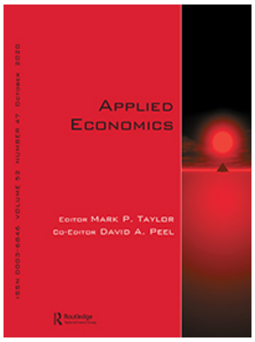

Applied Economics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/raec20

Retraining for the unemployed and the quality of the job match

Philipp Grunau & Julia Lang

APPLIED ECONOMICS
2020, VOL. 52, NO. 47, 5098–5114
https://doi.org/10.1080/00036846.2020.1753879

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Retraining for the unemployed and the quality of the job match

Philipp Grunau[a] and Julia Lang[b]

[a]Department Establishments and Employment, Institute for Employment Research, Nuremberg, Germany; [b]Department Active Labour Market Policies and Integration, Institute for Employment Research, Nuremberg, Germany

ABSTRACT

In Germany, retraining is an important measure in active labour market policy, providing unemployed individuals with extensive vocational training. Using administrative data, we show that retraining participants are more likely to take up jobs that require their educational degree and are more often employed in those occupations for which they have received vocational training. Moreover, retraining leads to higher earnings. As these effects may be driven by the positive impact of retraining on employment, we additionally try to isolate the direct effect by restricting our analyses to those formerly unemployed who find employment irrespective of participation in retraining.

KEYWORDS

Retraining; vocational training; evaluation of active labour market policies; overeducation; occupational mismatch

JEL CLASSIFICATION

I21; J24; J31; J62; J68

I. Introduction

In Germany, the catalogue of active labour market policy measures includes a long-term training programme called *retraining*, which aims at providing low-qualified unemployed individuals with extensive two- to three-year vocational training. To date, many studies have analysed the effectiveness of several training programmes for the unemployed (Card, Kluve, and Weber 2010, 2015). Some of these examined the specific case of retraining in Germany (e.g. Lechner, Miquel, and Wunsch 2007, 2011; Fitzenberger and Völter 2007; Fitzenberger, Osikominu, and Völter 2008; Doerr et al. 2017; Kruppe and Lang 2018), showing that retraining increases the probability of re-employment and earnings after a strong lock-in period. However, only few studies have addressed the qualitative dimensions of the programme's treatment effects. Lechner and Melly (2010) focus on wages and estimate bounds for the treatment effects of retraining beginning in the 1990s. They find a substantial increase in the earnings capacity of the retraining participants. Moreover, Dengler (2019) analyse the effects of different active labour market programmes for welfare recipients on various dimensions of job quality and work quality. For further vocational training, which includes retraining, she finds that participation increases the probability of finding a high-quality job, as there are positive effects on regular employment, earnings, stable employment and occupational exposure. Employing different measures of job (match) quality, we address another qualitative dimension of the treatment effect of retraining. In addition to earnings, we consider situations of educational mismatch in terms of both overeducation (vertical adequacy) and being employed in the same occupation for which vocational training has been received (horizontal adequacy).

Educational mismatch often occurs when unemployed people take up a new job. A (longer) period of unemployment may result in a depreciation of skills and could increase the job seeker's willingness to make concessions when taking up a job, i.e. lowering her/his reservation criteria (Munch and Skipper 2008) and considerations with respect to the choice of occupation. In this paper, we analyse the effects of retraining for unemployed individuals in Germany on the matching quality of subsequent jobs. Participants who successfully finish this active labour market policy programme obtain a vocational degree equivalent to that of initial vocational training. Retraining is designed to focus on occupations that are currently in high demand in the labour market (Kruppe and Lang 2018) and provides up-to-date skills. Thus, retraining should not only increase the

CONTACT Philipp Grunau Philipp.Grunau@iab.de Department Establishments and Employment, Institute for Employment Research, Regensburger Str. 104, Nuremberg 90478, Germany

likelihood to take up a job in general but also improve participants' job match quality.

Vertical mismatches refer to situations in which there is a discrepancy between the amount of formal education attained and the education required by the current job. On average, across selected OECD countries, one out of four workers reports such a mismatch between existing skills and those required for the job, where overeducation – a higher level of education than that required by the job – is more common than undereducation (McGowan and Andrews 2015). Situations in which workers' formal education is incommensurate with the education required by their current job can have negative consequences. For instance, overeducated workers have been shown to suffer a wage penalty in comparison to their correctly matched educational peers (for an overview, see Leuven and Oosterbeek 2011, or McGuinness 2006). Moreover, they show adverse mental health conditions (Chen, Smith, and Mustard 2010; Bracke, Pattyn, and von Dem Knesebeck 2013) and lower job satisfaction, although there is some disagreement regarding the latter issue. As the permeability between overeducation and adequate employment decreases, these disadvantages become more important. Therefore, findings showing that overeducation persists emphasize the long-term problems of this employment state (Dolton and Vignoles 2000; Sloane, Battu, and Seaman 1999; Battu, Belfield, and Sloane 1999 for Britain; Büchel and Mertens 2004; Pollmann-Schult and Büchel 2004 for Germany; Verhaest and Schatteman 2010, for Belgium). However, vertical mismatch affects not only individuals but also firms, as workers with an excess or lack of formal education affect firm-level productivity (Kampelmann and Rycx 2012, for Belgium; Grunau 2016, for Germany). Finally, misinvestment in education also entails costs for society and the economy in general (e.g. McGuinness 2006; OECD 2011).

The potentially detrimental effects of educational mismatches may appear not only in the vertical dimension but also in the horizontal dimension. In the latter, a worker is regarded as mismatched if the occupational tasks of his or her current job do not correspond with the occupation for which he or she received vocational training. For instance, this situation tends to be accompanied by substantial wage penalties (Nordin, Persson, and Rooth 2010, for Sweden; Robst 2007, for the US).

We use rich German register data that allow us to identify the entire population of retraining participants in 2004. We compare this group of trained unemployed to similar non-participants by applying propensity score matching with a large set of covariates that provide socio-economic characteristics, a detailed description of each unemployed person's employment and unemployment history and regional information. We estimate the effects of retraining on three measures of job match quality – earnings, overeducation, and working in the trained occupation – up to seven years after the start of the treatment in 2004. As the effects on these outcomes may be driven by the positive impact of retraining on employment, apart from standard regressions addressing the total effect, we also try to isolate the direct effect of the program by restricting the analysis to those formerly unemployed who find employment irrespective of participation in retraining.

II. Retraining and job match quality

As part of active labour market policy, retraining provides intensive vocational training for unemployed individuals in Germany. Those who either have never completed any vocational training or have not worked in their learned occupation for several years are eligible. Successful completion leads to a vocational degree that is equivalent to a degree obtained in the German apprenticeship system. Since the law requires a reduction of training duration by one third in the case of retraining for unemployed individuals, retraining is normally shorter than initial vocational training. Nevertheless, due to the two- to three-year duration of this programme, retraining can be considered by far the most extensive training programme for unemployed workers in Germany.

Since 2003, those unemployed individuals who are eligible for retraining can receive a training voucher indicating the objective – i.e. the target occupation – and the duration of training. With this voucher, workers can choose a suitable course by themselves within a defined period of time (at most three months, most often one month). Once the voucher has been redeemed, participation in the program is compulsory, and failure to participate may result in benefit sanctions.[1]

In this study, we concentrate on retraining starting in 2004 during a period of major labour market reforms in Germany. The most far-reaching part of the so-called Hartz reforms implemented between 2003 and 2005 was the introduction of means-tested welfare benefits (unemployment benefit II) in 2005. Since 2005, both the unemployed in the unemployment insurance system and the unemployed receiving unemployment benefit II in the welfare system can participate in retraining. During our observation period, however, only unemployed individuals in the unemployment insurance system, who are usually more closely linked to the labour market, were eligible for retraining programmes. Figure 1 shows the trends in the number of participants and the share of retraining in all subsidized training courses between 2000 and 2018. Starting with a high level of more than 95,000 entries into retraining in 2000, participation numbers strongly decreased in the course of the Hartz reforms, which attached greater importance to short-term measures and fast integration into the labour market. There was a decrease in not only the absolute participation numbers but also the relative importance of retraining. In 2003, retraining accounted for more than 25% of all entries into training courses, but only for approximately 7% between 2006 and 2009. After a minimum of approximately 16,000 entries in 2005, participation numbers increased again and amounted to more than 45,000 entries in 2018. In light of the new requirements on workers' qualifications and skills due to technological and structural changes and, at the same time, a lack of skilled labour in some occupations, retraining has moved back into the focus of attention in recent years.

Retraining strongly increases both the general and the occupation-specific human capital of participants, and thus, their employability. However, the overall effect of a higher employment probability of participants on the matching quality is not clear a priori. The likelihood of finding an adequate match will certainly be higher if the participants' new occupations obtained by retraining are indeed currently in demand. However, if retraining increases the participants' general employability,

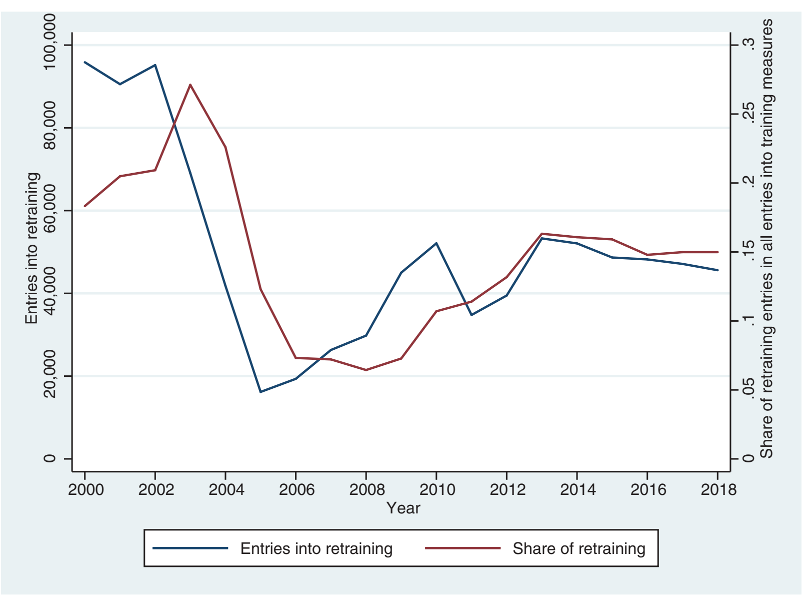

Figure 1. Number of entries into retraining and share of total training entries between 2000 and 2018.
Source: Statistics Service of the Federal Employment Agency.

[1]Caseworkers cannot sanction unemployed individuals for not redeeming the voucher. For a more detailed description of the voucher system in Germany, see Doerr et al. (2017). Before 2003, caseworkers directly assigned participants to specific retraining courses.

they may also be more likely to find a job in other occupational fields or a low-qualified job, and an increase in employment can also increase the share of mismatched workers.[2]

In contrast, predictions about other direct mechanisms through which retraining may affect job match quality are much clearer. Since 2004, employment agencies have been required to provide training plans for their region for the subsequent year, where they determine the occupations on which (re) training should focus and the number of people who should be trained in a certain occupation (IZA, DIW and Infas 2005). Based on these training plans, case-workers are instructed to choose retraining occupations according to the forecasted actual demand in the regional labour market. Because of this targeted assignment, we expect retraining to increase the likelihood that participants will find a job that matches their (new) qualifications.[3]

Moreover, unemployed people and employees who are not working in their learned occupations experience a depreciation of their (occupation-specific) human capital and find that the skills that they previously acquired become obsolete over time (Rosen 1975; Neuman and Weiss 1995). The longer a person is unemployed, the stronger these negative effects can become. Thus, another advantage that retraining participants have over those who do not attend such a program is their procurement of up-to-date skills. Even if both participants and non-participants ultimately hold the same vocational degrees, retraining participants should have better employment prospects because they acquired up-to-date skills and knowledge in this occupational field.[4] To conclude, besides an indirect effect via employment, there are different channels through which retraining can have a direct impact on our outcomes. Both due to the targeted assignment and the recency of skill attainment, we expect participation in retraining to improve the matching quality of participants' subsequent jobs.

Regarding the impact of retraining on earnings, there is an indirect positive effect via the increased likelihood of employment. As the share of employed individuals is expected to be higher for participants than for non-participants, the share of individuals with non-zero earnings is higher. However, the direct effect of retraining on (hourly) wages is unclear a priori. On one hand, due to the recent attainment of fresh skills and knowledge, treated individuals may achieve higher wages. However, on the other hand, untreated workers, even if they hold the same vocational degree, have the advantage of pre-unemployment occupation-specific experience, which can exert a positive impact on wages (e.g. Altonji and Williams 1998, 2005). Finally, yet importantly, regarding the targeted assignment into occupations of high demand, the overall impact may largely be contingent upon their specific location in the distribution of average remuneration across occupations. If high demand is non-randomly distributed, focusing on either above- or below-average-paying occupations, the impact of retraining on wages may look accordingly.

III. Estimation approach

To estimate the effects of retraining on the quality of the job match, we apply a matching approach through which we compare unemployed people who participate in retraining with comparable unemployed non-participants. The basic assumption of this matching approach is that, given all observable factors X that affect treatment status D and the outcome variable Y, the outcome without treatment is independent of the treatment status (conditional independence assumption, CIA). Rosenbaum and Rubin (1983) suggest matching on the treatment probability rather than matching on every single variable of vector X. In this case, the CIA is $Y_h^{D(t)=0} \perp P(X)$, with P(X) indicating the probability of being treated in month t given observable characteristics X. $Y_h^{D(t)=0}$ indicates the outcome Y h months after the treatment starts, given that an individual is not treated in month t ($D(t) = 0$). Then, the average treatment effect on the treated (ATET) can be written as follows:

[2]In a robustness check, we focus on the direct effect by comparing the job quality of employed participants with that of similar employed non-participants (see Section 5.2).

[3]For the specific case of (re)training in elderly care professions, Dauth and Lang (2019) show descriptively that approximately 25% to 50% of the participants have permanent jobs in the care sector after the training.

[4]However, non-participants' occupation-specific experience may reduce this advantage for participants.

$$\mathrm{ATET}^{h,t} = E\left(Y_h^{D(t)=1} | P(X), D(t) = 1\right) - E\left(Y_h^{D(t)=0} | P(X), D(t) = 0\right)$$

which is the difference in the average outcome Y of treated (D(t) = 1) and untreated (D(t) = 0) people after h months given their treatment probability P(X).

The ATET is the total effect of retraining on the outcome variable measure job match quality (earnings, horizontal or vertical (mis)match). As mentioned above, besides a direct impact of retraining on the different outcomes, there is an indirect effect working via a positive impact on employment, which we will address in the results section. For both analyses, we apply nearest neighbour matching with 1 neighbour and a calliper of 0.1.

Finally, because the standard derivative-based standard errors have been shown to be inconsistent after nearest neighbour matching and because bootstrapping is unreliable if it is performed with a fixed number of comparison observations (Abadie and Imbens 2008), we apply the adjusted estimation of standard errors proposed by Abadie and Imbens (2006, 2011, 2016).

IV. Data description

For our estimations, we use the Integrated Employment Biographies (IEB)[5] (see Dorner et al. 2010, for a description of a subsample of the data) from the Institute for Employment Research (IAB). These administrative data combine information for Germany from four different data sources: the IAB Employment History, the IAB Benefit Recipient History, the participants-in-measures data, and the data on job searches derived from the applicant pool database. The dataset includes all the individuals' employment (subject to social security contributions) and unemployment spells on a daily basis, as well as spells in which they receive unemployment benefits and information on their participation in active labour market programmes. Moreover, the dataset includes information on individual characteristics and regional information. For employed workers, there is also information on the employer and daily wages.

This rich data source enables us to apply a matching approach in which we compare unemployed individuals who participate in retraining with comparable unemployed non-participants. For our estimations, we define all individuals who began a retraining course in 2004 as the treatment group. Additionally, we draw a random sample of people who were unemployed at least once in 2004 as potential controls. To improve the matching process, we only match the unemployed who started retraining in a certain month in 2004 with the unemployed who were not in an active labour market programme in that month. However, we report aggregate results for all 12 months. The data allow us to follow the matched individuals for up to 7 years after the treatment began and to compare different outcomes on a monthly basis. Because retraining lasts up to three years, we present treatment effects 48 months and 84 months after the start of treatment.

We operationalize job quality using three different measures: First, we examine the effects of retraining on earnings. For this, we set earnings to zero for unemployed workers. Thus, as mentioned before, there is a positive indirect effect on earnings through employment as we assume that retraining increases the employment probability. The direct effect (wage effect), however, and thus also the overall effect, is unclear a priori.

Second, we address job match quality in terms of vertical educational mismatches, i.e. discrepancies in the amount of the workers' attained formal education and the education that their current jobs require. Because undereducation is not usually associated with individual-level negative effects, we focus on overeducation. To measure overeducation, we compare the workers' highest obtained educational degree with the degree that his or her current job usually requires. We determine required education by means of a large-scale administrative dataset from 2008, which contains sufficient observations on employees for each of the 334 occupations represented by the 3-digit level of the occupational classification *KldB 1988*. We take the mode of actually attained educational degrees over all workers within each occupation as its required education, which allows us to compare

[5]We use version V10.00 of the IEB and additional variables from the participants-in-measures data (MTH version V06.02-201204).

them with individuals' degrees and to identify overeducation.

Third, we examine the horizontal dimension of educational mismatches, i.e. when the occupation for which a worker received vocational training does not match the occupation of his or her current job. To examine the horizontal dimension of educational adequacy, we compare the occupation in which a worker is currently employed with the occupation for which s/he received her/his latest vocational training (for participants, this is retraining). To derive the latter information for non-participants, we draw on administrative data on the jobseeker histories of unemployed workers. These data contain information on the occupation for which they received their most recent educational training.

As mentioned previously, matching methods rely on the validity of the unconfoundedness assumption or conditional independence assumption (CIA). Because this assumption cannot be tested, it is crucial for identification to control for the relevant factors that affect treatment status and/or outcomes. Michael and Wunsch (2013) discuss the importance of different sets of variables for estimating the treatment effects of active labour market programmes. We follow their advice and include a rich set of control variables, which are listed in Table A1 in the Appendix. As employment agencies' and their caseworkers' success is partly evaluated according to the rate at which unemployed individuals are reintegrated into the first labour market, a common criticism is that case workers tend to reserve training measures for the unemployed persons who they believe will most likely be successful rather than those who could benefit most from participation. To address this potentially detrimental cream-skimming effect, we control for various variables that capture the pre-unemployment state in the propensity score matching (e.g. the employment history up to seven years prior to the current unemployment spell, the status before the current unemployment spell and the characteristics of the last job; see Table A1 in the Appendix). As an arguably crucial factor in our case, we also consider the job match quality in the last job before unemployment.

Our analysis sample comprises 27,427 treated unemployed workers who took part in a retraining measure in 2004 without missing values in the covariates and 134,094 untreated unemployed workers in the control group. In Table 1, we show mean values of some basic (pre-treatment) socio-demographics for the participants and non-participants and distinguish between workers who are employed and unemployed after 7 years. Moreover, we also

Table 1. Basic descriptive statistics for the treatment and control groups.

	Treatment group			Control group		
	All	Employed after 84 months	Unemployed after 84 months	All	Employed after 84 months	Unemployed after 84 months
Number of observations	27,427	16,965	10,462	134,094	49,054	85,040
Employed after 84 months	0.619			0.366***		
Female	0.471	0.470	0.471	0.330***	0.283***	0.355***
Age	33.23	33.31	33.10	36.07***	34.504***	36.978***
Marital status						
Single	0.418	0.394	0.458	0.427***	0.425***	0.428***
Not married, not living alone	0.077	0.075	0.080	0.076	0.075	0.077
Single parent	0.119	0.117	0.122	0.077***	0.068***	0.082***
Married	0.386	0.414	0.340	0.420***	0.432***	0.413***
Children	0.432	0.457	0.390	0.551***	0.553***	0.551***
Nationality: German	0.908	0.913	0.901	0.861***	0.882***	0.850***
Position in last job						
Blue-collar worker	0.355	0.347	0.368	0.422***	0.390***	0.441***
Skilled worker	0.180	0.194	0.158	0.260***	0.316***	0.228***
White-collar worker	0.265	0.262	0.268	0.170***	0.173***	0.169***
Part-time worker	0.196	0.192	0.202	0.144***	0.118***	0.159***
Highest vocational degree						
No degree	0.337	0.308	0.385	0.340	0.264***	0.384
Vocational degree	0.621	0.651	0.574	0.626	0.699***	0.583*
University degree	0.041	0.041	0.041	0.034***	0.037**	0.032***
Educational (mis)match in last job						
Overeducation	0.053	0.054	0.052	0.042***	0.042***	0.043***
Horizontal adequacy	0.164	0.170	0.153	0.263***	0.296***	0.240***

Source: Integrated Employment Biographies. */**/*** indicate significant differences in mean values between the treatment and control group at the 10/5/1% levels for the three samples (all/employed/unemployed after 84 months), respectively.

report the share of workers who were overeducated in their last job before unemployment (for retraining participants, this relates to their last vocational training before retraining) and the share of workers who have been employed in an occupation in which they also received training (horizontal adequacy). Statistically significant differences between the treatment and control group are denoted by asterisks.

First, participants are significantly younger than non-participants. Women appear to be overrepresented in the treatment group. Compared to all non-participants, retraining participants are significantly less likely to have children but more likely to be single parents. With regard to the position held in the last job before unemployment, non-participants were more often (skilled) blue-collar workers, whereas participants more often worked in white-collar jobs and more often were part-time employed. Although participants are slightly more likely to hold a university degree, there are no significant differences in the share of workers holding a vocational degree. Regarding the variables measuring qualification (mis)match in the last job before the current unemployment spell (i.e. the variables are constructed comparing the latest vocational degree before potential participation in retraining), workers in the control group had better job matches than those in the treatment group. Of the non-participants, 4.2% were overeducated, which is significantly less than the corresponding 5.3% of the retraining participants. Compared to a share of 16.4% in the treatment group, 26.3% of non-participants were also more likely to work in an occupation for which they were trained. Thus, there is no incidence of cream skimming with respect to these variables. In contrast, those workers who experienced a mismatch between existing skills and those required for the former job were more likely to receive retraining. However, this is not surprising because mismatched workers – who have not worked in their learned occupation for several years – belong to the programme's target group.

Moreover, Table 1 shows that the share of participants who are employed after 84 months is 61.9% and much higher than for non-participants, whose employment share is only 36.6%. Comparing workers who are employed after seven years with those who are not shows that women are as likely as men to take up a job again if they participated in retraining, whereas within the control group, women are significantly less likely to find a job (the results of the t-tests are not reported in Table 1). Overall, differences are much more pronounced between the treatment and control group than between employed and unemployed workers within the two groups.

To understand the dimension of educational (mis)match for (re-employed) workers in our sample, Table 2 presents the mean values of the outcome variables for the total sample and for those who are employed 84 months after (potential) treatment began, separately for men and women. As these variables are measured after participants have completed retraining, for the treatment group they refer to (mis)matches with regard to the qualifications acquired by retraining.

The shares of overeducated workers in our sample are comparably small: 2.2% of all employed men and 5.8% of all employed women. However, even though men earn significantly more than women and are less likely to be overeducated, the share of workers with an adequate horizontal job match is higher among women: 26.8% of all employed men and 37.7% of all employed women work in their learned occupation.

V. Econometric evidence

Because many studies find gender differences in the effects of subsidized training (e.g. Lechner, Miquel,

Table 2. Mean values of outcomes for workers in month 84 by gender.

	Men		Women	
	All (N = 104,340)	Employed (N = 44,017)	All (N = 57,181)	Employed (N = 22,002)
Earnings	24.70	58.65	17.00***	44.39***
Overeducation	0.008	0.022	0.019***	0.058***
Horizontal adequacy	0.083	0.268	0.120***	0.377***

Source: Integrated Employment Biographies. */**/*** indicate significant differences in mean values between men and women at the 10/5/1% levels, respectively.

and Wunsch 2011; Biewen et al. 2014), we carry out all estimations separately for men and women. We impose common support by dropping all observations from the treatment group with a propensity score below the minimum propensity score or above the maximum propensity score for the control group, which only applies to about 1 percent of our observations. To check the quality of the statistical matching, in our case, nearest neighbour matching, we calculate the mean standardized bias (Rosenbaum and Rubin 1983) across all included variables between both the treatment and the control group. In the following tables, these measures are displayed for both the pre- and post-matching state separately for each matching.

Treatment effects from retraining on the quality of the match

Before we present our results on the impact of retraining on job match quality, we briefly discuss the programme's effects on the likelihood of re-employment. We deliberately chose to keep this discussion brief as this nexus has already been addressed in many other papers before and the focus of our paper goes beyond that. Table 3 shows strong and significant positive ATET coefficients, implying that participation in the retraining program involves a higher likelihood to find a job. This holds for both points in time considered and is particularly noteworthy after 48 months, considering that retrainees were locked in the programme for a substantial part of this time. In more detail, retraining increases the probability of re-employment by 17.1 percentage points for men and by 26.6 percentage points for women 48 months after the start of the programme.

Returning to the outcome variables reflecting job match quality, first, our estimated ATETs (Table 4) imply that participation in retraining leads to higher earnings. This impact is more pronounced for women, for whom retraining increases earnings by approximately €13 a day – as opposed to a rise in daily earnings for men of about €10. This discrepancy in the ATETs between women and men can point towards three potential explaining conjunctures: First, as part of the earnings effect is due to higher employment shares among participants and our results in Table 3 show that employment effects are more pronounced for women than for men, the indirect effect via employment could matter. Second, as part of the targeted assignment of unemployed workers to high-demand vocational training occupations, female participants in particular tend to opt for/be assigned to better-paying occupations compared to what they would have chosen otherwise. However, with average daily earnings of almost €44 in month 84, treated women still earn less than their male counterparts, who receive daily earnings of almost €59 on average. Finally, this stronger impact on women's earnings may also be driven by an improved job matching quality in the other two dimensions, which are subsequently addressed.

Turning to the likelihood of becoming overeducated, our corresponding ATETs reveal a significant negative impact for both men and women. Again, the treatment effect is substantially larger for female participants, whose likelihood of overeducation is reduced by 2.7 percentage points. For men, program participation decreases the risk of overeducation by 1.8 to 2.2 percentage points. Although the effect sizes may seem small, considering the low risk of overeducation in general (see Table 2) certainly puts these effects in perspective. As mentioned above, retraining increases employment, and thus may also increase the probability that participants will be overeducated. Thus, since we find that retraining decreases the probability of being overeducated, the program seems to redirect individuals who would have found employment anyway to different jobs for which they are not overeducated. We discuss this direct effect in more detail in Section 5.1.

Table 3. ATETs of retraining on the likelihood of re-employment, by gender and observation point (in months).

	Men		Women	
After …	48 months	84 months	48 months	84 months
Re-employment probability				
ATET	0.171***	0.162***	0.266***	0.223***
SE	(0.009)	(0.009)	(0.011)	(0.011)
Post-matching mean bias (pre-match.)	1.8 (15.2)		2.1 (13.8)	
Observations	100,716	100,716	54,770	54,770

Source: Integrated Employment Biographies. Coefficients depict the average treatment effects on the treated (ATETs). Standard errors (in parentheses) and number of observations are displayed directly beyond the coefficients. */**/*** indicate the significance of the coefficients at the 10/5/1% levels, respectively.

Table 4. ATETs of retraining on several outcomes of job (match) quality by gender and observation point (in months).

	Men			Women		
After ...	48 months		84 months	48 months		84 months
Gross daily earnings (in Euros)						
ATET	10.413***		10.176***	13.654***		12.753***
SE	(0.589)		(0.620)	(0.586)		(0.610)
Post-matching mean bias (pre-match.)		1.8 (15.2)			2.1 (13.8)	
Observations	100,596		100,593	54,650		54,603
Overeducation						
ATET	−0.022***		−0.018***	−0.027***		−0.027***
SE	(0.002)		(0.002)	(0.003)		(0.003)
Post-matching mean bias (pre-match.)		1.9 (15.1)			1.9 (13.7)	
Observations	99,040		88,938	54,069		50,622
Horizontal adequacy (*3-digit level*)						
ATET	0.169***		0.164***	0.266***		0.239***
SE	(0.008)		(0.008)	(0.010)		(0.010)
Post-matching mean bias (pre-match.)		2.1 (14.7)			2.3 (12.8)	
Observations	70,214		63,220	40,616		38,150

Source: Integrated Employment Biographies. Coefficients depict the average treatment effects on the treated (ATETs). Standard errors (in parentheses) and number of observations are displayed directly beyond the coefficients. */**/*** indicate the significance of the coefficients at the 10/5/1% levels, respectively.

Addressing the horizontal dimension of educational mismatches, our estimates reveal strong ATETs from retraining on the likelihood of taking up employment in the same occupation for which vocational training (for participants, this is retraining) has been received. We compare occupations based on the 334 disjoint occupations specified by the 3-digit level of the occupational classification *KldB 1988*. The ATETs for men attain positive values of 16.9 percentage points after 48 months and 16.4 percentage points after 84 months, while for women, retraining increases the likelihood of horizontal adequacy by 26.6 percentage points after 48 months and 23.9 percentage points after 84 months. Again, the substantially larger absolute treatment effect for women could be both due to a direct effect of programme participation and an indirect impact via the bigger effect of retraining on the women's probability of re-employment.

Regarding the matching quality, the mean standardized biases are reduced to values between 1.8 and 2.3 and therefore are considerably lower than what has been proposed as sufficient in the literature (Caliendo and Hujer 2006).[6]

Conditioning upon employment

Naturally, policy evaluation requires comparing participants to non-participants who are similar to each other, the results of which we presented in the preceding section. However – although this is a standard procedure in this type of research – some effects may seem almost mechanical due to the circumstance that our outcomes can only be measured in cases of observed employment. Therefore, as retraining significantly increases employment (see Table 3), the coefficients of a standard estimation approach as presented above are a mixture of the direct and indirect effects of programme participation and – although the ATETs are valid from a programme evaluation perspective – may also be biased because they reflect this non-random selection into employment.

As we already established in Section 5.1, retraining exerts a positive impact on the probability of re-employment. Because of this, the indirect effect via employment will most likely increase both the shares of treated workers with positive earnings and with an adequate educational match, but also the share of mismatched (overeducated) retraining participants. In the extreme case, all of the additionally employed treated workers would, for instance, be overeducated. In the other extreme case, all additionally employed treated workers would find a job that matches their qualifications, resulting in an indirect effect on overeducation of zero.

In order to shed light on this issue, we ideally applied an exclusion restriction. However, although we have

[6]Identifying causal treatment effects using (propensity score) matching as a pre-processing step is sometimes criticized because of the potential problem that balancing certain covariates may lead to a decreased balance of others, hence counteracting the intended bias reduction (e.g. Ho et al. 2007). Therefore, as a robustness test, we use entropy balancing (Hainmueller 2012) instead of propensity score matching, which leads to similar results (cf. Table A2 in the Appendix).

a variety of explanatory variables, we do not have any variables at hand that are associated with employment and, at the same time, do not directly affect the quality of the job match. One solution to this problem of double selection – into participation and employment – is to estimate bounds for the treatment effects for specific subpopulations (e.g. Lee 2009; Zhang, Rubin, and Mealli 2008; Lechner and Melly 2010; Blanco, Flores, and Flores-Lagunes 2013), but most of these approaches are only applicable if assignment to treatment is random. Another approach, which is also valid when selection into treatment is based on a set of observable covariates, is suggested by Flores and Flores-Lagunes (2009), who estimate the local net average treatment effect (LNATET) for a subpopulation for which the treatment arguably does not affect the mechanism variable, i.e. employment.[7] We follow their approach, which in our case permits us to assess the effect of retraining on job (match) quality for the subgroup of people who are employed but whose employment probabilities are not affected by participation in a retraining programme.

To ensure this, we employ a matching approach involving predicted values of employment, given the covariates vector X. More specifically, we compare the employment status of every retraining participant with the status of its nearest neighbour and include only those pairs of treated and untreated individuals into our estimations that are employed at the same given point in time (because we only observe the outcome variables for employed individuals). The underlying idea of this approach is that if the participant and his/her nearest neighbour who did not participate in retraining have the same employment status, the participant's employment probability was arguably not affected by the programme. After the matching, we compare the difference in the mean outcomes only for the subgroup of pairs for which both the participant and his/her nearest neighbour are employed.[8] In contrast to the earnings outcome above, we now estimate the effect on daily wages, and thus can determine whether participants find better paid jobs.

Despite our restriction to those who arguably would have found employment irrespective of participation in retraining – which leads to a substantially reduced estimation sample – the coefficients displayed in Table 5 lead to similar conclusions as those in Table 4. Effect direction and significance mostly remain the same, but the magnitude of the effects differs.

For instance, the wage effect after 84 months is now only €6.5 for women and even insignificant for men, and thus much lower than in the standard approach. Consequently, a substantial part, albeit not all of the earnings effect we present in Table 4, can be explained by the indirect effect via employment.

Regarding the likelihood of overeducation, the effects are much more pronounced compared to the ATETs

Table 5. LNATETs of retraining on several outcomes of job (match) quality, by gender and observation point (in months).

	Men		Women	
After …	48 months	84 months	48 months	84 months
Gross daily wages (in Euros)				
LNATET	1.814***	.276	3.473***	6.455***
SE	(0.654)	(0.796)	(0.924)	(0.826)
Post-matching mean bias (pre-match.)	1.8 (15.2)		2.1 (13.8)	
Observations	5,531	5,992	3,333	3,822
Overeducation				
LNATET	−0.046***	−0.043***	−0.106***	−0.090***
SE	(0.005)	(0.006)	(0.011)	(0.010)
Post-matching mean bias (pre-match.)	1.9 (15.1)		1.9 (13.7)	
Observations	5,271	5,049	3,038	3,241
Horizontal adequacy (*3-digit level*)				
LNATET	0.218***	0.197***	0.299***	0.280***
SE	(0.015)	(0.017)	(0.021)	(0.021)
Post-matching mean bias (pre-match.)	2.1 (14.7)		2.3 (12.8)	
Observations	4,419	4,117	2,784	2,927

Source: Integrated Employment Biographies. Coefficients depict the local net average treatment effects on the treated (LNATETs). Standard errors (in parentheses) and number of observations are displayed directly beyond the coefficients. */**/*** indicate the significance of the coefficients at the 10/5/1% levels, respectively.

[7]For more details on the estimation approach, see Appendix A3.

[8]Employed participants whose nearest neighbour is not employed can be expected to have found a job because they were treated and are excluded from the analysis.

reported in Table 4. As only employed individuals can be overeducated, the indirect effect via employment is positive because it increases the probability to be overeducated among participants. Hence, including the indirect effect reduces the negative impact of retraining on overeducation in our standard estimates in Table 4 compared to the direct effects in Table 5.

Similarly, only employed individuals can work in the occupation they were trained for, and the indirect effect on horizontal adequacy via employment could be expected to be positive. The direct effects on having an adequate horizontal match that are presented in Table 5, however, are higher than the ATETs from Table 4. This can be explained by the inclusion of unemployed individuals in the standard ATET estimation approach at the point in time when we measure our outcome variables. Including the unemployed as zeros in the outcome variables biases any coefficient towards zero. In the case of horizontal adequacy, this effect seems to be even more pronounced than the indirect positive effect via employment.

Overall, our results on the direct effects imply that the selection into employment that is related to the participation status is not the sole driver of our main results. To put it differently, there is also a direct impact of retraining participation on job (match) quality beyond the indirect effect of increased employability.

Targeted assignment versus recency

Regarding the theoretical explanations behind our estimated treatment effects, we expect a positive impact of retraining on the job match quality because of two separate mechanisms: the targeted assignment of training occupations according to the predicted demand and the recency of educational attainment in the form of up-to-date learning contents. However, even though we find robust and substantial treatment effects to support our hypothesis of a positive impact of retraining on job match quality, thus far, we cannot adequately determine the decisive source of the estimated effects. Therefore, in this section, we attempt to disentangle the total treatment effects into its two components. To do so, we adjust the matching conditions: we only match individuals from the treatment and the control group who have received their most recent vocational training (for participants, this is retraining) in the same occupational group with respect to the 3-digit level of the occupational classification. We thus purge the ATETs of the impact stemming from the training occupation, hence

Table 6. ATETs of retraining with within-occupation comparisons and on several outcomes of job (match) quality, by gender and observation point (in months).

	Men		Women	
After …	48 months	84 months	48 months	84 months
Gross daily earnings (in Euros)				
ATET	8.822***	8.965***	11.979***	10.644***
SE	(0.982)	(1.035)	(1.098)	(1.133)
Post-matching mean bias (pre-match.)	5.8 (14.8)		5.8 (12.6)	
Observations	73,655	73,641	42,442	42,395
Overeducation				
ATET	−0.020***	−0.016***	−0.023***	−0.014**
SE	(0.004)	(0.004)	(0.006)	(0.006)
Post-matching mean bias (pre-match.)	6.2 (15.3)		6.0 (12.9)	
Observations	72,519	65,531	41,955	39,492
Horizontal adequacy (*3-digit level*)				
ATET	0.115***	0.107***	0.214***	0.191***
SE	(0.012)	(0.012)	(0.019)	(0.020)
Post-matching mean bias (pre-match.)	6.7 (14.7)		6.1 (12.8)	
Observations	69,770	62,783	40,217	37,756

Source: Integrated Employment Biographies. Coefficients depict the average treatment effects on the treated (ATETs). Standard errors (in parentheses) and number of observations are displayed directly beyond the coefficients. */**/*** indicate the significance of the coefficients at the 10/5/1% levels, respectively.

[9]However, this approach necessarily comes with a small drawback: Compared to all other reported estimates, where the share of treated observations off the common support is less than 1%, the proportion is somewhat higher, albeit still at a low level, in the case of an exact matching on occupations, at up to 4%. Accordingly, the matching quality in terms of the mean standardized bias also suffers somewhat as the number of potential controls is much lower when matching is performed within training occupations.

controlling for the targeted assignment. As such, we argue that the remaining treatment effects should be due to the recency of the vocational training.[9]

The ATETs displayed in Table 6 below show that the effects of retraining are significantly reduced in all our job match quality dimensions when compared to the ATETs in Table 4. This implies that targeted assignment is indeed at play as a mechanism driving the impact of retraining participation on job (match) quality. However, apart from decreasing the effect size to some extent, all coefficients retain their significance and even the greater part of their effect size. We interpret this as evidence for the importance of the recency of training for the treatment effect. Therefore, both theoretical mechanisms – targeted assignment and the recency of training contents – appear to be at play, with recency appearing to be (slightly) more important than targeted assignment.

Since we argue that parts of the impact may be indirect through increased employability, by analogy with the preceding estimations in Section 5.2, we rerun our regressions on a reduced subsample where matching is also conditional upon being employed in general, beyond the occupation for which the latest vocational training (retraining in case of participants) was received. The results (see Table A4 in the Appendix) also suggest that both mechanisms – targeted assignment and recency – are effective in the impact of retraining on job quality, although within that setting, targeted assignment appears as the main driver of the treatment effects.

VI. Conclusions

Previous studies have shown that extensive vocational training for the unemployed, i.e. retraining, improves participants' likelihood of finding a job thereafter. However, most analyses neglect the qualitative aspect of the treatment effect. Therefore, in this paper, we focus on the job match – that is, the quality of employment – after retraining.

Our results show that retraining not only improves the likelihood of re-employment but also increases earnings. Moreover, retraining considerably reduces the higher likelihood of unemployed workers ending up overeducated. Finally, treated individuals show a substantially higher likelihood of finding a job in the occupation for which they received vocational training. All the estimated ATETs can be considered large when analysed in relation to the level of job match quality without treatment.

Further restricting the matching of treated to untreated individuals within the same training occupation reveals that both theoretical mechanisms – the targeted assignment to occupations of high demand and the recency of educational attainment – appear to be at play. Since part of the estimated treatment effects (on the treated) are caused by an increase in the employment probability of participants, we additionally focus on the direct effects of programme participation using a subgroup of employed individuals whose re-employment probability was not affected by retraining. For this subsample, we also find that retraining has a positive impact on both wages and the likelihood of finding a job that matches the education of the participants.

However, an improved job (match) quality is desirable from more than an individual's perspective. Digitization and structural change are central challenges for the labour markets of many countries and are the focus of political attention. Due to these developments, jobs and their qualification requirements are expected to change considerably. Against this backdrop, further training can be a key to ensuring the employability of workers and addressing shortages in skilled labour. In this regard, it is important to know whether training programs are successful in integrating participants into the labour market via the intended occupation. Since retraining increases the likelihood of finding employment in the very occupation for which vocational training was received, our results suggest that retraining is not only a very promising strategy to integrate unemployed workers into the labour market but also to reallocate them to professions in which shortages of skilled labour exist and are most pronounced.

In our analysis we focus on participants who started retraining in 2004, i.e. before the introduction of means-tested welfare benefits in Germany in 2005. Thus, we only consider participants within the unemployment insurance system. However, unemployed individuals who receive unemployment benefit II within the welfare system differ from our group as they are generally less attached to the labour market and more often hard-to-place. Thus, for this group, the effects of retraining on the quality of the job match may differ and should be addressed in future research.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Abadie, A., and G. W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74 (1): 235–267. doi:10.1111/j.1468-0262.2006.00655.x.

Abadie, A., and G. W. Imbens. 2008. "On the Failure of the Bootstrap for Matching Estimators." *Econometrica* 76 (6): 1537–1557.

Abadie, A., and G. W. Imbens. 2011. "Bias-Corrected Matching Estimators for Average Treatment Effects." *Journal of Business and Economic Statistics* 29 (1): 1–11. doi:10.1198/jbes.2009.07333.

Abadie, A., and G. W. Imbens. 2016. "Matching on the Estimated Propensity Score." *Econometrica* 84 (2): 781–807. doi:10.3982/ECTA11293.

Altonji, J. G., and N. Williams. 1998. "The Effects of Labor Market Experience, Job Seniority, and Job Mobility on Wage Growth." In *Research in Labor Economics 17*, edited by S. Polachek, 233–276. Stamford, CT and London: JAI Press.

Altonji, J. G., and N. Williams. 2005. "Do Wages Rise with Job Seniority? A Reassessment." *Industrial & Labor Relations Review* 58 (3): 370–397. doi:10.1177/001979390505800304.

Battu, H., C. R. Belfield, and P. J. Sloane. 1999. "Overeducation among Graduates: A Cohort View." *Education Economics* 7 (1): 21–38. doi:10.1080/09645299900000002.

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. doi:10.1086/677233.

Blanco, G., C. A. Flores, and A. Flores-Lagunes. 2013. "Bounds on Average and Quantile Treatment Effects of Job Corps Training on Wages." *Journal of Human Resources* 48: 659–701. doi:10.1353/jhr.2013.0017.

Bracke, P., E. Pattyn, and O. von Dem Knesebeck. 2013. "Overeducation and Depressive Symptoms: Diminishing Mental Health Returns to Education." *Sociology of Health & Illness* 35 (8): 1242–1259. doi:10.1111/1467-9566.12039.

Büchel, F., and A. Mertens. 2004. "Overeducation, Undereducation, and the Theory of Career Mobility." *Applied Economics* 36 (8): 803–816. doi:10.1080/0003684042000229532.

Caliendo, M., and R. Hujer. 2006. "The Microeconometric Estimation of Treatment Effects – An Overview." *Allgemeines Statistisches Archives* 90 (1): 199–215. doi:10.1007/s10182-006-0230-4.

Card, D., J. Kluve, and A. Weber. 2010. "Active Labor Market Policy Evaluations: A Meta-Analysis." *The Economic Journal* 120 (548): F452–F477. doi:10.1111/j.1468-0297.2010.02387.x.

Card, D., J. Kluve, and A. Weber. 2015. "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *IZA Discussion Paper No. 9236*. Bonn, Germany.

Chen, C., P. Smith, and C. Mustard. 2010. "The Prevalence of Over-Qualification and Its Association with Health Status among Occupationally Active New Immigrants to Canada." *Ethnicity & Health* 15 (6): 601–619. doi:10.1080/13557858.2010.502591.

Dauth, C., and J. Lang. 2019. "Can the Unemployed Be Trained to Care for the Elderly? The Effects of Subsidized Training in Elderly Care." *Health Economics* 28 (4): 543–555. doi:10.1002/hec.3863.

Dengler, K. 2019. "Effectiveness of Active Labour Market Programmes on the Job Quality of Welfare Recipients in Germany." *Journal of Social Policy* 48 (4): 807–838. doi:10.1017/S0047279419000114.

Doerr, A., B. Fitzenberger, T. Kruppe, M. Paul, and A. Strittmatter. 2017. "Employment and Earnings Effects of Awarding Training Vouchers in Germany." *Industrial & Labor Relations Review* 70 (3): 767–812. doi:10.1177/0019793916660091.

Dolton, P., and A. Vignoles. 2000. "The Incidence and Effects of Overeducation in the U.K. Graduate Labour Market." *Economics of Education Review* 19 (2): 179–198. doi:10.1016/S0272-7757(97)00036-8.

Dorner, M., J. Heining, P. Jacobebbinghaus, and S. Seth. 2010. "The Sample of Integrated Labour Market Biographies." *Schmollers Jahrbuch* 130 (4): 599–608. doi:10.3790/schm.130.4.599.

Fitzenberger, B., A. Osikominu, and R. Völter. 2008. "Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany." *Annales d'Economie and de Statistique* 91/92: 321–355. doi:10.2307/27917250.

Fitzenberger, B., and R. Völter. 2007. "Long-Run Effects of Training Programs for the Unemployed in East Germany." *Labour Economics* 14 (4): 730–755. doi:10.1016/j.labeco.2007.05.002.

Flores, C. A., and A. Flores-Lagunes. 2009. "Identification and Estimation of Causal Mechanisms and Net Effects of a Treatment under Unconfoundedness." *IZA Discussion Paper No. 4237*. Bonn, Germany.

Frangakis, C. E., and D. B. Rubin. 2002. "Principal Stratification in Causal Inference." *Biometrics* 58 (1): 21–29. doi:10.1111/j.0006-341X.2002.00021.x.

Grunau, P. 2016. "The Impact of Overeducated and Undereducated Workers on Establishment-Level Productivity." *International Journal of Manpower* 37 (2): 372–392. doi:10.1108/IJM-01-2015-0007.

Hainmueller, J. 2012. "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis* 20 (1): 25–46. doi:10.1093/pan/mpr025.

Ho, D. E., K. Imai, G. King, and E. A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15 (3): 199–236. doi:10.1093/pan/mpl013.

IZA, DIW and Infas. 2005. "Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission - Modul 1b: Förderung beruflicher Weiterbildung und Transferleistungen." *BMAS Forschungsbericht (Research report of the Federal Ministry for Labour and Social Affairs).*

Kampelmann, S., and F. Rycx. 2012. "The Impact of Educational Mismatch on Firm Productivity: Evidence from Linked Panel Data." *Economics of Education Review* 31 (6): 918–931. doi:10.1016/j.econedurev.2012.07.003.

Kruppe, T., and J. Lang. 2018. "Labour Market Effects of Retraining for the Unemployed. The Role of Occupations." *Applied Economics* 50 (14): 1578–1600. doi:10.1080/00036846.2017.1368992.

Lechner, M., and B. Melly. 2010. "Partial Identification of Wage Effects of Training Programs." *Working Paper*. Brown University. 2010-8.

Lechner, M., R. Miquel, and C. Wunsch. 2007. "The Curse and Blessing of Training the Unemployed in a Changing Economy: The Case of East Germany after Unification." *German Economic Review* 8 (4): 468–509. doi:10.1111/j.1468-0475.2007.00415.x.

Lechner, M., R. Miquel, and C. Wunsch. 2011. "Long-Run Effects of Public Sector Sponsored Training in West Germany." *Journal of the European Economic Association* 9 (4): 742–784. doi:10.1111/j.1542-4774.2011.01029.x.

Lee, D. S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies* 76 (3): 1071–1102. doi:10.1111/j.1467-937X.2009.00536.x.

Leuven, E., and H. Oosterbeek. 2011. "Overeducation and Mismatch in the Labor Market." In *Handbook of the Economics of Education*, edited by E. A. Hanushek, S. Machin, and L. Woessmann, 283–326. Amsterdam: Elsevier.

McGowan, M. A., and D. Andrews. 2015. "Skill Mismatch and Public Policy in OECD Countries." *Working Paper No. 1210*. OECD Economics Department.

McGuinness, S. 2006. "Overeducation in the Labour Market." *Journal of Economic Surveys* 20 (3): 387–418. doi:10.1111/j.0950-0804.2006.00284.x.

Michael, L., and C. Wunsch. 2013. "Sensitivity of Matching-Based Program Evaluations to the Availability of Control Variables." *Labour Economics* 21: 111–121. doi:10.1016/j.labeco.2013.01.004.

Munch, J. R., and L. Skipper. 2008. "Program Participation, Labor Force Dynamics, and Accepted Wage Rates." In *Modelling and Evaluating Treatment Effects in Econometrics (Advances in Econometrics, Volume 21)*, edited by T. Fomby, T. C. Hill, D. L. Millimet, J. A. Smith, and E. J. Vytlacil, 197–262. UK: Emerald Group Publishing Limited.

Neuman, S., and A. Weiss. 1995. "On the Effects of Schooling Vintage on Experience-Earnings Profiles: Theory and Evidence." *European Economic Review* 39 (5): 943–955. doi:10.1016/0014-2921(94)00019-V.

Nordin, M., I. Persson, and D.-O. Rooth. 2010. "Education–Occupation Mismatch: Is There an Income Penalty?" *Economics of Education Review* 29 (6): 1047–1059. doi:10.1016/j.econedurev.2010.05.005.

OECD. 2011. "Right for the Job: Over-Qualified or Under-Skilled?" Chapter 4 In *OECD Employment Outlook 2011:* 191–233. Paris: OECD Publishing.

Pollmann-Schult, M., and F. Büchel. 2004. "Career Prospects of Overeducated Workers in West Germany." *European Sociological Review* 20 (4): 321–331. doi:10.1093/esr/jch027.

Robst, J. 2007. "Education and Job Match: The Relatedness of College Major and Work." *Economics of Education Review* 26 (4): 397–407. doi:10.1016/j.econedurev.2006.08.003.

Rosen, S. 1975. "Measuring the Obsolescence of Knowledge." In *Education, in Income and Human Behavior*, edited by F. T. Juster, 199–232. New York: Carnegie Foundation for the Advancement of Teaching and National Bureau of Economic Research.

Rosenbaum, P. R., and D. B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55. doi:10.1093/biomet/70.1.41.

Rubin, D. B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5): 688–701. doi:10.1037/h0037350.

Sloane, P. J., H. Battu, and P. T. Seaman. 1999. "Overeducation, Undereducation and the British Labour Market." *Applied Economics* 31 (11): 1437–1453. doi:10.1080/000368499323319.

Verhaest, D., and T. Schatteman. 2010. "Overeducation in the Early Career: An Analysis Using Sequence Techniques." *HUB Research Papers in Economics and Management*.

Zhang, J., D. Rubin, and F. Mealli. 2008. "Evaluating the Effect of Job Training Programs on Wages through Principal Stratification." In *Modelling and Evaluating Treatment Effects in Econometrics (Advances in Econometrics, Volume 21)*, edited by T. Fomby, T. C. Hill, D. L. Millimet, J. A. Smith, and E. J. Vytlacil, 117–145. UK: Emerald Group Publishing Limited.

Appendix

A1. Descriptive statistics

Table A1. Mean values of the pre-treatment covariates for participants and non-participants.

	Participants	Non-participants
Female	0.471	0.330***
Marital status		
Single	0.418	0.427***
Not married, not living alone	0.077	0.076
Single parent	0.119	0.077***
Married	0.386	0.420***
Children	0.435	0.552***
Age	33.23	36.07***
Age groups		
Age 20–24	0.131	0.110***
Age 25–29	0.236	0.169***
Age 30–34	0.211	0.167***
Age 35–39	0.201	0.181***
Age 40–44	0.141	0.170***
Age 45 and older	0.080	0.202***
German	0.908	0.861***
Health problems	0.071	0.101**
Disabled	0.012	0.028**
Last occupation (BIBB major occupational fields)		
Occup. involving extraction/production of raw materials	0.026	0.041***
Manufacturing, processing, repair/maintenance occup.	0.256	0.379***
Occup. in operation and servicing of plants/machinery	0.065	0.048***
Occup. involving sale/marketing of goods	0.096	0.079***
Transport, storage, security occup.	0.141	0.149***
Hotel/restaurant and cleaning occup.	0.109	0.126***
Office and commercial occup.	0.129	0.084***
Technical and scientific occup.	0.037	0.025***
Legal, management and business occup.	0.006	0.006
Occ. in media sciences, humanities, social sciences, art	0.016	0.012***
Health care, social and personal care occup.	0.111	0.047***
Teaching occup.	0.008	0.005***
Position in last job		
Blue-collar worker	0.355	0.422***
Skilled worker	0.180	0.260***
White-collar worker	0.265	0.170***
Part-time worker	0.196	0.144***
Highest vocational degree		
No degree	0.337	0.340
Vocational degree	0.621	0.626
University degree	0.041	0.034***
Employment history 7 years before the beginning of the unemployment spell		
Days in employment	1187.60	927.30***
Days of benefit receipt	367.91	637.18***
Number of spells with benefit receipt	2.15	3.75***
Days in unemployment	376.91	700.02***
Number of unemployment spells	2.30	4.15***
Days in labour market programs	132.36	144.22***
Number of spells with program participation	0.824	1.131***
Days without information	475.58	568.94***
Number of spells without information	0.417	0.548***
Participation in short-term training two years before the unempl. spell	0.227	0.351***
Participation in further training two years before the unempl. spell	0.138	0.152***
Mainly employed …		
One year prior to unemployment	0.592	0.345***
Two years prior to unemployment	0.530	0.345***
Three years prior to unemployment	0.505	0.381***
Four years prior to unemployment	0.191	0.242***
Daily wage last job	44.44	45.82***
Mean daily wage (7 years)	42.74	43.30***
Education		
No school degree	0.042	0.140***
Secondary schooling degree (Hauptschulabschluss, Mittlere Reife)	0.809	0.777***
Secondary schooling degree (Abitur)	0.150	0.083***
Classification of local labour market		
Areas in East Germany with the poorest labour market conditions	0.085	0.060***
Areas in East Germany with poor labour market conditions	0.126	0.135***

(*Continued*)

Table A1. (Continued).

	Participants	Non-participants
Areas mainly in East Germany, high unemployment, some on the border to the west	0.074	0.064***
Areas characterized by big cities and high unemployment	0.128	0.099***
Areas mainly characterized by big cities and moderately high unemployment	0.040	0.069***
Areas with above-average unemployment but moderate dynamics	0.105	0.092***
Areas with average unemployment	0.104	0.103
Areas with below-average unemployment and weak dynamics	0.116	0.114
Centres with a good labour market situation and strong dynamics	0.062	0.074***
Rural areas with a good labour market situation and strong seasonal dynamics	0.031	0.030
Areas with SME structure and a good labour market situation	0.084	0.108***
Areas with the best labour market situation and strong dynamics	0.045	0.052***
Elapsed duration of the unemployment spell in months		
<1	0.183	0.447***
1–2<	0.048	0.108***
2–3<	0.056	0.090***
3–4<	0.055	0.059***
4–5<	0.056	0.046***
5–6<	0.052	0.033***
6–7<	0.051	0.028***
7–8<	0.050	0.023***
8–9<	0.039	0.018***
9–10<	0.039	0.016***
10–11<	0.036	0.012***
11–12<	0.035	0.011***
12–13<	0.029	0.010***
13–24	0.165	0.066***
>24–36	0.061	0.021***
>36	0.043	0.010***
State before beginning of the current unemployment spell		
Employed	0.684	0.458***
Apprentice	0.020	0.005***
No information 1–3 months	0.090	0.179***
No information 4–6 months	0.052	0.113***
No information 7–12 months	0.057	0.099***
No information 13–24 months	0.041	0.071***
No information more than 24 months	0.055	0.076***
Federal state		
Schleswig-Holstein	0.041	0.037***
Hamburg	0.006	0.027***
Lower Saxony	0.112	0.096***
Bremen	0.010	0.009
North-Rhine-Westphalia	0.156	0.187***
Hesse	0.067	0.054***
Rhineland-Palatinate	0.033	0.046***
Baden-Württemberg	0.063	0.089***
Bavaria	0.125	0.123
Saarland	0.011	0.013**
Berlin	0.095	0.062***
Brandenburg	0.076	0.044***
Mecklenburg-Vorpommern	0.047	0.038***
Saxony	0.078	0.072***
Saxony-Anhalt	0.026	0.057***
Thuringia	0.053	0.046***
Regional unempl. rate (dependent civilian labour force)	13.739	13.178***
Educational (mis)match in the last job		
Overeducation	0.053	0.042***
Horizontal match (3-digit level)	0.164	0.263***
Number of observations	27,427	134,094

Source: Integrated Employment Biographies. */**/*** indicate significant differences in the mean values between the treatment and control group at the 10/5/1% levels, respectively.

A2. Robustness

Identifying causal treatment effects using (propensity score) matching as a pre-processing step is sometimes criticized because of the potential problem that balancing certain covariates may lead to a decreased balance of others, hence counteracting the intended bias reduction (e.g. Ho et al. 2007). Moreover, observations are either matched or discarded, implying a reduction of the analysed sample. An alternative solution to achieve covariate balance is entropy balancing, where weights are assigned to every observation. As the approach achieves balance by adjusting the covariate distributions of the control group to those of the treatment group, only the weights for the control group units vary, whereas each treated individual is assigned a weight of 1. While the formation of these weights aims at covariate balance, it also strives to remain close to some uniform base weights to preserve information (Hainmueller 2012). However, a re-estimation based on entropy balancing instead of propensity score matching reveals no substantial differences in the estimated ATETs (cf. Table 4).

Table A2. ATETs of retraining on several outcomes of job (match) quality, by gender and observation point (in months), using entropy balancing.

	Men		Women	
After …	48 months	84 months	48 months	84 months
Gross daily earnings (log)				
ATET	10.105***	9.680***	13.618***	12.370***
SE	(0.473)	(0.503)	(0.474)	(0.503)
Observations	100,614	100,611	54,652	54,605
Overeducation				
ATET	−0.016***	−0.013***	−0.031***	−0.026***
SE	(0.002)	(0.002)	(0.006)	(0.006)
Observations	99,051	88,949	54,070	50,623
Horizontal adequacy (3-digit level)				
ATET	0.167***	0.163***	0.269***	0.241***
SE	(0.005)	(0.005)	(0.006)	(0.008)
Observations	70,236	63,242	40,630	38,164

Source: Integrated Employment Biographies. Coefficients depict the average treatment effects on the treated (ATETs). Standard errors (in parentheses) and number of observations are displayed directly beyond the coefficients. */**/*** indicate the significance of the coefficients at the 10/5/1% levels, respectively.

A3. Estimation approach for local net average treatment effects on the treated (LNATET)

Let Y_i be the outcome (measuring qualification mismatch) of an individual i with treatment status D_i and a (post-treatment) mechanism variable S_i (employment). Both S_i and Y_i depend on D_i, and Y_i additionally depends on S_i. In our case, $D_i = 0$ for non-participants and $D_i = 1$ for retraining participants, and $S_i = 0$ if an individual i is not employed and $S_i = 1$ if s/he is employed. Within the potential outcome framework (Rubin 1974), the average treatment effect on the treated (ATET) is

$$\text{ATET} = E[Y(D=1) - Y(D=0)|D=1]$$

.

Flores and Flores-Lagunes (2009) decompose this total treatment effect into a mechanism effect and a net effect:

$$\text{ATET} = E[Y(D=1, S(D=1)) - Y(D=1, S(D=0))|D=1]$$

$$+E[Y(D=1, S(D=0)) - Y(D=0, S(D=0))|D=1]$$

The first term is the mechanism effect and is caused by a change in S due to a change in D, which in our case represents an indirect effect on job match quality as retraining affects participants' employment prospects. The second term is the net effect and equals the part of the treatment effect of D on Y holding S fixed at S (D = 0). In our case, this is the effect of retraining on job match quality, given no effect of retraining on employment. Flores and Flores-Lagunes (2009) employ the concept of principal stratification (Frangakis and Rubin 2002) to define a set of comparable individuals based on the potential values of the post-treatment variable S (here: (likelihood of) employment). One of the two approaches they suggest enables us to estimate the local net average treatment effect on the treated (LNATET) by focusing on a specific subpopulation for which D does not affect S, and thus $S_i(D_i = 1) = S_i(D_i = 0)$.

A4. LNATETs applying within-occupation comparisons

Table A4. LNATETs of retraining with within-occupation comparisons and on several outcomes of job (match) quality, by gender and observation point (in months).

	Men			Women		
After …	48 months		84 months	48 months		84 months
Gross daily earnings (in Euros)						
LNATET	1.197		0.369	1.824***		1.893***
SE	(0.688)		(0.784)	(0.743)		(0.804)
Post-matching mean bias (pre-match.)		5.8 (14.8)			5.8 (12.6)	
Observations	6,998		6,859	5,677		5,614
Overeducation						
LNATET	−0.007**		−0.012***	−0.022***		−0.018***
SE	(0.003)		(0.003)	(0.005)		(0.005)
Post-matching mean bias (pre-match.)		6.2 (15.3)			6.0 (12.9)	
Observations	6,608		6,066	5,506		5,085
Horizontal adequacy (*3-digit level*)						
LNATET	0.069***		0.073***	0.084***		0.103***
SE	(0.011)		(0.011)	(0.014)		(0.017)
Post-matching mean bias (pre-match.)		6.7 (14.7)			6.1 (12.8)	
Observations	5,353		5,090	4,772		4,593

Source: Integrated Employment Biographies. Coefficients depict the local net average treatment effects on the treated (LNATETs). Standard errors (in parentheses) and number of observations are displayed directly beyond the coefficients. */**/*** indicate the significance of the coefficients at the 10/5/1% levels, respectively.