives from the observation that, in countries such as Germany, nearly all unemployed persons would receive a treatment if their unemployment spells were sufficiently long (Frederiksson and Johansson, 2008). Accordingly, individuals who find jobs rapidly are less likely to receive training, as the treatment definition is restricted to unemployment periods.

[11] The share of job returners among all the voucher recipients is approximately 9%. Hence, the number of observations in this analysis is relatively small, although it consists of the population of the job returners that received a training voucher in the years under consideration.

Table 2: Number of observations in the different evaluation samples

	Static evaluation approach			
Time window	**12 months + (pseudo) elapsed duration**	12 months	6 months	3 months
Num. treated	**2,986**	2,986	2,087	1,243
Num. controls	**3,461**	3,748	4,647	5,491

	Dynamic evaluation approach											
Elapsed duration	1	2	3	4	5	6	7	8	9	10	11	12
Num. treated	439	440	360	306	286	252	192	180	154	137	129	107
Num. controls	6,269	5,746	5,323	4,954	4,605	4,307	4,066	3,845	3,652	3,482	3,322	3,186

Note: The main sample is constructed using a static evaluation approach over a period of 12 months. I condition on simulated start dates for non-treated job returners (pseudo elapsed duration). The employment and earnings effects using all different evaluation approaches are presented in Figure B.1 in Appendix B.

Lechner and Smith, 2007). To make the treatment definitions comparable between the treatment and control samples, I consider only individuals who remain unemployed at the start of their (pseudo) treatment. I use this constructed duration to control for the fact that immediately treated job returners are not necessarily comparable to job returners who are treated later. In total, 2,986 job returners who received a training voucher (treatment group) and 3,461 job returners who received another or no program (control group) are included in the main evaluation sample.[12] The employment and earnings effects using all evaluation approaches are presented in Figure B.1 and discussed in Appendix B.

I present a conservative estimate of the treatment effect relative to the never treated sample because the outcomes of the control group include those who receive a training voucher beyond one year after registration in the PES. This phenomenon reveals another weakness of the static approach: the treatment effect is re-defined each time the considered period in which treatment may occur changes. Another strand of the literature acknowledges this problem and identifies the treatment effect relative to the never treated (see Crépon et al., 2009, van den Berg and Vikström, 2019, Vikström, 2017) by right-censoring the outcomes of the control sample when entering treatment. This approach is not feasible in this study because voucher assignments are observed in the data from 2003-2005 only. However, I observe a voucher assignment beyond 12 months after the

[12] In fact, 93% of the control group did not participate in any other ALMP programs in the first 12 months after registration. The remaining 7% mainly participated in short training and application measures or received targeted wage subsidies.

registration period for 667 job returners. I implement a robustness analysis in which I exclude these women from the control group. As an additional check, I exclude from the control sample those who participate in another ALMP program within and beyond the first 12 months after registration. The results remain very similar when I use these alternative control groups and are presented in Figure B.2.

4.2 Identification

Randomized trials are the gold standard for determining causality. However, randomization rarely occurs when public policies are implemented, and training vouchers are not randomly assigned to job returners. One strategy to address selectivity into treatment is to rely on a selection-on-observables identification, which is motivated by the richness of the data used in this study.[13] To identify the treatment effect of interest, which is the effect of awarding a training voucher to job returners in the first year after their registration, I control for pretreatment variables X that jointly influence the treatment assignment and the outcome. Then, conditional on these variables, the voucher assignment and the potential outcomes of job returners are independent. This condition is established as the conditional mean independence assumption (CIA). Furthermore, the identification requires overlap in the distributions of the propensity scores between the treated and control samples (see the discussion in Lechner and Strittmatter, 2019). The formalization of the treatment effect and the identifying assumption are shown in Appendix C.

In the following, I argue that the CIA is satisfied in the case of assigning training vouchers to job returners. Doing so requires that I can control for all factors that jointly determine the award of a training voucher and the potential outcomes in the estimations. The voucher assignment is determined by the eligibility of the job returners to receive a voucher, the assignment decision of caseworkers, and the re-entrants motivation to return.

The eligibility for a training voucher depends on the job returners' registration status. Only a registration as unemployed renders them eligible to receive a voucher, which may

[13] Another way to address selection into treatment could be to find a strong and convincing instrument that generates random variation. Such an instrument is not available in this setting.

induce a bias if the job returners select into different registration states. Ebach and Franzke (2014) conducted a case study with caseworkers and job returners. The authors report that registered job returners are often not informed about the differences between the registration states and observe many switches among the states. In this sample, the share of job returners that register as unemployed directly after return amounts to 64% and 68% in the control and treatment group, respectively, and, I observe up to four switches between registration states in both groups. I control for possible selection from the registration process by including information on the registration status directly after return and eligibility for unemployment benefits in the estimations. Benefit-eligible job returners likely register as unemployed mainly for monetary reasons rather than for the possibility of participating in training.

Job returners differ in their education, occupational background, pre-interruption attachment to the labor market and family situation. All these factors are observable to the caseworker and play a role in the decision to assign a training voucher. Table 3 summarizes the potential confounders, their availability in the data, and how they are included in the estimation. I condition on variables that describe the job returners' personal, educational and vocational background in all estimations. Furthermore, I control for their employment histories and their employment status directly before interruption. The probability of receiving a training voucher also depends on regional labor market conditions. Caseworkers must consider the local labor demand in their assignment decisions. Therefore, I include detailed regional information and condition on state dummies and the share of the employed in the production industry per local employment district.

The group of registered job returners is quite homogeneous with respect to their motivation to return. The reason is that registration as a job returner produces obligations (e.g., attend regular meetings with a caseworker and be available for job offers), but, in contrast to regular unemployed persons, job returners are mostly not eligible to receive benefits. In this sample, only 25% of the job returners receive benefits after registration. Therefore, I expect that selection based on motivation is less severe compared to that in

Table 3: Overview of potential confounders and available and used control variables

Potential confounders	Information available in the data	Control variables included in the estimation
Return process	Daily spell information, registration states for each spell	Elapsed duration until treatment, dummies for registration as unemployed directly after return, eligibility for unemployment benefits, quarter dummies of return
Personal characteristics	Age, nationality, health problems, family status, number of children (living in the household), birth dates of the youngest child, Missing: partner information	Dummy for health problems, dummy for being married, dummy if child is between 1-3 year old, dummy if child is between 3-5 year old
Educational background	Educational degrees, vocational degrees	Dummies for educational degree at return, dummy for no vocational degree
Former occupation	Occupational classification (FEA 2010), working time, working position	Dummies if last reported occupations was in business, accounting or law, dummy if last occupation was white-collar work, dummy if last occupation was full-time
Employment history	Employment spells, unemployment spells, spells with benefit receipt, daily wage	Dummy for being unemployed directly before return, months unemployed the last 6 years before interruption, cumulated wages and benefits in the last 6 years before interruption
Regional characteristics	Regional identifiers, number of employed per industry, unemployment rate, population density, number of vacant full-time jobs	State dummies, share of employed in production industry per local employment district
Attitude/Motivation	–	Interruption duration in months

Note: Because of the sample size and potential multi-correlation, I did not include all variable that are available in the data in the estimations. I followed Huber et al. (2015) and used omitted variable test to select the variables with the largest explanatory power. In the descriptive statistics, I show that also the omitted variables are balanced after matching.

studies focusing on regular unemployed individuals. Despite the homogeneity with regard to their motivation to return, the job returners might differ considerably in factors that are not directly observable in the data, such as their abilities, toughness, attitudes toward employment or situation with a partner and family. These factors influence the length of a woman's employment interruption and the expected benefits of training. I account for them by conditioning on the interruption duration as an additional control variable.

Past studies have assessed the CIA in the context of the evaluation of policy programs in the US (e.g., Heckman et al., 1999, Mueser et al., 2007) and Germany (e.g., Biewen et al., 2014, Lechner and Wunsch, 2013) and support its plausibility in cases of flexible conditioning on the labor market history, personal characteristics, and regional labor market information. These studies focus on the plausibility of the CIA with respect to participation in a certain program. In this study, I aim to identify the effects of the intention to participate in training (being awarded a voucher) instead of the participation decision. I argue that the selection process used to identify the intention to treat effect is less demanding because starting the program is not part of the selection. Thus, the CIA must not hold for the selection process of voucher redemption.

To further check the robustness of my results and support a causal interpretation, I perform several robustness checks. First, I show the results of an impact estimation of pretreatment (here, pre-interruption) outcomes. If there are unmeasured constant differences between the treated and control individuals, I would observe them in the effect estimates before interruption. Second, I implement a placebo validation in the pre-interruption period to show that I condition on all the variables that are necessary to balance the employment outcomes of the treated and control samples. Third, another robustness check is performed by estimating the treatment effects for the job returners who do not redeem their training vouchers. If effects exist other than the one I aim to measure (e.g., threat or anticipation effects), they should be discernible in the group of job returners with unredeemed training vouchers. The results of all the robustness checks are presented in Section 5.3 and strongly support a causal interpretation of the results.

4.3 Estimation and inference

I apply radius matching on the propensity score with bias adjustment as proposed by Lechner et al. (2011). Thus, I compare the outcomes job returners that are similar with respect to their probability of being treated conditional on all observed characteristics but that differ only in terms of receiving voucher.[14] To reduce the bias that may result from incorrectly specifying the propensity score model, I additionally match on a Mahalanobis distance specified by selected control variables in addition to the propensity score (see discussion in Lechner et al., 2011, Huber et al., 2014).

All variables included in the specification of the propensity score and added as control variables in the Mahalanobis distance are listed in Table 3. The inference is based on the standard errors of the estimated parameters obtained by nonparametric bootstrapping

[14]Radius matching uses a one-to-many matching algorithm that matches each treated unit to control units located at a predefined distance around its propensity score (Dehejia and Wahba, 1999, 2002). Radius matching is found to perform well in an empirical Monte Carlo simulation study conducted by Huber et al. (2013), highlighting its good performance relative to linear regressions, especially in evaluation studies that rely on small samples.

(sampling individual observations with replacement) with 499 replications.

4.4 Descriptive statistics

I summarize the sample characteristics available in the data in Table 4 and Table D.1 in Appendix D. I report the means of the treated and control samples as well as the standardized differences before and after matching for both the control variables used in the estimation (see column (3) of Table 3) and for all other characteristics available in the data.[15] To facilitate the presentation of the vast set of available variables, Table 4 includes the variables with standardized differences larger than 20 before matching, and Table D.1, those with smaller standardized differences. The information regarding individual characteristics refers to the time of return and registration as a job returner at a local employment agency. The information regarding former employment refers to the last job before interruption. Only the characteristics of the local employment agency districts refer to the (pseudo) treatment time.

The sample means of the characteristics of the treated and non-treated job returners are reported in columns (1) and (2). The training voucher recipients are older and are more likely to be married and to have children younger than six years but are less likely to have children younger than three years. These returners have finished school more often with a university entry degree and are less likely to have no vocational education. The voucher recipients were more often employed in occupations belonging to business accounting and law. The share of re-entrants who had previously worked full-time is higher in the control group, but those in the treatment group were more often employed in a white-collar job. The duration of the employment interruption differs considerably between the treated and non-treated women. The training voucher recipients had an average interruption duration of 72 months (more than 6 years). This duration among the non-treated women is considerably shorter (28 months), which is reflected in their

[15] For each covariate j, I calculate the standardized difference as $SD_j = 100 \times (\bar{x}_{1j} - \bar{x}_{0j}) / \sqrt{0.5(s^2_{x_{1j}} + s^2_{x_{0j}})}$, where $\bar{x}_{1j}$ and $\bar{x}_{0j}$ are the sample means of the j-th covariate in the treated and control group, respectively, and $s^2_{x_{1j}}$ and $s^2_{x_{0j}}$ are the corresponding sample variances.

Table 4: Mean values of observed characteristics with large standardized differences

	(1) Treatment-group	(2) Control-group	(3) SD between (1) and (2)	(4) Matched Control group	(5) SD between (1) and (4)
Personal and family characteristics					
Age in years	36.237	33.900	39.524	35.539	11.898
Age 25-29 years	0.163	0.270	26.178	0.167	1.143
Age 40-44 years	0.241	0.140	26.120	0.188	13.102
Married	0.687	0.571	24.161	0.620	13.991
Children 3–5 yrs	0.256	0.171	20.818	0.302	8.211
Children $\leq$ 3 yrs	0.235	0.483	53.475	0.285	11.416
Educational and vocational degrees					
University entry degree	0.236	0.141	24.410	0.219	3.971
No vocational degree	0.171	0.255	20.526	0.194	6.032
Last occupation					
Business accounting and law	0.402	0.256	31.466	0.351	10.551
Working time: Missing	0.276	0.161	28.139	0.242	7.704
White-collar	0.647	0.447	40.933	0.599	9.960
Blue-collar	0.217	0.359	31.671	0.252	8.071
Interruption and return characteristics					
Interruption duration (months)	72.369	28.432	90.324	54.274	29.949
Eligible for unemployment benefits	0.166	0.409	55.833	0.224	14.660
Status prior interruption					
Employed	0.464	0.271	40.844	0.503	7.771
Unemployed	0.521	0.723	42.565	0.481	8.092
Return 3th quarter 2004	0.110	0.180	20.177	0.150	12.057
Return 4th quarter 2004	0.081	0.196	33.740	0.097	5.767
Labor market history					
Months employed last 1 years	3.510	1.838	39.927	3.374	2.916
Months employed last 3 years	11.801	7.270	39.801	11.317	3.859
Months employed last 6 years	25.672	16.678	41.400	25.080	2.461
Months unemployed last 1 years	2.725	5.349	57.587	3.085	8.219
Months unemployed last 3 years	4.978	11.383	72.003	5.618	7.952
Months unemployed last 6 years	6.362	16.055	75.946	7.310	8.781
Months with benefits last 3 years	5.482	8.767	39.135	5.088	5.230
Months with benefits last 6 years	7.986	13.501	43.853	7.212	7.215
Cumulated wages last 3 years (Euros)	19,329	10,714	40.473	17,575	7.198
Cumulated wages last 6 years (Euros)	46,336	25,831	46.608	44,179	4.225
Cumulated benefits last 3 years (Euros)	3,508	5,961	43.988	3,429	1.523
Cumulated benefits last 6 years (Euros)	5,238	8,743	41.740	4,678	7.630
Regional characteristics					
Bavaria	0.211	0.127	22.482	0.203	1.986
Hamburg, Schleswig-Holstein, Mecklenburg Pomerania	0.042	0.096	21.490	0.051	4.313
Northrhine-Westphalia	0.217	0.136	21.326	0.204	3.388
Saxony-Anhalt, Saxony, Thuringia	0.107	0.267	41.979	0.132	7.709
West Germany	0.781	0.529	55.131	0.754	6.583
Share of employed agriculture	0.011	0.017	50.289	0.012	12.505
Share of employed construction	0.062	0.070	39.929	0.063	4.246
Share of employed trade	0.150	0.144	32.962	0.149	3.955
Share of employed banking, insurance	0.039	0.033	41.665	0.038	7.440
Share of employed public service	0.064	0.071	37.242	0.065	7.421
Share of employed education	0.039	0.047	46.775	0.040	6.032
Share of employed private households	0.001	0.001	50.700	0.001	4.346
Share of vacancies: full-time	0.786	0.751	31.514	0.780	5.814
Unemployment rate	11.792	14.770	52.041	12.160	6.893
N	2,986	3,461		3,461	

Note: In columns (1) and (2), the mean values of observed characteristics for the treated and non-treated sub-samples are reported. Column (4) shows the mean values of the matched control group. Information on individual characteristics refers to the time of registration as job returner in local employment agencies, with the exception of the monthly regional labor market characteristics which refer to the (pseudo) treatment time. I report the standardized differences (SD) between the samples before and after matching multiplied by 100. The variables included in the estimations are listed in Table 3.

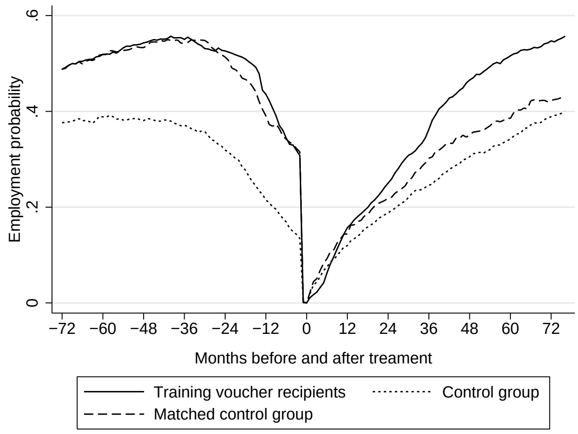
eligibility to receive unemployment benefits. Only 17% of the job returners who received a training voucher are eligible, whereas the share is 41% in the control group. The share of the women who returned in 2003 and the first quarter of 2004 is larger in the treatment group. Regarding the labor market history, I find that the voucher recipients were employed longer, had higher wages and received fewer benefits. These returners are more likely to reside in West Germany and the local employment agency districts with a lower unemployment rate and a higher share of vacancies that are full-time jobs.

In column (3), I report the standardized difference in the mean values between the treated and non-treated job returners before matching. I report the mean values of the covariates for the matched control group in column (4) and the standardized differences after matching in column (5). The standardized differences decrease remarkably after matching in both the included and the omitted variables. According to Austin (2011) there is no universally agreed-upon criterion to indicate important imbalances. However, studies refer to a benchmark of 10 to indicate negligible differences between treatment groups (e.g., Normand et al., 2001). As reported in column (5) of Tables 4 and D.1, the standardized differences after matching are, in most cases, smaller than 10, or, if higher, very close to 10.[16] I interpret this finding as an indicator of high match quality, which is confirmed when I investigate the average employment and earnings levels in the treated, control and matched control groups separately over the pre-interruption and post-treatment time horizons in Figure 2. I find a large difference in the employment and earnings levels of the treated and control samples before the interruption. Among the treated job returners, the employment rates range between 50-60%, and the earnings range between approximately 900-1000 Euros per month. The employment and earnings rates in the control group are lower, confirming the positive selection of the treated samples with respect to employment and benefit histories. After I condition on the variables listed in the third row of Table 3, the differences between the two groups nearly disappear.

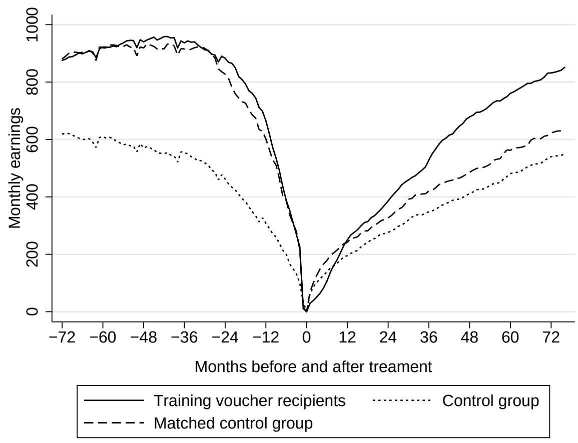
[16]The only exception is the interruption duration in months. For this characteristic, the standardized difference before matching is higher than 90 and reduces remarkably after matching to a value of 30. In order to check if this drives the empirical results, I implemented a subgroup analysis by the length of the interruption duration (interrupted 12-24 month, 24-36 months, or more than 36 months). I find no noteworthy differences between these subgroups. The results are available upon request.

Figure 2: Outcome levels of comparison groups for the pre-interruption and post-treatment period

(a) Employment probability



(b) Monthly earnings (in Euros)



Note: The levels of employment and earnings are presented for each of the 72 months before the interruption and the 76 months following the treatment (148 months).

Overall, these levels appear reasonable because I do not restrict the sample to those who are employed before the interruption. The definition of job returners also includes those who have interrupted unemployment or apprenticeship.

5 Empirical results

5.1 Employment and earnings effects

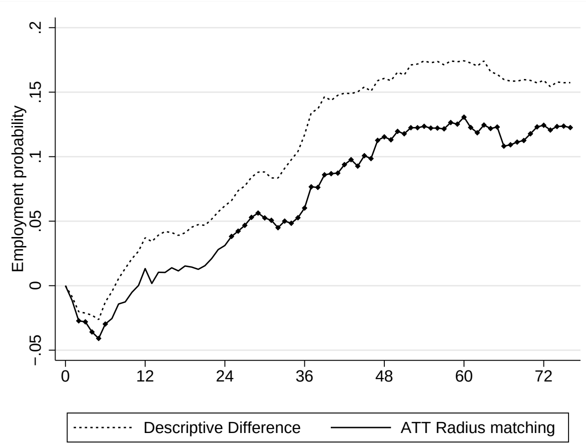
The effects on the probability of being employed and monthly earnings are presented in Figure 3. The lines show the point estimates for each month. The diamonds indicate significant effects at the 5%-level. I compare the women who receive a training voucher to those who do not within a twelve-month period after return and conditional on the (pseudo) start dates of the treatment. The dashed line indicates the unconditional difference between the treatment and control groups. The solid line displays the average treatment effect of the treated (ATT) obtained by radius matching with a bias adjustment for each month after the voucher receipt.

I find negative lock-in effects over the short term because the training participants reduce their search activity during participation, which reduces both the probability of being employed and the monthly earnings compared to the control persons. The length of the lock-in effects corresponds to the average program duration of 12 months. After two years, the point estimates become significantly positive. Four years after the voucher award, the job returners who received a training voucher have a 12 ppoints higher probability of being employed than had they not received the voucher. The employment effect remains constant until the end of the observation period. The long-term effects on monthly earnings amount to 200 Euros (220 USD) per month. These effects are quite large compared to the effect estimates found with the vouchers awarded to regular unemployed persons during the same period (comp. Doerr et al., 2017).

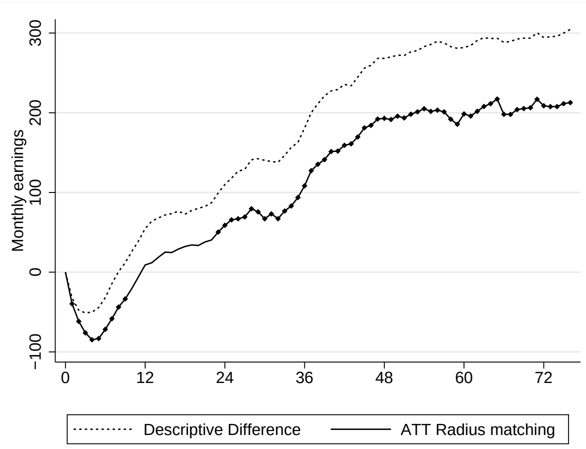
Higher employment and earnings levels may emerge in the long term from the high indirect costs due to short-term foregone employment and earnings during lock-in periods. As shown in Figure 3, the positive effects on both outcomes stabilize after approximately 48 months indicating a positive long-term equilibrium. This implies that the success of vocational training net the lock-in effect will always be positive. To quantitatively assess the net benefit at the end of the observation period, I accumulate the employment months

Figure 3: ATT on employment and earnings

(a) Employment probability



(b) Monthly earnings (in Euros)



Note: Effects are estimated for each of the 76 months following the treatment. Diamonds report significant point estimates at the 5%-level. Significance levels are bootstrapped with 499 replications. Lines without diamonds indicate that point estimates are not significantly different from zero.

and earnings over more than six years. I show the results in Table E.1 in the Appendix. The net benefit starts to be positive and significantly different from zero after four years. Compared to the first month after treatment, the training voucher recipients are employed more than five months longer and earn 8,360 Euro (10,000 USD) more than comparable non-recipients after 76 months. Assuming that the treatment effect remains constant beyond the observational period, the net benefits are likely to be huge at the end of job returners working life.

The results obtained by different evaluation approaches are presented in Figure B.1 in Appendix B. The effect estimates obtained from the static approaches differ only slightly using different time windows and the (pseudo) elapsed duration as a control variable. The lock-in effect on the employment outcomes is somewhat less pronounced when I control for the (pseudo) elapsed duration compared to the pure static approaches. This result might indicate that I correct for the bias predicted by Frederiksson and Johansson (2008) by retaining individuals in the control sample only to the extent that they remain unemployed at the pseudo start of the treatment. In the long term, the pattern of all the static approaches is very similar. The results from the preferred main sample lie in the middle. Using a dynamic approach, I find a different pattern, which hints at attenuation effects. These effects can be rationalized because the voucher recipients in the dynamic concept are compared with a control sample that had partially received a voucher to participate in training later. A detailed discussion can be found in Appendix B.

5.2 Job quality and stability

I aim to analyze if the receipt of a training voucher and potential subsequent participation in a training course affect the quality of the employment obtained by the job returners. I use the probabilities of being employed full-time, earning at least 100% of the previous earnings, and being marginally employed. Marginal employment is also a good measure of employment stability because marginal employment tends to involve short-term contracts and no or reduced contributions to social systems. As a direct measure of job stability, I consider the probability of being employed in a job that lasts for at least 36 months.

I find that the probability of being employed full-time significantly increases among the voucher recipients approximately two years after the treatment (see Table 5). The point estimate after 76 months amounts to a 6.5 ppoint higher probability of full-time employment. Likewise, I find positive effects on the probability to earn at least as much as in the year before the interruption of 4-6 ppoints. The accumulated effects on these job quality measures in Table E.1 show that the voucher recipients are employed full-time more than three months longer than the comparable job returners without training vouchers. Similarly, they are two months longer employed in jobs in which they earn at least their previous wage.

Table 5: ATT on job quality and job stability measures

	6	12	24	36	48	60	72	76
Full-time employment	-0.018	0.011	0.030	0.047	0.074	0.085	0.066	0.065
	(0.009)	(0.012)	(0.012)	(0.015)	(0.016)	(0.017)	(0.017)	(0.018)
At least 100% of previous earnings	-0.008	0.011	0.011	0.018	0.043	0.064	0.052	0.050
	(0.008)	(0.010)	(0.013)	(0.016)	(0.017)	(0.017)	(0.018)	(0.018)
Marginal employment	-0.127	-0.105	-0.068	-0.051	-0.030	-0.047	-0.049	-0.061
	(0.024)	(0.022)	(0.017)	(0.018)	(0.016)	(0.017)	(0.019)	(0.020)
Job lasts at least 36 months	-0.015	0.005	0.025	0.041	0.081	0.101	0.095	0.096
	(0.010)	(0.013)	(0.017)	(0.019)	(0.022)	(0.023)	(0.022)	(0.022)

Note: Significance levels are bootstrapped with 499 replications.

The outcomes of being full-time employed and employed with at least 100% of previous earnings are coded as zero for non-employment and employment of another type. Hence, these measures of employment quality should be interpreted relative to the main employment effect.[17] If the main employment effect is 12 ppoints and the effect on full-time employment is 6.5 ppoints, approximately 55% of the employment gain is in terms of full-time employment, and 45% of the gain is in terms of part-time employment. Accordingly, 40% of the employment gain is from employment relationships in which the job returners earn at least as much as before the interruption.

The positive effects on job quality are also confirmed when we consider the probability of being marginally employed. After a strong lock-in effect, the negative effect estimates on marginal employment stabilize 36 month after treatment at minus 4-6 ppoints. Thus, the probability of marginal employment in permanently reduced for job returners who received a training voucher. The cumulative effects show that the job returners with training vouchers were marginally employed for five fewer months than they would have been without a training voucher (see Table E.1). As outlined in Section 3.2, marginal employment can be characterized as precarious employment, especially if it is the only job and not a second one in addition to regular employment. In the unreported estimations, I check whether the effect is driven by a reduction in marginal employment as the main job or a reduction in the probability of having a marginal side job. I find that awarding

[17] The reason for this is to avoid a double selection problem (e.g., Heckman, 1979, Lee, 1978, Sorensen, 1989). In a scenario in which I would interpret the employment quality measures conditional on finding employment, I would have to control for two selection processes: into treatment and into employment.

training vouchers to female job returners has no effect on the probability that they engage in marginal employment as a side job. The entire effect is driven by a reduction in marginal employment as the main job.[18]

Finally, I present the job stability results in the last row of Table 5. The job returners with training vouchers benefit from a significantly higher probability of being employed for at least 36 months by almost 10 ppoints. This effect is remarkable if interpreted relative to the main employment effect, which suggest that at the end of the observation period nearly 80% of the employment gain is from stable employment relationships that last at least three years. Furthermore, the voucher recipients were employed almost four months longer in such jobs (see Table E.1).

5.3 Robustness checks

I seek to interpret the change in the employment outcomes of female job returners as a causal effect of receiving a training voucher. I perform several checks to support the causal interpretation and robustness of the results. First, I implement a pre-interruption outcome evaluation. Figure 2 illustrates that the outcomes of the treated and matched control groups already become similar when I condition on the observed characteristics. If time-constant unobserved characteristics lead to a selection bias that is not controlled for by the included variables, it should be apparent in an impact evaluation prior to the interruption. Therefore, I use the same matched sample and empirical specification as in the main analysis in the post-treatment period and estimate the treatment effects on various points during the pre-interruption period.[19] The results are shown in Table F.1. The effect estimates are nearly never significantly different from zero.

One concern with this test is that the pre-interruption outcomes are related to the control variables that are used to balance the sample, particularly the employment history controls. Although the outcomes are measured on a monthly basis, and I use aggregated

[18]The results are available upon request.

[19]I use the employment probability, monthly earnings, and probability of being marginally employed in this robustness test. Since working full-time, earning at least 100% of previous earnings, and having a stable job are categories of the main employment variable, I do not use them in the pre-interruption outcome evaluation.

measures to control for the employment history, this procedure is more a test of internal validity than of external validity. Therefore, I implement a second test for external validity. I use the same sample of treated and control persons as for the main analysis. Then, I define a placebo treatment start at exactly 12 months prior to the date when the interruption occurred and redefine the control variables accordingly. This method is possible for the large majority of controls but not for those capturing the interruption characteristics, e.g., the lengths of the interruption period. The results for this test are presented in Table F.2. None of the effects estimates are significantly different from zero. Both tests together support a causal interpretation of the effects.

In the third test, I use the unredeemed training vouchers for a specification test. The purpose is similar to that of the test presented above, but I focus on the post-treatment period instead of the pre-interruption period. This approach allows me to rule out selection effects that may occur shortly before the treatment (e.g., anticipation effects). If such effects exist, they should also be visible in the effect estimates of the job returners who did not redeem their training vouchers. The corresponding results are shown in Figure F.1. Because the number of observations is markedly reduced (90% of all job returners redeem their training vouchers), I present additional effect estimates obtained by a linear regression. Both estimators imply that the treatment effects of the unredeemed vouchers are almost never significantly different from zero. Again, I interpret this result as a confirmation of the robustness of the presented results.

5.4 Heterogeneous effects by vocational degree

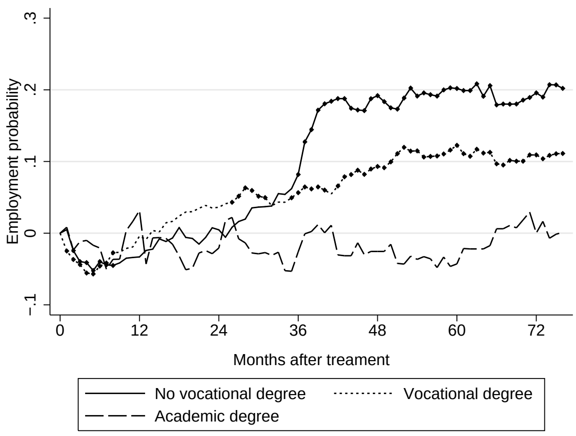
In this type of study, it is interesting to explore the effect heterogeneity in greater depth. First, I focus on differences among different qualified job returners. Second, I show how the effects vary by program type (see Section 5.5).[20] Previous studies that focus on training for regular unemployed persons find the largest effects for the lower skilled participants

[20] I have implemented a heterogeneity analysis by the employment state of job returners directly before the interruption (employed, in apprenticeship, or unemployed) and by the length of the interruption duration (interrupted 12-24 month, 24-36 months, or more than 36 months). Since I find no noteworthy differences and the effects for all subgroups follow the pattern of the main results, I decided to not include this analysis in the paper. Results are available upon request.

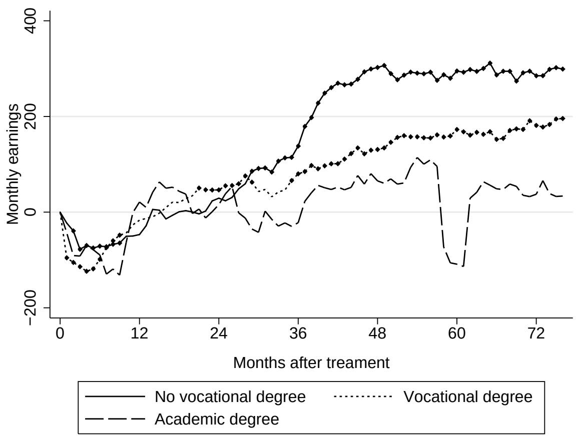
(e.g., Doerr et al., 2017, Lechner et al., 2011). Since training programs provide occupation-specific skills and sometimes even a vocational degree to participants, training vouchers could be especially beneficial for low-skilled job returners. Qualifications are measured by vocational degrees at return. I distinguish between job returners with and without vocational degrees and identify high-skilled job returners with academic degrees. I stratify the sample according to these dimensions and perform the estimations separately for each stratum. The heterogeneous effects by vocational qualification are presented in Figure 4.

Figure 4: ATT on employment and earnings by vocational degree

(a) Employment probability



(b) Monthly earnings (in Euros)

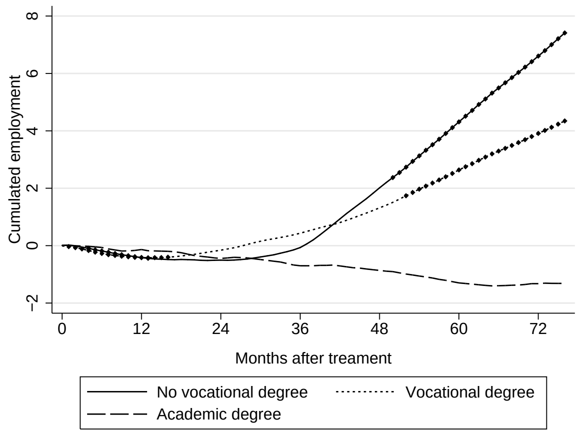


Note: Effects are estimated for each of the 76 months following the treatment. Diamonds report significant point estimates at the 5%-level. Significance levels are bootstrapped with 499 replications. Lines without diamonds indicate that point estimates are not significantly different from zero.

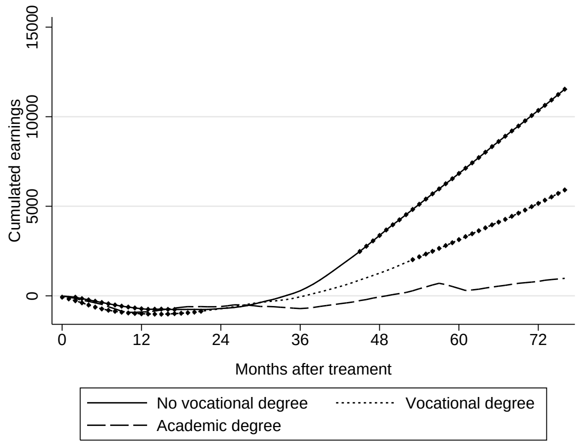
The highest effect is found among the lowest-skilled job returners. I find some negative short-term effects, but after 36 months, the effect estimates are significantly positive on

Figure 5: Cumulative employment and earnings by vocational degree.

(a) Accumulated employment



(b) Accumulated earnings (in Euros)



Note: Accumulated effects are presented for each of the 76 months following the treatment. Diamonds report significant point estimates at the 5%-level. Significance levels are bootstrapped with 499 replications. Lines without diamonds indicate that point estimates are not significantly different from zero.

a high and stable level of 20 ppoints for employment. The pattern of monthly earnings is very similar. After 36 months, the effects amount to almost 300 Euros (360 USD) and stabilizes until the end of the observation period. The pattern and effect size of the estimates for the medium-skilled job returners are very similar to the main effect. The medium-skilled job returners represent the largest skill group with 75% in the treated and 70% in the control sample. After two years, the effects become significantly positive, and increase to a 12 ppoint higher employment probability in the long term. Earnings increase by nearly 210 Euros (250 USD). The accumulated effects show significant employment and earnings gains in both groups (see Figure 5). The net gains from being awarded a training voucher are higher for low-skilled job returners without vocational degrees than for the medium-skilled, although the cumulated effects for the low-skilled remain negative over a longer period. Over the 76 months considered in this study, treated job returners without vocational degree were more than seven months longer employed and earned 11,000 Euros more (13,000 USD) than had they not received the voucher.

The results of the highest-skilled job returners with academic degrees paint a more negative picture. I find no effects of a voucher award over the entire observation period on either employment or monthly earnings. Thus, awarding training vouchers to high-skilled job returners with academic degrees is ineffective. The accumulated effects on employment indicate employment losses, but these effects are not significant. The results of the specification test of the pre-interruption outcomes are provided in Table G.1 in Appendix G. The effects on the pre-interruption employment probabilities and earnings are not significant in the large majority of cases. Again, this result supports a causal interpretation for the heterogeneous effects by vocational degree.

Overall, the finding that the low-skilled job returners profit most from the voucher award confirms the results by Doerr et al. (2017). However, there are interesting differences between the results for regular unemployed and job returners. Doerr et al. (2017) find very strong and pronounced lock-in effects in all skill groups that are significantly more negative the higher the qualification of the unemployed is. This is not the case for the job returners. Their lock-in effects vary somewhat in lengths, which comes from a different course type composition in the skill groups, but not in their depth, because job returners of all skill levels have a similar chance to find employment in the counterfactual situation of not being awarded a training voucher. This implies that even if a voucher award is not effective in the long term, as it is the case for the high-skilled job returners, the net loss is not huge. Another difference to the findings in Doerr et al. (2017) is the magnitude of effects. The long-term effect of receiving a training voucher is zero for high-skilled regular unemployed and job returners alike, but the effects for the medium- and low-skilled are much higher for job returners compared to regular unemployed.

The heterogeneity by skill level and the remarkable positive effects for low-skilled job returners could simply result from the composition of training types within the skill groups. In fact, almost 50% of job returners without vocational degree participate in retraining programs whereas this share is only 20% and 14% among the medium- and high-skilled, respectively. In Table G.3 in the Appendix, I present a subgroup analysis by skill level and training program type. I find that the low- and medium-skilled job returners similarly benefit when they participate in short and long training programs. Focusing on retraining programs, I again find that the low-skilled job returners profit most. They outperform all other skill groups in terms of employment and earnings effects. Thus, the high positive effects for low-skilled job returners not only come from the course composition, but also because they seem to benefit most from training holding the program constant. Furthermore, this analysis reveals that awarding a training voucher to high-skilled job returners is not effective irrespective of the program type.

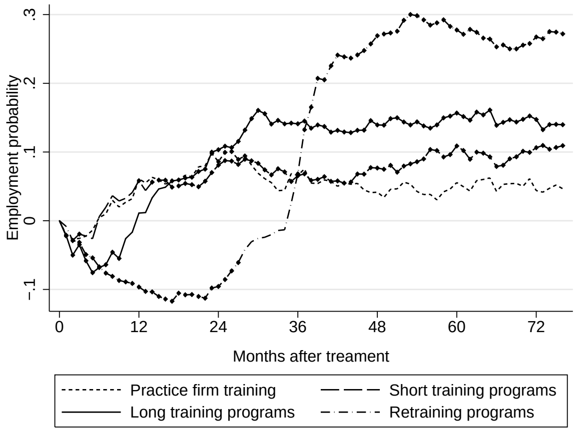
5.5 Heterogeneous effects by training type

In this section, I estimate the effectiveness of the different program types conditional on voucher redemption (90% of all the job returners redeem their training vouchers). The results by program type cannot be interpreted in a causal way because the redemption of a voucher may also include a particular selection. The results are nevertheless suggestive because they indicate the channel through which the causal effect operates.

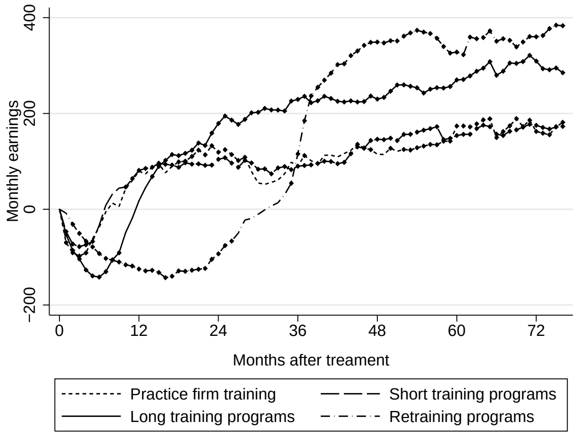
The effect estimates clearly show that the success of the different program types increases with the human capital provided in these courses (comp. Figure 6). Furthermore, the pattern of the lock-in effects clearly corresponds to the length of the different program types. I find short negative lock-in effects for participants in practice firm training and short training over a period of six months. One year after treatment the effects on employment and monthly earnings turn significantly positive. Whereas the earning effects evolve very similar for the two programs and result in a monthly increase of 180 Euros (220 USD) in the long term, the effects on employment diverge after 3.5 years. The long term employment effects of practice firm training are not significant. For short training

Figure 6: ATT on employment and earnings by course type

(a) Employment probability



(b) Monthly earnings (in Euros)



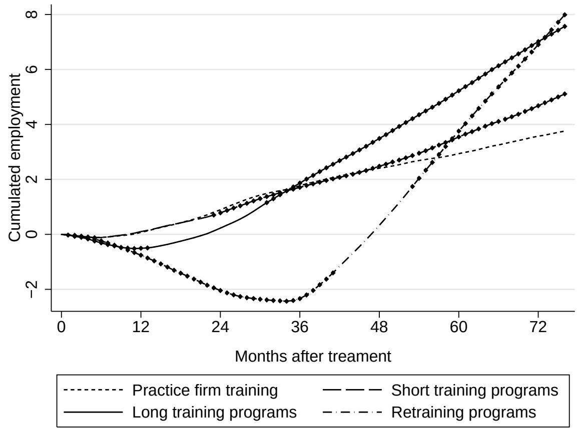
Note: Effects are estimated for each of the 76 months following the treatment. Diamonds report significant point estimates at the 5 percent level. Significance levels are bootstrapped with 499 replications. Lines without diamonds indicate that point estimates are not significantly different from zero.

courses, I find a significant employment increase that stabilizes around 10 ppoints four years after treatment until the end of the observation period.

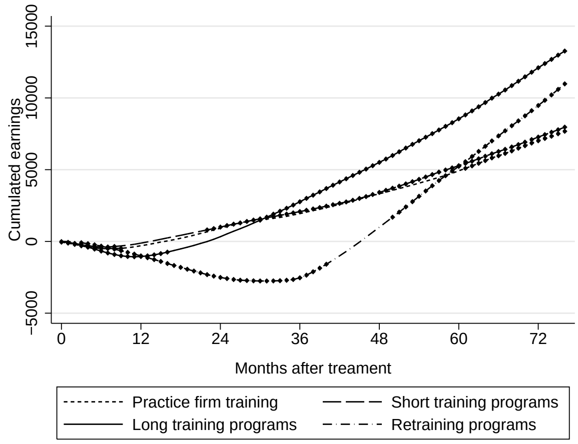
The long training programs provide occupation-specific knowledge and have a relatively long average duration of nine months. After a pronounced lock-in period, the employment and earnings effects become significantly positive 1.5 years after treatment. The effects on employment probability increase after two years to constant and significant 14 ppoints until the end of the observation period. The earnings effects show a similar pattern, and increase from 220 Euros (260 USD) over the medium-term to 280 Euros (340 USD) six years after the training voucher is awarded.

Figure 7: Cumulative employment and earnings by course type.

(a) Accumulated employment



(b) Accumulated earnings (in Euros)



Note: Accumulated effects are presented for each of the 76 months following the treatment. Diamonds report significant point estimates at the 5%-level. Significance levels are bootstrapped with 499 replications. Lines without diamonds indicate that point estimates are not significantly different from zero.

The retraining programs have the longest durations (up to three years). Participating in such programs provide job returners the possibility to obtain a vocational degree. Therefore, unsurprisingly, I find much longer lock-in effects with this program type. The short-term employment probability is reduced by 11 ppoints, and earnings are decreased by 160 Euros (190 USD). After 36 months, the employment gain jumps rapidly, to over 25 ppoints, and the earnings gains are large, at over 370 Euros (450 USD) per month.

I present the results of cumulative employment and earnings in Figure 7. I do not find any effect on cumulative employment among the job returners who participated in practice firm training, but a significant gain in cumulative employment of five months with short training programs. The cumulative earning gains are similar for practice training and short programs and amount to 7,000 Euros (8,400 USD). Participants in long training programs and retraining gain almost 8 months of employment until the end of the observation period. The slope of the cumulated effect is steeper for the retraining programs, thus, in the longer term beyond the 76 months of observation, I expect the cumulative effects of retraining to become much larger than those of long training programs. Regarding cumulative earnings, I find earning gains of more than 13,000 Euros (15,500 USD) with the long training programs and gains of 11,000 Euros (13,000 USD) with the retraining programs. The participants in these programs benefit from higher employment and high monetary returns from their investment in human capital.

6 Conclusion

This paper is the first to investigate the impact of publicly funded vocational training for female job returners on their subsequent working careers. It focuses on the case of Germany, where re-entrants to the labor market after a family-related interruption are defined as individuals eligible to participate in ALMP programs if they register at local employment agencies. Vocational training is the most prominent qualification instrument among ALMP programs available and assigned through a training voucher. Using the rich administrative data of a 100% sample of job returners who are awarded a training voucher from 2003-2005 and a randomly generated control group, I identify the treatment effect of a training voucher award for the female job returners.

The findings show that training programs are very effective in improving the labor market outcomes of women after family-related withdrawal from the labor force. The award of a training voucher translates into a higher employment probability of 12 ppoints and higher monthly earnings (200 Euros/220 USD). Importantly, training vouchers also improve the quality and stability of employment of the job returners. They face a higher probability of earning at least as much than before the interruption and obtaining a full-time job. In contrast, the probability of acquiring unstable marginal employment permanently decreases. Nearly 80% of the employment gain is from stable employment relationships. Several robustness checks support a causal interpretation of these results.

The investigation of the effect heterogeneity reveals some interesting insight into vocational degrees and the effectiveness of different course types. Training vouchers are an effective instrument to improve the labor market outcomes for low- and medium-skilled re-entrants, and the highest positive effects are found for low-qualified job returners without a vocational degree. In contrast, awarding a voucher to job returners with academic degree is ineffective. I find no effect of unredeemed training vouchers. Conditional on the redemption of the voucher, the effectiveness of training increases with the human capital intensity provided by the respective course. This results strongly imply that the positive effects of training vouchers work through the human capital channel.

The analysis in the paper exclusively focuses on the effect of publicly funded vocational training for the job returners who decided to register in local employment offices to receive consultation and support in finding a job. On the one hand, based on the observed socio-economic characteristics of the registered job returners, those who received treatment seem to be positively selected compared to the non-treated women. Thus, displacement effects could be an issue if some of the gains of the job returners who received a training voucher came at the expense of displacing the not treated or not-registered job returners. On the other hand, the treated job returners experienced longer interruption durations on average. When job returners with long interruptions are discriminated by employers, displacement effects might be less of an issue. Unfortunately and similar to most program evaluation studies that use individual data at the micro level, it is not possible to analyze such general equilibrium effects.

This study shows that vocational training helps the long term labor market outcomes of female job returners that received a training voucher. In particular, awarding a training voucher to low- and medium-skilled job returners contributes to the successful reintegration of these women. They benefit from a significant higher employment probability, higher earnings and higher job quality and stability, especially if they participate in programs that provide a substantial amount of human capital. However, for high-skilled women with an academic degree other policy instruments should be considered.

References

Andersson, F., H. J. Holzer, J. I. Lane, D. Rosenblum, and J. Smith (2013). Does federally-funded job training work? nonexperimental estimates of wia training impacts using longitudinal data on workers and firms. NBER Working Papers 19446, National Bureau of Economic Research.

Angelov, N., P. Johansson, and E. Lindahl (2016). Parenthood and the gender gap in pay. *Journal of Labour Economics 34*(545–579), 399–424.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research 46*, 399–424.

Baker, M., J. Gruber, and K. Milligan (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy 116*(4), 709–745.

Beblo, M., S. Bender, and E. Wolf (2009). Establishment-level wage effects of entering motherhood. *Oxford Economic Papers 61*, i11–i34.

Bergemann, A. and G. van den Berg (2008). Active labor market policy effects for women in europe: A survey. *Annales d'Economie et de Statistique 91/92*, 377 –399.

Bergemann, A. and G. van den Berg (2014). From giving birth to paid labor: The effects of adult education for prime-aged mothers. *IFAU Working Paper 15/2014*.

Biewen, M., B. Fitzenberger, A. Osikominu, and M. Paul (2014). The effectiveness of public sponsored training revisited: The importance of data and methodological choices. *Journal of Labor Economics 32(4)*, 837–897.

Calvo-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. *American Economic Review 94*, 426–454.

Crépon, B., M. Ferracci, and G. J. Jolivet, Grégory van den Berg (2009). Active labor market policy effects in a dynamic setting. *Journal of the European Economic Association 7*, 595–605.

Dahl, G. B., K. V. Løken, M. Mogstad, and K. V. Salvanes (2016). What is the case for paid maternity leave? *The Review of Economics and Statistics 98*(4), 655–670.

Dehejia, R. H. and S. Wahba (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association 94*(448), 1053–1062.

Dehejia, R. H. and S. Wahba (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics 84*(1), 151–161.

Diener, K., S. Götz, F. Schreyer, and G. Stephan (2013). Beruflicher Wiedereinstieg von Frauen nach familienbedingter Erwerbsunterbrechung. *IAB-Forschungsbericht 9*.

Doerr, A., B. Fitzenberger, T. Kruppe, M. Paul, and A. Strittmatter (2017). Employment and earnings effects of awarding training vouchers in germany. *Industrial and Labor Relations Review 70*(3), 767–812.

Drange, N. and M. Rege (2013). Trapped at home: The effect of mothers' temporary labor market exits on their subsequent work career. *Labour Economics 24*, 125–136.

Ebach, M. and B. Franzke (2014). *Nichtleistungsberechtigte Wiedereinsteigerinnen in Westdeutschland und die Arbeitsförderung nach SGB III.* Bertelsmann.

Evertsson, M. (2016). Parental leave and careers: Women's and men's wages after parental leave in sweden. *Advances in Life Course Research 29*, 26–40.

Evertsson, M. and D. Grunow (2012). Women's work interruptions and career prospects in germany and sweden. *International Journal of Sociology and Social Policy 32*(9/10), 561–575.

Evertsson, M., D. Grunow, and S. Aisenbrey (2016). Work interruptions and young women's career prospects in germany, sweden and the us. *Work, Employment and Society 30*(2), 291–308.

Frederiksson, P. and P. Johansson (2008). Dynamic treatment assignment - the consequences for evaluations using observational studies. *Journal of Business Economics and Statistics 26*(4), 435–445.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica 47*, 153–161.

Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999). The Economics and Econometrics of Active Labor Market Programs. In O. C. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, Vol. 3*, pp. 1865–2086. North Holland.

Heinrich, C. J., P. R. Mueser, K. R. Troske, K.-S. Jeon, and D. C. Kahvecioglu (2013). Do public employment and training programs work? *IZA Journal of Labor Economics*.

Huber, M., M. Lechner, and A. Steinmayr (2014). Radius Matching on the Propensity Score with Bias Adjustment: Finite Sample Behaviour, Tuning Parameters and Software Implementation. *Empirical Economics 49*(1), 1–31.

Huber, M., M. Lechner, and C. Wunsch (2013). The Performance of Estimators Based on the Propensity Score. *Journal of Econometrics 175*(1), 1–21.

Huber, M., M. Lechner, and C. Wunsch (2015). Work health promotion and labour market performance of employees. *Journal of Health Economics 43*, 170–189.

Imbens, G. W. (2004). Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *Review of Economics and Statistics 86(1)*, 4–29.

Kluve, J. and S. Schmitz (2018). Back to work: Parental benefits and mothers' labor market outcomes in the medium run. *Industrial and Labor Relations Review 71*(1), 143–173.

Lalive, R., A. Schlosser, A. Steinhauer, and J. Zweimüller (2014). Parental leave and mothersćareers: The relative importance of job protection and cash benefits. *The Review of Economic Studies 81*(1), 219–265.

Lalive, R. and J. Zweimüller (2009). How does parental leave affect fertility and return to work? evidence from two natural experiments. *The Quarterly Journal of Economics 124*(3), 1363–1402.

Lechner, M., R. Miquel, and C. Wunsch (2011). Long-run effects of public sector sponsored training. *Journal of the European Economic Association 9(4)*, 742–784.

Lechner, M. and J. Smith (2007). What is the Value Added by Caseworkers? *Labour Economics 14(2)*, 135–151.

Lechner, M. and A. Strittmatter (2019). Practical procedures to deal with common support problems in matching estimation. *Econometric Review 38*(2), 193 – 207.

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

Lee, L. (1978). Unionism and wage rates: a simultaneous equations model with qualitative and limited dependent variables. *International Economic Review 19*, 415–433.

McCall, B., J. A. Smith, and C. Wunsch (2016). Chapter 9 - government-sponsored vocational education for adults. Volume 5 of *Handbook of the Economics of Education*, pp. 479 – 652. Elsevier.

Mincer, J. and S. Polachek (1974). Family investments in human capital: Earnings of women. In *Marriage, Family, Human Capital, and Fertility*, pp. 76–110. National Bureau of Economic Research, Inc.

Mueser, P., K. Troske, and A. Gorislavsky (2007). Using state administrative data to measure program performance. *Review of Economics and Statistics 89*(4), 761–783.

Normand, S. L. T., M. B. Landrum, E. Guadagnoli, J. Z. Ayanian, T. J. Ryan, P. D. Cleary, and B. J. McNeil (2001). Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: A matched analysis using propensity scores. *Journal of Clinical Epidemiology 54*(4), 387–398.

Puhani, P. A. and K. Sonderhof (2011). The effects of parental leave extension on training for young women. *Journal of Population Economics 24*, 731–760.

Rinne, U., A. Uhlendorff, and Z. Zhao (2013). Vouchers and caseworkers in public training programs: Evidence from the hartz reform in germany. *Empirical Economics 45(3)*, 1089–1127.

Rossin-Slater, M., C. J. Ruhm, and J. Waldfogel (2013). The effects of californias paid family leave program on mothers leave-taking and subsequent labor market outcomes. *Journal of Policy Analysis and Management 32*(2), 224–245.

Rubin, D. (1974). Estimating the Causal Effect of Treatments in Randomized and Non-Randomized Studies. *Journal of Educational Psychology 66*(5), 688–701.

Schönberg, U. and J. Ludsteck (2014). Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth. *Journal of Labor Economics 32*, 469–505.

Schrenker, A. and A. Zucco (2020). Gender pay gap steigt ab dem alter von 30 jahren stark an. *DIW-Wochenbericht 10*, 137–145.

Sianesi, B. (2004). An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s. *Review of Economics and Statistics 86*, 133–155.

Sorensen, E. (1989). Measuring the pay disparity between typically female occupations and other jobs: A bivariate selectivity approach. *Industrial and labor relations review 42*(4), 624–639.

van den Berg, G. J. and J. Vikström (2019). Long-Run Effects of Dynamically Assigned Treatments: A New Methodology and an Evaluation of Training Effects on Earnings. *IZA Discussion Paper No. 12470*.

Vikström, J. (2017). IPW estimation and related estimators for evaluation of active labor market policies in a dynamic setting. *Labour Economics 49*, 42–54.

A Institutional details

Table A.1: Services available in each registration state

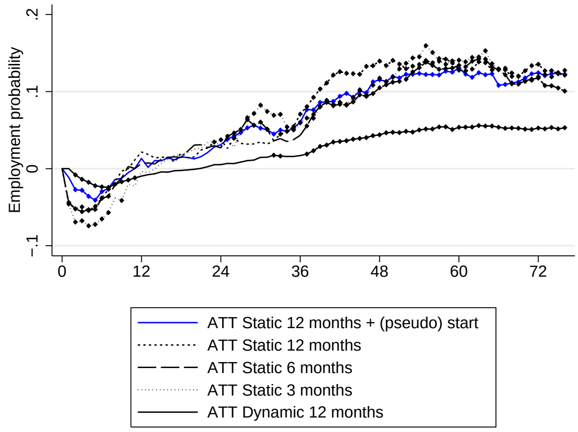
Services	Seeking advice	Search for employment	Unemployed
Information services Material Events KURSNET, BERUFENET	✓	✓	✓
Counseling services Consultation (§29-30 SGB III)	✓	✓	✓
Placement services Intensive placement counseling Placement offers Online job market (§35 SGB III)		✓	✓
Promotion services Placement promotion (§44 SGB III) Activation and integration (§45 SGB III) Training voucher (§81-87 SGB III) Targeted wage subsidy (§88 SGB III)			✓

B Alternative evaluation and control samples

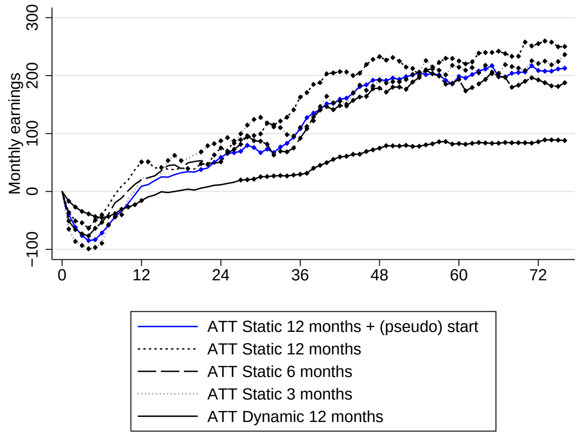
The results obtained by different evaluation samples are presented in Figure B.1. There are only small differences between the effect estimates obtained from static approaches using different time windows and the (pseudo) elapsed duration as an additional control variable (indicated by the blue line). I find very different effects when applying a dynamic evaluation framework. The results from that approach are indicated by the solid black line. Frederiksson and Johansson (2008) argue that estimates from a static approach might be underestimated due to a positively selected control group (see discussion in Section 4.1). In this case, the effects from a dynamic approach would be more positive compared to those obtained from a static approach. That means I would find less negative effects during lock-in periods and higher effects in the longer run. When I control for the (pseudo) elapsed duration in the static approach I find that the lock-in effects on employment are

Figure B.1: ATT using different evaluation samples

(a) Employment probability



(b) Monthly earnings (in Euros)



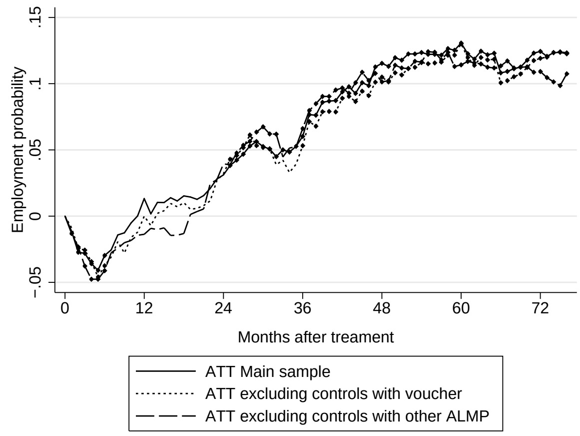
Note: Effects are estimated for each of the 76 months following the treatment. Diamonds report significant point estimates at the 5%-level. Significance levels are bootstrapped with 499 replications. Lines without diamonds indicate that point estimates are not significantly different from zero.

somewhat less pronounced. This might indicate that I correct for the bias predicted by Frederiksson and Johansson (2008) by retaining individuals in the control sample only to the extent that they are still unemployed at the pseudo start of the treatment. However, in the long run the patterns of all static approaches are very similar.

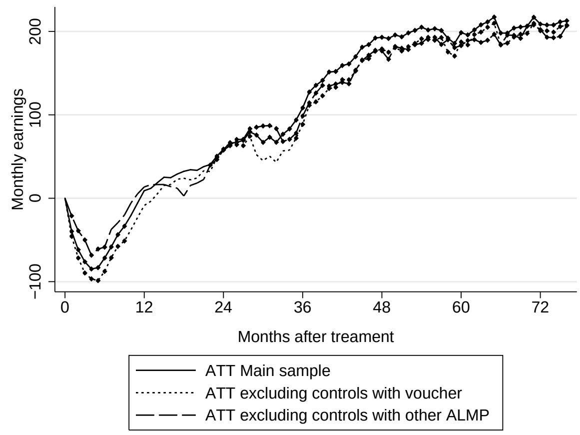
The results from the dynamic approach do not support the predicted pattern in case of a positively selected control group. Instead, I find evidence of an attenuation effect. Using the dynamic concept, I estimate the effect of receiving a training voucher now vs. receiving it later or never for each months over the treatment period. Thus, a fraction of control persons each month m will likely receive a training voucher in a following month $m+t$. The short-term effect estimates (lock-in effects) are closer to zero because I compare voucher recipients during their potential subsequent training participation with control persons who partly participate in training shortly thereafter. The lower effect estimates over the long term can be rationalized by the same argument. Voucher recipients are compared with control persons that had partially participated in training and thus attenuate the estimated effects. The results from using alternative control group definition are presented in Figure B.2.

Figure B.2: ATT using different control group definitions

(a) Employment probability



(b) Monthly earnings (in Euros)



Note: Effects are estimated for each of the 76 months following the treatment. Diamonds report significant point estimates at the 5%-level. Significance levels are bootstrapped with 499 replications. Lines without diamonds indicate that point estimates are not significantly different from zero.

C Details on the identification

The treatment of interest is a voucher awarded during the first twelve months after registration as a job returner between January 2003 and December 2005. Each woman is observed for at least 76 months. The voucher award as an intention to treat is denoted as $D_i \in 0, 1$ (for individuals $i = 1, ..., N$). The outcome variable is denoted as Y_{it} (where $t = 1, ..., 76$) and indicates the number of months since the award of the voucher. To account for the dynamics of the assignment process, I match on the time elapsed between return and treatment in the main evaluation.

Following the notation described in Rubin (1974), the potential outcomes are indicated as Y_{it}^d, where $d = 1$ under treatment and $d = 0$ otherwise. For each job returner, only the realized outcome is observed. Thus, the potential outcome $E[Y_{it}^1|D_i = 1]$ is directly observed from the data. $E[Y_{it}^0|D_i = 1]$ is the counterfactual expected potential outcome as Y_{it}^0 is never observed in the treated sub-population and is the expected non-treatment outcome of those who received a training voucher. I aim to identify the expected difference γ_t between the outcomes Y_{it}^1 and Y_{it}^0 of the women who received a training voucher in each month of the post-treatment period. This effect is known as the average treatment effect of the treated (ATT) in the literature (e.g., Imbens 2004)

$$\gamma_t = E[Y_{it}^1|D_i = 1] - E[Y_{it}^0|D_i = 1].$$

Assumption 1 *(Conditional Mean Independence).* For all $d \in \{0, 1\}$,

$$E[Y_{it}^d|D_i = 0, X_i = x] = E[Y_{it}^d|D_i = 1, X_i = x] \text{ for } \forall x \in \mathbb{X},$$

and $t \in 1, ..., 76$, where $\mathbb{X}$ denotes the support of X_i, and all necessary moments exist. The (pseudo) start dates M of the treatment are included in the vector of control variables ($M \in X$). When the CIA holds, conditional on the pretreatment control variables X_i, individuals are randomly assigned to the treatment, and the expected potential outcomes are independent of the treatment status D_i.

Assumption 2 *(Common Support).*

$$0 < p(x) < 1, \text{where } \ p(x) = Pr(D_i = 1|X_i = x)$$

where $p(x)$ is the conditional treatment probability (propensity score). I enforce Assumption 2 by excluding all observations outside of the common support from the estimations (only 0.03%). This approach does not affect the results.

D Descriptive statistics

Table D.1: Means of observed characteristics with small standardized differences

	(1) Treatment-group	(2) Control-group	(3) SD between (1) and (2)	(4) Matched Control group	(5) SD between (1) and (4)
Personal and family characteristics					
Age 30-34 years	0.226	0.300	16.945	0.270	10.163
Age 35-39 years	0.287	0.241	10.537	0.309	4.733
Age 45-49 years	0.082	0.049	13.369	0.066	6.117
Health problems	0.042	0.088	18.921	0.039	1.528
Not German	0.055	0.070	6.369	0.070	6.571
Children	0.857	0.913	16.793	0.903	13.432
Number of children	1.655	1.694	3.901	1.729	7.390
Educational and vocational degrees					
No educational degree	0.017	0.054	19.957	0.030	8.022
Schooling degree without Abitur	0.743	0.803	14.377	0.749	1.332
Missing	0.004	0.001	4.430	0.002	2.352
Vocational degree	0.751	0.700	11.413	0.722	6.520
Academic degree	0.078	0.046	13.478	0.084	2.105
Last occupation					
Full-time work	0.369	0.457	17.786	0.367	0.393
Part-time work	0.355	0.383	5.772	0.391	7.370
Worker type: missing	0.136	0.194	15.672	0.150	4.018
Agriculture, forestry and farming	0.019	0.046	15.163	0.019	0.036
Production and manufacturing	0.084	0.108	8.139	0.078	2.475
Construction and architecture	0.014	0.022	5.538	0.023	6.319
Science and information services	0.021	0.018	2.593	0.015	4.595
Traffic and logistics	0.110	0.143	9.971	0.143	9.771
Commercial service, trading and sales	0.160	0.205	11.680	0.176	4.220
Health, social sector and teaching	0.133	0.139	1.612	0.138	1.348
Humanities and social science	0.017	0.016	0.989	0.016	0.829
Missing	0.038	0.046	4.417	0.041	1.887

< table continues on next page >

Table D.1: < continued >

	Treatment-group	Control-group	SD between (1) and (2)	Matched Control group	SD between (1) and (4)
Interruption and return characteristics					
Status prior interruption					
Apprenticeship	0.015	0.006	8.845	0.016	1.310
Elapsed duration until treatment	3.961	3.678	8.874	3.938	0.711
Return 1st quarter 2003	0.121	0.066	19.137	0.114	2.094
Return 2nd quarter 2003	0.110	0.066	15.617	0.078	11.048
Return 3rd quarter 2003	0.148	0.096	15.800	0.133	4.311
Return 4th quarter 2003	0.146	0.105	12.417	0.125	6.338
Return 1st quarter 2004	0.166	0.147	5.368	0.168	0.541
Return 2nd quarter 2004	0.118	0.144	7.712	0.135	5.009
Registration: unemployed	0.680	0.635	9.458	0.646	7.189
Registration: searching	0.320	0.364	9.237	0.354	7.144
Labor market history					
Months with benefits last 1 years	3.253	4.082	18.041	2.538	8.433
Regional characteristics					
Baden-Württemberg	0.097	0.083	5.098	0.105	2.491
Berlin, Brandenburg	0.085	0.124	12.909	0.078	2.451
Lower Saxony, Bremen	0.118	0.087	10.238	0.126	2.457
Hesse	0.072	0.045	11.547	0.064	3.067
Rhineland Palatinate, Saarland	0.051	0.034	8.305	0.038	6.511
Share of employed mining	0.005	0.005	0.338	0.004	6.603
Share of employed production	0.249	0.233	18.757	0.256	2.974
Share of employed energy and water	0.009	0.010	14.519	0.009	3.218
Share of employed trade	0.151	0.148	17.546	0.152	5.873
Share of employed tourisms	0.029	0.029	2.788	0.028	0.910
Share of employed communication	0.055	0.058	18.577	0.055	3.550
Share of employed real estate	0.120	0.115	13.104	0.116	9.315
Share of employed health and social	0.117	0.117	0.742	0.118	1.143
Share of employed services	0.047	0.048	10.226	0.047	1.681
Share of employed ext. organisation	0.001	0.001	9.163	0.001	1.958
Population density (per km^2)	722	569	10.830	574	11.084

Note: In columns (1) and (2), the mean values of observed characteristics for the treated and non-treated sub-samples are reported. Column (4) shows the mean values of the matched control group. Information on individual characteristics refers to the time of registration as job returner in local employment agencies, with the exception of the monthly regional labor market characteristics which refer to the (pseudo) treatment time. I report the standardized differences (SD) between the different samples before and after matching multiplied by 100. The variables included in the estimations are listed in the third row of Table 3.

E Cumulated effects

Table E.1: Cumulated months of employment and earnings (all outcomes)

	6	12	24	36	48	60	72	76
Months employed	-0.17	-0.21	-0.03	0.57	1.69	3.16	4.58	5.07
	(0.49)	(0.12)	(0.28)	(0.45)	(0.64)	(0.84)	(1.04)	(1.10)
Monthly earnings in Euro	-437	-589	-193	734	2,681	5,049	7,521	8,360
	(101)	(178)	(320)	(552)	(843)	(1,202)	(1,630)	(1,787)
Months full-time employed	-0.09	-0.10	0.05	0.52	1.30	2.21	3.11	3.38
	(0.03)	(0.08)	(0.18)	(0.30)	(0.43)	(0.58)	(0.72)	(0.76)
Months at least 100% previous earnings	-0.05	-0.02	0.04	0.29	0.70	1.34	1.95	2.16
	(0.03)	(0.07)	(0.19)	(0.31)	(0.46)	(0.63)	(0.80)	(0.86)
Months marginally employed	-0.80	-1.50	-2.45	-3.14	-3.57	-4.14	-4.67	-4.89
	(0.15)	(0.27)	(0.46)	(0.59)	(0.71)	(0.84)	(0.98)	(1.04)
Months in stable employment	-0.08	-0.12	-0.01	0.37	1.19	2.29	3.47	3.85
	(0.03)	(0.10)	(0.27)	(0.44)	(0.66)	(0.90)	(1.14)	(1.21)

Note: Accumulated effects are presented for each of the 76 months following the treatment. Significance levels are bootstrapped with 499 replications.

F Results of robustness checks

Table F.1: ATT pre-interruption period

	Months before interruption									
	-1	-2	-3	-4	-5	-6	-12	-24	-36	-72
Employment probability	-0.001	-0.008	-0.004	-0.001	-0.005	0.002	0.045	0.013	0.008	0.000
	(0.003)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.018)	(0.018)	(0.018)
Monthly Earnings	-5	10	-10	3	16	1	63	55	21	-5
	(2)	(23)	(24)	(25)	(26)	(28)	(43)	(43)	(40)	(41)
Marginal employment	-0.017	-0.007	-0.005	0.002	0.003	0.002	-0.010	-0.017	-0.026	-
	(0.015)	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)	(0.011)	(0.008)	-

Note: Effects are presented for each of the 6 months directly before interruption, as well as for 12, 24, 26, and 72 months before interruption. Marginal employment can be observed in the data since 1999. Hence, I cannot implement the impact evaluation 72 months before treatment for this outcome. Standard errors are bootstrapped with 499 replications.

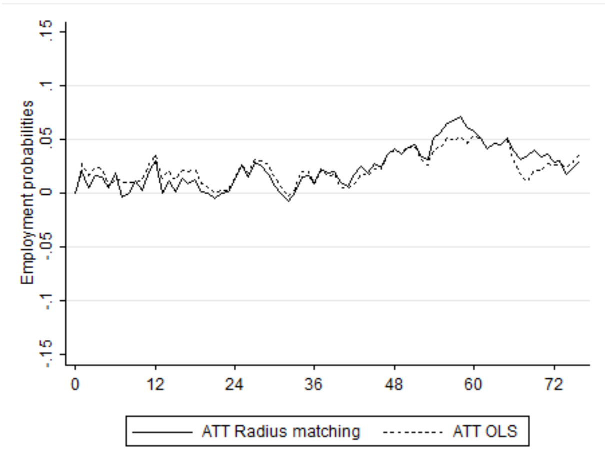
Table F.2: ATT pseudo treatment

	Months before interruption											
	1	2	3	4	5	6	7	8	9	10	11	12
Employment probability	0.016 (0.021)	0.006 (0.021)	0.003 (0.021)	0.012 (0.020)	0.002 (0.020)	0.000 (0.020)	- 0.015 (0.020)	-0.019 (0.020)	-0.025 (0.020)	-0.012 (0.019)	-0.010 (0.018)	-0.008 (0.018)
Monthly earnings	-25 (103)	-5 (97)	-12 (99)	12 (97)	-11 (98)	-11 (95)	-21 (90)	7 (85)	4 (81)	41 (81)	46 (82)	48 (84)
Marginal employment	0.033 (0.018)	0.027 (0.018)	0.024 (0.017)	0.025 (0.017)	0.021 (0.018)	0.018 (0.018)	0.022 (0.017)	0.020 (0.017)	0.021 (0.018)	0.008 (0.018)	0.017 (0.016)	0.007 (0.015)

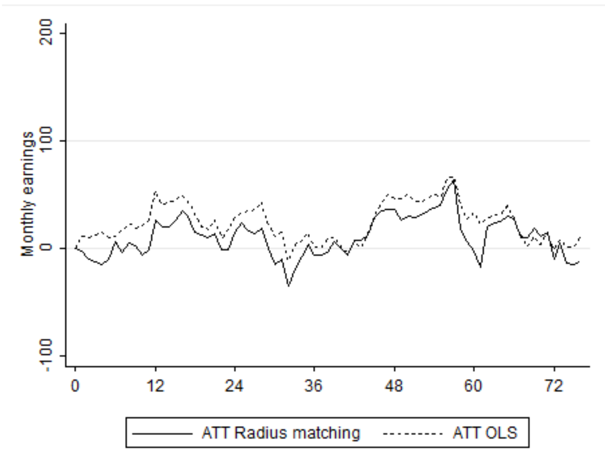
Note: Effects are presented for each of the 12 months following the pseudo treatment date. Standard errors are bootstrapped with 499 replications.

Figure F.1: ATT for unredeemed training vouchers

(a) Employment probability



(b) Monthly earnings (in Euros)



Note: Effects are estimated for each of the 76 months following the treatment. Diamonds report significant point estimates at the 5 percent level. Significance levels are bootstrapped with 499 replications. Lines without diamonds indicate that point estimates are not significantly different from zero.

G More results on heterogeneous effects

Table G.1: ATT pre-interruption period by vocational degree

	Months before interruption									
	-1	-2	-3	-4	-5	-6	-12	-24	-36	-72
No vocational degree										
Employment probability	0.000 (0.022)	0.000 (0.034)	0.006 (0.035)	0.020 (0.035)	0.003 (0.036)	0.005 (0.034)	0.089 (0.036)	0.063 (0.036)	0.054 (0.036)	0.009 (0.034)
Monthly Earnings	-24 (12)	-17 (46)	-31 (51)	-12 (45)	11 (45)	-1 (49)	114 (58)	119 (60)	78 (60)	-18 (54)
Vocational degree										
Employment probability	-0.011 (0.008)	0.029 (0.021)	0.021 (0.016)	0.039 (0.023)	0.041 (0.023)	0.049 (0.024)	0.039 (0.022)	0.022 (0.021)	0.025 (0.019)	0.008 (0.026)
Monthly Earnings	-54 (12)	46 (28)	32 (30)	17 (37)	24 (39)	55 (42)	100 (48)	60 (39)	-19 (45)	29 (52)
Academic degree										
Employment probability	0.000 (0.030)	0.015 (0.075)	0.010 (0.078)	0.007 (0.076)	-0.012 (0.076)	-0.012 (0.077)	0.071 (0.073)	0.079 (0.084)	-0.037 (0.081)	0.006 (0.084)
Monthly Earnings	-1 (8)	-106 (104)	-154 (134)	-67 (136)	-24 (135)	-85 (136)	47 (141)	107 (180)	-69 (183)	40 (188)

Note: Effects are presented for each of the 6 months directly before interruption, as well as for 12, 24, 26, and 72 months before interruption. Standard errors are bootstrapped with 499 replications.

Table G.2: ATT pre-interruption period by course type

	Months before interruption									
	-1	-2	-3	-4	-5	-6	-12	-24	-36	-72
Practice firm training										
Employment probability	-0.002 (0.003)	0.025 (0.043)	0.022 (0.045)	0.030 (0.045)	0.019 (0.045)	0.024 (0.045)	0.079 (0.045)	-0.013 (0.036)	-0.021 (0.038)	-0.004 (0.042)
Monthly Earnings	-8 (6)	50 (44)	31 (60)	52 (67)	12 (72)	3 (77)	95 (74)	113 (98)	-32 (91)	9 (103)
Short training courses										
Employment probability	-0.004 (0.003)	0.009 (0.020)	0.002 (0.020)	0.004 (0.021)	0.010 (0.020)	0.016 (0.021)	0.051 (0.025)	0.015 (0.023)	0.019 (0.022)	0.026 (0.026)
Monthly Earnings	-2 (2)	35 (37)	-7 (35)	-10 (39)	19 (42)	12 (50)	97 (60)	86 (64)	102 (58)	67 (66)
Long training courses										
Employment probability	-0.004 (0.004)	-0.008 (0.035)	0.000 (0.036)	0.010 (0.037)	-0.005 (0.035)	0.001 (0.032)	0.033 (0.031)	0.017 (0.026)	0.033 (0.024)	0.030 (0.034)
Monthly Earnings	-6 (3)	-15 (38)	-10 (44)	38 (49)	30 (54)	10 (58)	49 (67)	49 (70)	110 (62)	82 (66)
Retraining										
Employment probability	-0.001 (0.002)	0.015 (0.020)	0.035 (0.021)	0.039 (0.021)	0.044 (0.025)	0.049 (0.026)	0.072 (0.026)	0.009 (0.024)	0.028 (0.028)	-0.037 (0.024)
Monthly Earnings	-4 (3)	19 (22)	35 (24)	43 (27)	65 (28)	60 (33)	84 (39)	29 (34)	-26 (38)	-85 (35)

Note: Effects are presented for each of the 6 months directly before interruption, as well as for 12, 24, 26, and 72 months before interruption. Standard errors are bootstrapped with 499 replications.

Table G.3: Heterogeneity by program type and vocational degree

	6	12	24	36	48	60	72	76
				Practice firms and short training programs				
No vocational degree								
Employment	0.007	0.027	0.047	0.022	0.035	0.070	0.068	0.102
probability	(0.037)	(0.038)	(0.060)	(0.061)	(0.065)	(0.061)	(0.070)	(0.069)
Monthly	-26	52	73	52	76	86	77	102
earnings	(42)	(52)	(70)	(78)	(90)	(77)	(95)	(95)
Vocational degree								
Employment	0.016	0.085	0.101	0.078	0.085	0.120	0.115	0.115
probability	(0.019)	(0.026)	(0.027)	(0.031)	(0.032)	(0.033)	(0.033)	(0.032)
Monthly	-44	65	112	101	127	174	199	205
earnings	(25)	(32)	(34)	(39)	(41)	(43)	(48)	(50)
Academic degree								
Employment	-0.082	0.031	-0.019	-0.027	-0.119	-0.141	-0.043	-0.042
probability	(0.073)	(0.089)	(0.097)	(0.111)	(0.100)	(0.100)	(0.108)	(0.106)
Monthly	-136	104	118	117	81	-22	75	58
earnings	(100)	(146)	(166)	(172)	(174)	(218)	(210)	(221)
				Long training programs				
No vocational degree								
Employment	-0.077	-0.060	-0.016	0.087	0.139	0.161	0.189	0.183
probability	(0.043)	(0.055)	(0.076)	(0.076)	(0.083)	(0.077)	(0.091)	(0.087)
Monthly	-115	-63	9	239	301	249	286	256
earnings	(46)	(69)	(86)	(113)	(118)	(110)	(136)	(113)
Vocational degree								
Employment	-0.063	0.044	0.155	0.152	0.140	0.207	0.211	0.194
probability	(0.023)	(0.034)	(0.038)	(0.045)	(0.042)	(0.046)	(0.046)	(0.044)
Monthly	-174	10	216	211	197	300	375	354
earnings	(32)	(41)	(43)	(47)	(50)	(55)	(60)	(63)
Academic degree								
Employment	-0.045	-0.045	-0.003	-0.104	-0.042	-0.055	0.004	-0.007
probability	(0.070)	(0.096)	(0.111)	(0.116)	(0.116)	(0.117)	(0.116)	(0.114)
Monthly	-182	-117	1	-149	-61	-428	99	81
earnings	(114)	(166)	(192)	(201)	(220)	(477)	(275)	(272)
				Retraining programs				
No vocational degree								
Employment	-0.060	-0.046	-0.048	0.107	0.282	0.333	0.321	0.332
probability	(0.024)	(0.025)	(0.038)	(0.044)	(0.050)	(0.049)	(0.054)	(0.053)
Monthly	-49	-56	-27	179	413	456	478	506
earnings	(28)	(29)	(41)	(56)	(68)	(74)	(79)	(80)
Vocational degree								
Employment	-0.103	-0.143	-0.127	0.042	0.248	0.242	0.214	0.223
probability	(0.019)	(0.026)	(0.028)	(0.034)	(0.037)	(0.039)	(0.039)	(0.040)
Monthly	-149	-189	-127	100	349	325	322	337
earnings	(23)	(30)	(34)	(42)	(53)	(53)	(55)	(54)
Academic degree								
Employment	-0.047	-0.178	-0.142	-0.083	0.167	0.216	0.106	0.185
probability	(0.085)	(0.095)	(0.143)	(0.163)	(0.117)	(0.166)	(0.171)	(0.167)
Monthly	-166	-257	-221	-217	72	-29	-81	95
earnings	(137)	(186)	(213)	(270)	(310)	(588)	(392)	(377)

Note: Significance levels are bootstrapped with 499 replications.