New Evidence on the Determinants of Firm-based Training*

Susanne Steffes† and Arne Jonas Warnke‡

November, 2016

Preliminary draft!

Abstract

In this paper we analyze firm- and worker-level observable and unobservable heterogeneity in participation in job-related training. We use a novel panel dataset which links firm and worker surveys, and contains detailed information on the incidence, duration, initiative and funding of training. Using multilevel methods, we analyze selection processes for participation in training and consider the complementarity of investments in training made by the firm and worker. Our results point towards significant differences in the determinants of participation in courses financed by the firm, or training co-financed by the worker. Another original result from our study concerns the significance of job characteristics in determining individuals' participation in training. The information included in the analysis on workers' job characteristics, has played a significant role in the identification of important determinants. Firm-financed and worker (co-)financed training seem to be neither substitutes nor complements of one another. In addition, we show that whilst unobserved firm-heterogeneity is of little significance, worker characteristics remain important determinants of participation in training, even after controlling for variables.

JEL Classification: I24, J24, M53

Keywords: Human Capital, Training, Linked-Employer-Employee Data (LEE), Decomposition, Unobserved Heterogeneity

*Special thanks go to Daniel Dietz, Bernd Fitzenberger, Katja Görlitz, François Laisney, Henrik Stryhn, Thomas Zwick and seminar participants at ZEW Mannheim and University of Freiburg, University of Würzburg, EEA-ESEM (2014), Colloquium on Personnel Economics (2014), Workshop on Economics of Education (2014), DFG Conference on the German Labour Market in a Globalized World (2015), The Economics of Vocational Education and Training (2015).

†ZEW Centre for European Economic Research Mannheim. E-mail: steffes@zew.de.

‡ZEW Centre for European Economic Research Mannheim. E-mail: warnke@zew.de.

1 Introduction

In contrast to schooling decisions, investments in on-the-job training (formal or informal) are determined jointly by firms and workers. Given that firms and workers are not alike, training investments most probably depend on characteristics of both parties. In addition, training investments are driven by the job context. Certain occupations and tasks require that workers regularly learn new, or preserve existing skills. In other words, the heterogeneity of firms and the diversity of jobs and workers within these firms are crucial elements of the training decision.

The bilateral nature of investment decisions makes it difficult to empirically disentangle the determinants of training (Lynch and Black, 1998). Many empirical studies have investigated determinants of work-related training and showed that both individual and firm characteristics are important (this literature is summarized by Arulampalam et al., 2004; Bassanini et al., 2005; Asplund, 2005). Very few studies, however, address both dimensions simultaneously by looking at multiple workers in several firms at the same time.[1] Despite the importance given to the interdependency of training investment decisions in the theoretical literature, empirical evidence rarely acknowledges such considerations.

In this paper, we raise the question as to whether, and to what extent workers' participation in training differs across firms (between-firm heterogeneity) or between employees in the same firm (within-firm heterogeneity). Between-firm heterogeneity considers the extent to which participation in training differs between two comparable workers who work for two different firms. Within-firm heterogeneity looks how much training rates differ between co-workers in a given firm. We focus on neither the specific determinants of participation in training, nor the expected returns. Our interest rather lies in the general interplay between firms and workers when it comes to job-related training. Therefore, we are the first who simultaneously investigate the extent to which individual training rates vary between different firms and workers, and suggest reasons for these heterogeneities in training, including for example, job characteristics. We distinguish formal training on the basis of the party initiating it and the party financing it - two other vital aspects addressed in the theoretical literature.

Several studies in the literature on training such as Lynch and Black (1998), or Zwick (2005) for Germany, have used firm-level data with aggregated information about, for example, the proportion of highly skilled workers to show the interrelation between workforce composition and training rates at the firm-level. These studies use information about the share of workers who participate in training but they do not reveal the dynamics of workers' participation in training over time (given the establishment dynamics). We know neither how often certain

[1] Most studies look either at individuals only using household data, or at firms using establishment data (sometimes aggregating individual information), or at the workforce of single firms. Whether the focus is on the firm or the worker dimension depends usually on the aggregation level of the data at hand.

workers participate in training, nor anything about the workers' investments. Furthermore, a change in a firm's average training rate may be caused by either an adjustment of the firm policy or due to changes in the composition of workers (e.g. within the group of highly-skilled workers). Aggregated data does not enable us to identify which mechanism has been decisive for individual participation in training.[2] Here we use matched employer-employee data where individual workers give information about their training history, which combines the advantages of disaggregated data on individual employees and aggregated firm data. We therefore have detailed information, not only about the participation in training of an individual worker, but also about the participation of their co-workers in a firm (and changing compositions of the workforce).

We are aware of only two studies which use linked-employer-employee data with information about individual participation in training: Frazis et al. (1999) is an early exception which uses a US matched dataset to analyze firm and worker characteristics correlated with provision of, and participation in training. The authors can mostly validate results of previous studies using unmatched data. An interesting finding concerns firm size. The positive relationship between firm size and training seen in the dataset is primarily driven by the presence of fringe benefits and innovative work practices. Almeida Santos and Mumford (2004) use Australian linked-employer-employee data of almost 1,500 workplaces and 14,000 employees. The authors investigate firm-specific variables such as (voluntary) turnover, percentage of unionized members or wage compression. Neither study provides estimates of the significance of firms for workers' participation in training (or the heterogeneity between them), nor assess the interdependence of firm and worker co-financed training.

We are not aware of any such study that simultaneously observes several firms and workers being employed in these firms in Germany. Instead, a number of studies conducted for Germany use worker-level data in order to investigate determinants of work-related training. Using the German Socio-Economic Panel (SOEP) from the late 1980's, Pischke (2001) distinguishes between formal training in general and training which is (co-)sponsored by the employer.[3] In addition to the worker-level characteristics, this data provides some basic information about the employer such as the firm size and sector. The results point towards higher rates of participation amongst men, more highly educated, and younger workers. Firm and job characteristics are also important. Pischke finds that workers in larger firms, in the public sector and in managerial or high-skilled white collar jobs, receive more training. The results are very similar for both forms of training. Using an updated version of the same data, Grund and Martin (2012)

[2]It is also questionable to what extent a manager knows the exact number of (unique) workers who attended a training in a given year when asked about it.

[3]Employer sponsored training is defined in this study as "if the training either took place during work hours or the employer is named as the organizer of the training or the employer bore at least some of the monetary cost of training".

investigate formal training patterns of workers in the private sector between 1989 and 2008. Training rates have increased over this time period in Germany. In contrast to Pischke's findings, this study does not identify any differences in average training rates between men and women. Training profiles are inversely U-shaped with respect to age and U-shaped with tenure. Participation rates are once again higher for more highly educated workers and for those in more complex professions. Grund and Martin confirm previous results with respect to firm-size and show pronounced differences in training between different sectors. It remains unclear, however, whether firm-size or sector differences are proxies for other unobserved characteristics or attributes which lead to additional differences in training between workers. Fitzenberger and Muehler (2014) use personnel records from a single firm to analyze gender differences in the age-profile of firm-financed training. The use of single-firm data offers the advantage that the confounding impact due to male and female sorting into certain firms can be excluded. The results indicate differences in age-related patterns in training amongst male and female employees. Women recieve less firm-sponsored training than men, particularly when they are in their thirties. Female workers participate relatively more often in training at higher ages but they do not catch up with male training rates.

First, we shed new light on the interplay between the firm and worker dimensions. Estimating the explanatory power of different sets of observables make it easier for the empirical work to test appropriate hypotheses. A worker's characteristics and the market conditions of firms depend upon one another as certain workers will choose to sort into certain firms. This makes it difficult to rely on either models which focus on the firm-level only or models which focus exclusively on worker heterogeneity.

Secondly, this does not only apply to the empirical work on training determinants but also to the studies analyzing the effects of training. These studies try to measure the benefits of training such as higher wages on the worker-level or productivity growth on the firm-level. This literature has to tackle the issue of unobserved heterogeneity as training investments are typically not allocated on the basis of characteristics which are observable for researchers. Some establishments might have higher training rates due to the introduction of new technologies while certain workers are more bound to the firm and receive accordingly more training. Knowledge about the relative importance of firm and worker heterogeneity (observable but especially unobservable) is crucial if an appropriate identification strategy to estimate causal returns is to be looked for.

The results of this paper show that in comparison to worker heterogeneity, firm heterogeneity is far less important. This means that differences between co-workers in a given firm have a much greater impact on determining investments in training than aggregated differences between the workforce of two firms do. In addition, we can explain a large share of aggregated training differences between two firms if we control for firm, worker and job characteristics. In contrast,

most of the worker-level variance remains unexplained even after adding a large set of control variables. These results are not only found for job-related training in general but also hold if we distinguish between purely firm-financed and worker co-financed training. Unsurprisingly, worker heterogeneity plays a more significant role in determining training investments where courses are co-financed by workers.

The paper is organized as follows. In the next section, we will present our research questions (Section 2). Section 3 and 4 will present the available linked-employer-employee panel dataset and the methods we will use to make the most of the available structure. Section 5 will show the results of our analyzes while Section A discusses further topics and gives robustness checks. In Section 6 we will provide a conclusion, detailing what has been learnt from the study and will outline implications for policy recommendations.

2 Research Questions

Our first research question uses detailed linked employer-employee data from multiple firms, along with information about the participation of workers in training, in order to calculate the overall importance of firms for workers' attendance of training. We contrast this with the heterogeneity in participation in training between workers within a firm to see whether average training rates differ between co-workers. We first add information on observable characteristic of workers, then firms, and finally on job attributes. This helps us to explain why average training rates differ between firms, or which groups of workers have high training propensities. Furthermore, we give information about the relative importance of unobserved variables, i.e. remaining and unexplained heterogeneity between firms and between workers. We compare results for training in general to analyses based on firm-financed and worker co-financed training courses only.

Research Question 1 *How important are firms and workers for individuals' participation in training? Are there differences between firm-financed and worker co-financed training?*

It is difficult to derive unambiguous predictions about the association between certain characteristics and training investments from theory. The recent empirical literature has found a large number of relevant predictors for investments in training. However, the majority of these studies either ignore the firm dimension or do not account (properly) for the compositional differences of the workforce between firms. Assume, for instance, that older workers operate in labor market conditions where firms exert more monopsony power (because these workers are less mobile). According to the theory outlined above, the propensity of a firm to provide training should generally be higher in such markets. Accordingly, if we do not take a firm's monopsony

power into account (e.g. a measure of wage compression), we should find that rates of investment in training for older workers are confounded by the omission of the market condition. Thus, a

Research Question 2 *Which are the relevant firm-, worker-, and job-level predictors for participation in training?*

Previous literature has often neglected the fact that decisions to participate in different forms of training are not independent of one another. An early exception is Royalty (1996) who uses a multinomial probit model to investigate participation in on-the-job and off-the-job training as well as no training. We offer a new approach to this matter.

3 Data

We use the German linked employer-employee dataset WeLL: *Berufliche Weiterbildung als Bestandteil Lebenslangen Lernens* (Further training as a part of lifelong learning, see Huber and Schmucker, 2012) which comprises four waves of a worker survey conducted between 2007 and 2010. Data has been collected by the Research Data Centre of the Federal Employment Agency at the Institute for Employment Research (FDZ). The sample of survey participants was selected in two steps. First, a random sample of 149 establishments in the manufacturing and service sector was drawn from establishments which participated in the IAB Establishment Panel in 2005. This is an annual employer survey of approximately 16,000 businesses (see Kölling, 2000). This sample is stratified to establishments with 100 to 2,000 employees from three West-German and two East-German states. Second, an employee sample was randomly drawn from all employees who were covered by the social security system and were employed on December 31st 2005 in one of the 149 establishments. The first wave consisted of 5,819 interviews, the subsequent waves of 4,560, 4,667 and 3,720 interviews respectively.[4] It was possible to link the survey data to the administrative records of each employee (provided that he or she granted the relevant permission). Records include information on wages per day and the duration of employment. Survey data was also linked to the IAB Establishment Panel (Schmucker et al., 2014).

WeLL is a particularly well-suited dataset for analyzing training effects as it provides comprehensive information on workers' recent participation in training. The questionnaire includes (retrospectively asked) questions about participation in formal training courses such as the duration, content, initiative, financing and whether the training overlapped with leisure time. Survey participants are asked about the number of job-related training courses they had attended since

[4]This includes only those interviews in which participants agreed that their survey data can be linked to social security.

a reference date (which was the last interview for panel participants and a maximum of two years). Basic information about participation in training (start and end date) are available for all training courses while detailed information has been collected for up to three most recent training courses per wave. In addition, the survey includes questions about socio-demographic characteristics, personality traits and satisfaction with one's life and work. Moreover, respondents were asked about their health, labor attachment and career aspirations.

The administrative data provides information on individuals' employment histories in regard to earnings, employment status, unemployment status and respective benefits going back to 1975.[5] This information is available, not only for participants in the WeLL panel, but for all workers in a firm. This gives us information about the size of an establishment, its average wage and wage growth rates without much measurement error.

In addition, it provides us with information regarding firm heterogeneity relevant to participation in training. This information includes sector, legal form, age, and whether a workers' council was present.

We limit the analyses to individuals still working in one of the 149 establishments from which they were drawn in 2006.[6] We also excluded observations which were missing information in relevant variables such as training or socio-demographic characteristics. In addition, we disregarded all workers who attended more than three training courses as information was not adequately detailed.[7] Finally, we limited the analyses to workers aged 21-64 and excluded apprentices and workers in partial retirement. We focus on full-time or regular part-time workers and disregarded those with a usual working time of less than 15 hours a week. We were subsequently left with 5,785 workers and 12,560 observations. The mean number of unique WeLL-participants per establishment is 39 (median 24) with an interquartile range of 15 to 45.

3.1 Variables

The main variable of interest is the participation in formal training courses. In the first wave for instance, individuals where asked "Did you participate from January 1, 2006 up to now in any job-related seminars or training courses?". If they responded positively to this question they were then asked "Who launched the initiative for the participation in training?" and were given the option of selecting the answers "my own initiative", "by order of the firm/supervisor", "mostly upon advice of my firm", "mandatory occupational training" and "mostly upon the advice of someone else". We have excluded mandatory training such as obligatory first-aid

[5]Periods of employment which are not subject to social security, such as periods of self-employment or employment in the civil service, are not observed. For further information see Jacobebbinghaus and Seth, 2007.

[6]For an analyses of the effects of training on labor mobility using WeLL see Dietz and Zwick, 2016.

[7]This concerns ca. 4.0% of all workers. Sensitivity analyses regarding participation in any training course are shown in Table 10 and discussed in Section A.

courses, fire safety training and equal opportunities courses from analysis. This reduced the number of training courses by around 15.1% in the first wave. We also disregarded training courses completed on advice from a third party (ca. 1%). This gives us an average training participation rate of 44.7% per wave which is comparable to other sources such as the Adult Education Survey (e.g. Autorengruppe Bildungsberichterstattung, 2012). Training participation rates vary statistically significantly by wave and by establishment (according to a F-test). The distribution of number of training courses per worker and wave is shown in Table 1. 27.9% of of workers participate in a given waven for example in one training course.

Table 1: Number of Training Courses Attended per Wave

Observations	Percent	Number of training courses
6,942	53.0%	no course attended
3,652	27.9%	participated in one course
1,481	11.3%	two courses attended
485	3.7%	participated in three courses
529	4.0%	attended more than three courses

Note: Number of job-related training courses (excl. mandatory training) per worker and wave. Workers attended more than three courses were excluded for the main analyses but are considered in Section A.

Almost all training courses (more than 85%) included in the survey are rather general in nature in the sense that skills learnt are completely or predominantly transferable to other firms. Therefore, we do not distinguish between general and firm-specific training, but consider who bears the costs for the course.[8] To this end, we distinguish between firm-financed and worker-co financed training courses by identifying whether a worker participated in training exclusively during his or her working hours, or whether training also took place during workers' leisure time. 60% of all training courses took place during working hours only. 20% occurred either partially or only in leisure time.[9] If we take participants who attended no training courses into account, we find that 24.4% of workers participate in firm-financed training only, whilst 14.9% participate in training overlapping with leisure time, and 5.3% of workers participate in a given wave in both forms of training (e.g. attend at least two courses where one is completely in working time and the other overlaps with leisure time).

Summary statistics are presented in Table 5.

[8]Participations in training have been asked whether the skills learnt in the course are transferable to other firms. A training course is here defined as *general* if trained skills are completely or predominantly transferable to other firms while *truly firm-specific* training implies that skills are not at all transferrable.

[9]We do not consider a monetary investment by the worker because only 16.2% of all training courses are not completely paid by the firm and often workers only contribute a small sum. This is also in line with conversations with practitioners: Firms often pay for training even if this is not directly relevant for the tasks involved in a job itself, but for which attendance is required at the weekend or in the evening.

We look at observable individual as well as establishment characteristics with possible relevance for participation in training. We distinguish between establishment, worker, and job attributes. Given that the time period about which workers are retrospectively asked differs between and within waves (depending on the month of the (last) interview) we add a variable to each specification which captures the length of the time period in months.

Our choice of characteristics in the empirical model was influenced by the training literature (e.g. Asplund, 2005). Establishment characteristics include the state in which the establishment operates (these states differ in public training subsidy programs), information about establishment size and age, whether the establishment is a public enterprise and whether it operates in the service sector. We use information about the presence of a works council to capture possible training differences due to labor institutions.[10] From social security data of all employees of the establishments (survey participants or not), we derive the median wage and a linear growth rate for establishments' median wages from 2000-2010. Furthermore, we add a measure for the wage-compression of an establishment which, according to the model provided by Acemoglu and Pischke (1999), might induce firms to invest in general training: This measure is the standard deviation of establishments' log daily wages of full-time workers in 2006.[11]

As worker characteristics we consider sex, age, age squared, foreign citizenship and living alone. In addition, we look at levels of education (vocational qualification or tertiary degree), experience of unemployment in the last five years (derived from social security data) and a proxy for a worker's labor attachment and health status. Information on labor attachment was collected via a question which asked workers how likely they consider it that they will be still be employed in a year's time (on a scale ranging from 0 to 10). Health status was assessed subjectively on a Likert scale ranging from very good to very bad (1-5).

Job characteristics included working hours, temporary work contract, tenure and an indicator for recent recruiting. The latter is included because we expect higher training rates for new workers who lack firm-specific skills. We are also interested in job requirements since we assume that workers performing complex tasks have greater training requirements. We proxy job requirements by job-tasks (see Figure 1), occupational information (Figure 2) and managerial responsibilities. Job-tasks measure actitivities workers frequently perform in the job.[12]

[10] In a study for Germany, Pfeifer (2015) finds significant positive effects of works councils but not of unions on participation in training, so we do not include information about the presence of unions.

[11] Median wages are calculated for all employees while the wage compression refers to full-time workers only; we impute censored wages using a tobit model.

[12] Information regarding the job-tasks are comparable to information used for instance, by Spitz Oener (2006) who investigates technology-driven changes in the skill demand in Germany in recent decades. Respondents in the survey were asked whether they perform a task "frequently", "seldom" or "never". We consider here only job-tasks which are frequently performed.

4 Methods

We aim to analyze establishment and worker determinants of participation in training simultaneously for which we use multilevel generalized linear models (for a textbook introduction see Gelman and Hill, 2007)[13].

These models allow us to simultaneously consider firm and worker heterogeneity in participation in training. We combine the advantages of individual level data (rich information, not only about the individual worker, but also about workers in the same firm) with attributes of the firms. Firms' average training rates are not, for example, confounded by the changing composition of their staff. We can estimate the relevance of firms' and workers' observable and unobservable heterogeneity with regard to training. Furthermore, we are able to perform all analyses which can be done by either disaggregated household data or aggregated firm level data because standard linear least squares estimates can be considered a special case of multilevel models.

In Section A, we show that results are quite similar if either the incidence or the intensity of training is considered (training yes/no or the number of courses a worker attended). We therefore focus primarily on the incidence of training – or whether a given worker participates in any formal training at a certain period.

The model gives us an estimate of the total variation in participation in training between establishments and between workers within the same establishment. The training differences can be either observable or unobservable for us. We can explain observable variation by adding establishment, worker or job characteristics in the following specification while unobservable variation remains the same (in the form of a positive variance of establishment and/or worker random effects). Since our outcome is binary, we use a random effects logit model.

$$\Pr\big[\,\text{Training=1}_{it}\,\big|\,\alpha_{J(i,t)},\,\theta_i,\,\tau_t\,\big] = \text{logit}^{-1}\big(\alpha^{(1)}_{J(i,t)} + \theta^{(1)}_i + \tau_t\big) \tag{1}$$

Here and in the following specifications, we examine whether worker i in establishment $J(i,t)$ participates in training at period t. $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ are establishment respective worker random effects and $\boldsymbol{T}$ year dummies (we do not assume normality regarding the time dimension because there are at maximum four observations for each worker).[14] We approximate the log-likelihood, which has no closed form solution, by Gauss-Hermite quadrature (see StataCorp, 2013). We use the variance of establishment and worker random effects to assess the variation attributable to each dimension. We identify establishment random effects by observing multiple workers per establishment (within and across waves), whilst it is the panel structure of the data which

[13] Other terms are for example mixed or two-way random effects models.

[14] We denote here the random effects of Equation 1 with a superscript "(1)" to allow for the variance decomposition a distinction to the random effects of other equations.

enables us to identify worker random effects. We assume that $\alpha \sim \mathcal{N}(0, \sigma_\alpha^2)$ and $\theta \sim \mathcal{N}(0, \sigma_\theta^2)$ (and mutual independence of the random effects). We must assume that workers do not self-select into certain firms because of favorable unobserved characteristics.[15] The independence assumption is necessary because we observe only four workers who appear in two different WeLL establishments (for a formal discussion see Abowd and Kramarz, 1999). In Section A.2 we discuss a further specification using establishment fixed effects.

We first estimate a *basic* model without any control variables other than for time effects. We cluster standard errors on the establishment level.[16] In a second step, we add establishment characteristics $\boldsymbol{Z}$ to our equation to see which attributes of establishments are associated with higher or lower training rates of the workers. Furthermore, we investigate how well establishment characteristics explain inter-firm variation in training. All workers of the same establishment share the same establishment attributes. Therefore, by construction, establishment characteristics $\boldsymbol{Z}$ can only explain variation between establishments and not between workers within the same firm.

$$\Pr\big[\text{Training=1}_{it} \,\big|\, \alpha_{J(i,t)},\, \theta_i,\, \tau_t,\, \boldsymbol{Z}_{J(i,t),t}\,\big] = \text{logit}^{-1}\big(\alpha^{(2)}_{J(i,t)} + \theta^{(2)}_i + \tau_t + \boldsymbol{Z}_{J(i,t),t}\delta\big) \qquad (2)$$

We then once again exclude observable establishment information and look at the association between participation in training and workers' observable attributes. On the one hand, differences in worker characteristics $\boldsymbol{X}$ can explain training differences between workers in the same establishment in terms of their participation in training. This may be the case, for example, where higher ability correlates with increased levels of training. On the other hand, it might be that certain workers tend to work at certain establishments - high-ability workers for example might gravitate towards a certain employer, this might also explain why average differences between workers within the same establishment, for example if worker's ability is correlated with training. On the other hand, if certain workers such as high-ability workers tend to cluster at certain establishments, worker characteristics would explain why average training rates differ between firms.

$$\Pr\big[\text{Training=1}_{it} \,\big|\, \alpha_{J(i,t)},\, \theta_i,\, \tau_t,\, \boldsymbol{X}_{it}\,\big] = \text{logit}^{-1}\big(\alpha^{(3)}_{J(i,t)} + \theta^{(3)}_i + \tau_t + \boldsymbol{X}_{it}\beta\big) \qquad (3)$$

Finally, we look simultaneously at firms and workers to see whether associations identified between establishment characteristics and training for example, are driven by differences between

[15] Our (necessary) assumption that establishment and worker random effects are mutually independent (plus independent of the other covariates) excludes for example job mobility related to the idiosyncratic training provision of an establishment.

[16] This should give us a conservative estimate of the standard errors since there are (almost) no mover between establishments in our sample.

the workforces of establishments.

$$\Pr\big[\text{Training=1}_{it} \,\big|\, \alpha_{J(i,t)},\, \theta_i,\, \tau_t,\, \boldsymbol{X}_{it},\, \boldsymbol{Z}_{J(i,t),t}\big] = \text{logit}^{-1}\big(\alpha^{(4)}_{J(i,t)} + \theta^{(4)}_i + \tau_t + \boldsymbol{X}_{it}\beta + \boldsymbol{Z}_{J(i,t),t}\,\delta\big) \tag{4}$$

In a final step, we add information about job characteristics $\boldsymbol{W}$ such as tenure or job requirements (our "full model").

$$\begin{aligned}\Pr\big[\text{Training=1}_{it} \,\big|\, \alpha_{J(i,t)},\, \theta_i,\, \tau_t,\, \boldsymbol{X}_{it},\, \boldsymbol{Z}_{J(i,t),t},\, \boldsymbol{W}_{J(i,t),t}\big] = \\ = \text{logit}^{-1}\big(\alpha^{(5)}_{J(i,t)} + \theta^{(5)}_i + \tau_t + \boldsymbol{X}_{it}\beta + \boldsymbol{Z}_{J(i,t),t}\,\delta + \boldsymbol{W}_{J(i,t),t}\,\gamma\big)\end{aligned} \tag{5}$$

Research Question 2 investigates predictors for participation in training. We will answer this question by looking at observable predictors of participation in training $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{W}$. The relative importance of single variables can be easily assessed by looking at either the size of their coefficients as we have standardized all right-hand-side variables including binaries, or by comparing t-statistics in order to incorporate uncertainty.

In order to investigate the role of firm and worker heterogeneity (Research Question 1), we look at the variance of worker and firm random effects before and after adding observable information. The reduction of the variance of the random effects in the full model (e.g. after controlling for worker and firm characteristics $\text{Var}\big[\alpha^5_{J(i,t)}\big]$, $\text{Var}\big[\theta^5_k\big]$) compared to the basic model ($\text{Var}\big[\alpha^1_{J(i,t)}\big]$, $\text{Var}\big[\theta^1_k\big]$) gives an estimate of the relevance of observable vs. unobservable heterogeneity.[17]

To investigate the importance of establishments and workers with regard to training, we use a variance decomposition approach. The variance decomposition for mixed models without random coefficient is easily derived as the covariance in participation in training between two workers of the same establishment is simply $\text{Cov}\big(\text{Train}_{i_1t}\,,\,\text{Train}_{i_2t}\big) = \sigma^2_{\alpha^{(1)}}$. We are therefore able to partition the total variance $\sigma^2 = \sigma^2_\alpha + \sigma^2_\theta + \frac{\Pi^2}{3}$ into the three variance terms - the variance of the establishment random effects $\sigma^2_{\alpha^{(1)}}$, the variance of the worker random effects $\sigma^2_{\theta^{(1)}}$ and the variance of the latent error which is assumed equal to the variance of the logistic distribution $\frac{\Pi^2}{3}$.[18] This gives us an estimate of the relevance of inter-establishment training differences compared to heterogeneity of participation in training between workers of the same establishment over time (this measure is usually called the intra-class correlation). Equation (6) shows a measure of the overall or unconditional relevance of establishments for workers' participation in training, whilst Equation (7) shows the relevance of establishments after controlling for establishment, worker and job characteristics. Sections A.2 and A.3 discuss other specifications

[17] We abstract here from the complication that observable heterogeneity is time-varying while random effects are by definition constant over time.

[18] This is a standard assumption in the literature, see for example Hox (2010). In Section A.3, we show results using a linear approximation as suggested in Goldstein et al. (2002).

including different approaches to the variance decomposition.

$$\frac{\sigma^2_{\alpha^1}}{\sigma^2_{\alpha^1} + \sigma^2_{\theta^1} + \frac{\Pi^2}{3}} \quad (6) \qquad \frac{\sigma^2_{\alpha^5}}{\sigma^2_{\alpha^5} + \sigma^2_{\theta^5} + \frac{\Pi^2}{3}} \quad (7)$$

Finally, we investigate the relationship of firm- and worker-financed training. Workers might participate in both firm. and worker co-financed training, or indeed, in neither of them. We first analyze whether establishments which provide a great deal of training are also those in which many workers participate in training which they themselves co-finance. We hereby run the specifications from Equation (1) and Equation (5) for participating in training entirely during working hours, and for training which overlaps with workers' leisure time. We then correlate the establishment random effects $\text{Corr}\big(\alpha^{(1)}_{\text{Firm financed}}, \alpha^{(1)}_{\text{Worker co-financed}}\big)$ in order to shed light on the relationship between aggregated firm- and worker co-financed training rates in a first step.

For the individual worker perspective, we differentiate between not participating in training, attending firm-financed, worker co-financed training courses or participation in both forms of training in a given wave. We use a multinomial logit model.[19]

$$\Pr\big[\text{Training}=c_{it} \,\big|\, \tau_t,\ \boldsymbol{X}_{it},\ \boldsymbol{Z}_{J(i,t),t},\ \boldsymbol{W}_{J(i,t),t}\big] = \frac{\exp^{\boldsymbol{T}_t\tau_c+\boldsymbol{X}_{it}\beta_c+\boldsymbol{Z}_{J(i,t),t}\delta_c+\boldsymbol{W}_{J(i,t),t}\gamma_c}}{\sum_{i=1}^{4}\exp^{\boldsymbol{T}_t\tau_c+\boldsymbol{X}_{it}\beta_c+\boldsymbol{Z}_{J(i,t),t}\delta_c+\boldsymbol{W}_{J(i,t),t}\gamma_c}} \quad (8)$$

For computational reasons, we do not include establishment or individual random effects here. Standard errors are clustered on the establishment level to account for intra-establishment correlation. We refrain from using multiway clustering for reasons of simplicity since there is almost no mobility of workers between establishments and clustering on the highest level gives conservative estimates.

5 Results

5.1 Research Question 1

In the following, we present results gained for our research questions introduced above in Section 2, namely, how important firm and worker heterogeneity is for the individual participation in training. And whether there are differences between firm-financed and worker co-financed training. To this end, we first compare the intra-class correlation for establishments and workers

[19]Unfortunately, we do not have variables available which vary with the alternative training outcomes. This means that we cannot relax the independence of irrelevant alternatives assumption by using a multinomial probit model without relying on the correct specification of the functional form. As discussed in Keane (1992), results for the multinomial probit without exclusion restriction are "extremely tenuous". A nested logit model has been numerically instable so far, so we do not present its results here.

by looking at the relative importance of firms (inter-firm correlation), and workers (correlation of workers within the same establishment) within the overall variance in participation in training. Table 2 shows the intra-class correlation for establishments and workers for all job-related training courses (Equation 6 and 7). Our estimate of the unconditional intra-establishment correlation (basic model) attributes 12% of the total variation in participation in training to differences between establishments (across-firm heterogeneity). The importance of workers is more than twice as large (28%) (within-firm heterogeneity), while according to the latent-response formulation, most of the variance in training remains unexplained (slightly above 60%). According to a likelihood-ratio test, both establishment and worker random effects are highly significant.

Table 2: Intra-class Correlation (job-related training in general)

Level	**Basic Model** Equation 1	**Firm Char.** Equation 2	**Worker Char.** Equation 3	**Firm+Worker** Equation 4	**+ Job Char.** Equation 5
Establishment	11.9%	6.4%	9.5%	5.6%	1.4%
Worker	27.6%	29.4%	25.7%	26.9%	19.3%
Latent Error	60.5%	65.6%	65.0%	69.0%	79.6%

Note: This table reports the estimated intra-class correlation (Equation 6 to 7) for different specification of a multilevel logit model for participation in job-related training courses (see Table 7). Establishment and worker random effects are statistically significant in all specifications according to a likelihood-ratio test.

In the next step, we take the fact that businesses and workers differ in observable characteristics into account. We thereby compare intra-class of the basic models with the full model and other specifications. Comparison of the first column (basic model) and the last column (full model with establishment, worker, and job characteristics) of the first row in Table 2, indicates that almost 90% of the across-establishment correlation can be explained by observables. The comparison of the basic model with other specifications (columns 2-4) shows that the variation in average training rates between establishments is mostly reduced by establishment and job characteristics, and to a lesser extent by worker characteristics. This indicates that establishment-variation in terms of workers' participation in training courses can primarily be explained by firm-level or job-level differences, but to a lesser extent by the composition of the workforce itself.[20]

Next, we look at training differences between workers within the same establishment. Here, we can only explain 30% of the variance in worker random effects by including a large set of observables ($1 - 19.3/27.6 = 0.30$). This is in stark contrast to the previous results for the establishment level where little unexplained heterogeneity remained.[21] For the explanation of

[20]In our final specification, including establishment, worker and job characteristics, the intra-establishment correlation drops to meager 1.4%. However, establishment random effects variation remains still significant according to a likelihood-ratio test (as do worker random effects).

[21]Of course the number of observations per establishment is much larger than that is the case for workers.

training incidence differences between workers we find that job characteristics are more powerful predictors than worker characteristics (establishment characteristics cannot explain training heterogeneity of workers within an establishment by construction). Note, however, that the set of observed job characteristics is also larger than the set of observed worker characteristics (36 vs. 10).

The intra-class correlation (and the variation in random effects) is an abstract concept which is not always immediately intuitive. The relevance of the variation in worker random effects can be expressed with a simple statistical example. For this example, we look at 1,090 workers who participate in all four waves of the survey (18.8% of all survey participants). If we ignore the establishment dimension here, we can expect slightly more than 9% of workers to never participate in training. If training were to be randomly allocated, we could expect 4% of workers to participate in training in all waves, with a probability of 45%.[22] The actual numbers are strikingly greater, 26.6% do not participate in training at any point, whilst 9.1% of workers participate in training in all four waves. Here we see a much larger variation than would be expected from random allocation.

Table 2 reveals three main findings. Firstly, the largest share of the variance in participation in training can be attributed to neither firm differences, nor to worker differences. Secondly, whilst most of the establishment differences can be explained by observable characteristics, this does not translate to the worker level. The majority of within-establishment worker heterogeneity remains unexplained (the intra-class correlation goes down from 27.6% to 19.2%). Unexplained heterogeneity is therefore a major issue. Finally, Table 2 indicates that job characteristics are a significant source of training heterogeneity on both the firm and worker level.

Before we discuss the findings for purely firm-financed and co-financed training, we aim to illustrate the magnitude of differences in training participation. We wish to show how the variance of establishment random effects is reflected in average training rates between establishments.[23] One must bear in mind that all variables including binaries are standardized such that we can read the estimated training probability for the worker-firm match simply by considering the intercept (and of course, the logit link function). The estimated training probability for the worker-firm match in the basic model is $\frac{1}{1+\exp(-0.38)} \approx 40.6\%$. The variance of establishment random effects is $\sigma^2_{\theta^{(1)}} = 0.65$.[24] This means that an establishment with an unobserved random effect which is one standard deviation below the average has an unconditional training probability of just $\frac{1}{1+\exp(-0.38-0.65)} \approx 26.3\%$. However, looking at the full model (last column of Table 7), we see that $\sigma^2_{\theta^{(5)}} = 0.05$. Tis means that differences in training rates between establishments

[22] $(1-0.45)^4 \approx 0.0915$ and $0.45^4 \approx 0.041$.

[23] Random effects variance components are available in Table 7 in the second last block (θ^2_{Firm} / θ^2_{Worker}).

[24] The estimated training probability for the worker-firm match is somewhat lower than the (overall) average training rate of 45% due to Jensen's inequality.

can, to a large extent, be explained by observables. This is a result which we already have seen in the intra-class correlations. The analogous difference translates into a range of 2.4 percentage points ($\left[\mu - \sigma_\theta\,,\ \mu + \sigma_\theta\right] = \left[41.1\%\,,\ 43.5\%\right]$).

We then go on to provide a decomposition for the variance for training which is either fully financed by the firm, or at least partially financed by the worker (part two of Research Question 1, see Table 3). This analysis delivers very similar results to those seen following the decomposition of the sum of all training.[25] Again, we find that the worker dimension is more important than the establishment dimension, and we can explain a much larger share of the variance on the establishment level with reference to observables than we can on the worker level. Most interesting, perhaps, is the finding that co-financed training exhibits a larger share of worker level variance compared to training which is purely financed by the firm. The share on the establishment level meanwhile, is $\approx$11%, making it similar across all types on training.

Table 3: Variance Partitioning (firm-financed / worker co-financed training)

	Firm-financed Training		**Worker co-financed Training**	
Level	**Null**	**Occ**	**Null**	**Occ**
Establishment	11.2%	2.3%	10.8%	3.7%
Worker	23.4%	18.0%	34.8%	30.4%

Note: Table 3 reports the estimated intra-class correlation (Equation 6 and 7) for different specification of a multilevel logit model for participation in job-related training courses either completely in working-time of partly in leisure time (see Table 8 and 9). Establishment and worker random effects are statistically significant in all specifications according to a likelihood-ratio test.

5.2 Research Question 2

In order to answer Research Question 2, we look at the regression coefficients for the different specifications in Table 7. With regard to establishment characteristics, we do not find differences between smaller and larger firms in our sample.[26] The same applies to establishment age. We find higher training rates for workers in public enterprises and, in line with the theoretical literature (Acemoglu and Pischke, 1999), in establishments with a compressed wage structure. However, given that adding job characteristics results in a loss of statistical significance, training advantages in these firms seem to be driven by differences in the tasks involved in specific jobs (see Column 5 in Table 7). We find consistently higher training rates amongst workers in the service sector, in high wage establishments and for workers in establishments which experienced a positive linear wage trend. Having said this, the magnitude of this coefficient

[25] Here, we only show the comparison between the basic and the full model. Results of the other specifications can be delivered on request.

[26] Note, that our sample consists of establishments ranging from 100 to 200 employees.

falls considerably if worker and job characteristics are included. Once again, in comparison to worker characteristics, job characteristics can be considered more important.

Various worker characteristics, in particular workers' level of education, are found to be strong predictors of their likelihood of participating in training. (see Table 7, Column III and IV). However, after adding job characteristics, the majority of coefficients fall and once again become insignificant. Less well educated workers and individuals with health issues exhibit yet lower rates of participation in training. Despite this, after the addition of job characteristics, highly educated workers and those with a high degree of labor force attachment no longer show an advantage in participation in training. The connection between higher levels of education, particularly tertiary education, and increased on-the-job training has been identified in the literature on a number of occasions (e.g. Carneiro and Heckman, 2003). We have been able to confirm such observations. In the full model, however, the coefficient of tertiary education reduces almost to zero, thereby becoming statistically insignificant. This means that a high level of education correlates with a significant number of job characteristics such as with professional occupations. The causes of training advantages remain unclear. Our results do at least point strongly to the finding that the dimension of job characteristics plays a considerable role in the assignment of training courses within businesses. In addition, we find an interesting result regarding the selection of certain groups of workers. The tendency of women and older workers to participate in training is lower only after job characteristics are added. This is an indication that these groups perform more training-intensive tasks but receive less training than co-workers in similar jobs. Lower training attendance is particularly pronounced amongst older workers (see Zwick, 2015). This has been identified by a large negative coefficient for age squared.

As shown by the variance decomposition approach, job characteristics are important predictors for participation in training. We find considerably higher training rates for workers who have managerial responsibilities. In contrast to other studies (Grund and Martin, 2012), we do not find lower training rates for part-time workers. Moreover, participation in training varies depending on job content, that is the specific occupation and the tasks involved in a job, as shown in Figure 1 (see also Görlitz and Tamm, 2016). Here we find the largest difference between manufacturing and knowledge-based jobs. If we translate the tasks into a classification of routine versus non-routine and analytic, complex, interactive, or manual jobs (Spitz Oener, 2006), we find that it is in particular non-routine, complex activities, such as analytic or interactive tasks which are associated with higher rates of training. There are, however, no negative associations for routine tasks. Workers who perform manufacturing tasks for example, very seldom participate in training. If we look at occupations, we find marked differences in participation in training. This is shown in Figure 2. The reference occupations are *other services* which consists mostly of low-skilled white-collar jobs such as security contractors and cleaners. We

see that many occupations show higher training rates (only assembling and food processing jobs have actual insignificantly lower participation rates) than *other services*. Training is particularly common in certain female-dominated occupations such as health and education professions in which more ca. two-thirds of all workers are female. In addition, rates of training are high amongst engineers and electricians, as well as in certain white-collar jobs including merchants and in a number of professions such as consultancy and accountancy. Training differences between workers who perform certain tasks or work in certain occupations are significant, as illustrated by the large (standardized) coefficients.

We now turn to the question as to whether findings for Research Question 2 depend on the type of training concerned; training which takes place during working hours only, and training which overlaps in leisure time. The answer is: partly yes. Here, we compare results for the total participation in training with results of training taking place during working time (see Table 8) and training overlapping with leisure time (see Table 9). We have found some interesting results with regard to the variables measuring firm characteristics. While both proxies for the economic situation (the median wage and its linear trend) are positively related with training courses taking place in working time, this is not the case for training overlapping with leisure time. Indeed, establishments' median wage is negatively correlated with workers' participation in co-financed training in the full model including job characteristics (at the 10% significance level). It is to be expected that training which is fully financed by the firm is more sensitive to the economic situation of the business. Moreover, the negative association between the median wage and worker co-financed training might point to a possible substitution effect. Workers make up for the lower investments made by the firm by themselves contributing more to finance the human capital investment.

A comparison of the worker characteristics also shows very interesting results. First, in the models excluding job characteristics signs and significance levels of almost all worker characteristics are similar across all types of training, including workers' level of education. One exception to this is female workers who tend to participate less in training courses occurring in working time, and more in training courses which overlap with leisure time. Furthermore, the shapes of the age-training profiles differ for both forms of training. Whilst the age structure was significantly u-shaped for training co-financed by the worker, there is no such relationship between age and training in working time. Again, we see that most coefficients decrease in magnitude and become insignificant after job characteristics are taken into account. Still, some differences are striking: after the inclusion of job characteristics, only female workers and those with health issues exhibit significantly lower rates of training in working time. That means, on the one hand, that the assignment of purely firm-financed training is strongly determined by the worker's specific job. On the other hand, it would seem that firms investment less in women

and workers with health problems.[27]

The picture is quite different if we look at the full model of training courses overlapping with leisure time. Firstly, women participate significantly more often in such training courses. Secondly, differences between groups of workers depending on their levels of education remain significant but coefficients are not very large. In addition, it is interesting that after including controls for the job cell, age differences become even larger. Finally, other characteristics which might be proxies for worker productivity, labor attachment or health status for example, become insignificant if job heterogeneity is taken into account. Thus, we can conclude that job and worker attributes are important determinants of participation in training courses for which workers bear some of the costs in that training overlaps with their leisure time. The economic rationale of a worker, such as the expected amortization period of training investments, seems to play an important role in this context. This is in line with theoretical considerations we presented in Section 2.

Finally, we look at the findings for job characteristics. In line with theoretical considerations, workers with limited contracts and those in part-time participate less in training courses taking place during working time. This does not apply to training which overlaps with leisure time. Here we even find a positive correlation with part-time work which might again suggest that substitution effects play a role here (of part-time working women). Workers with managerial responsibilities display a greater probability of participating in both types of training. As shown in Figure 1, there are some minor differences between performing certain job-tasks such as conducting negotiations and training (strong positive relationship with firm-financed training but virtually no for worker co-financed training). In general, we find very similar results for both types of training.

In contrast, we find large occupational differences between training in working time and training overlapping with leisure time. Figure 2 shows, on the one hand, that higher training rates in most occupations (compared to *other services*) are driven by worker co-financed training. Only four occupational groups, engineers, merchants, clerical professionals and teaching professions are provided with significant more training. On the other hand, workers in many other jobs are trained more often in their leisure time. This is particularly true for health-related professions.

In Table 6, we arrange the coefficients from the previous specifications in Tables 7 to 9 according to the absolute value t-statistic. This is a simple but common measure of relative importance which takes the magnitude of the association with training as well as the uncertainty into ac-

[27]Note, we do not argue that firms discriminate against women and sick persons. We just find that both groups attend less in (fully) firm provided training, even after controlling for various characteristics. The actual reasons therefore are difficult to establish, e.g. whether this is voluntary or not. We investigate the possible reasons for lower female job-related training rates more in detail in Steffes and Warnke (2016) where we also show that this is probably not due to higher turnover rates, see also Royalty (1996).

count.[28] Job characteristics which constitute somewhat more than 50% of all variables (36 out of 64) dominate Table 6. For our overall training regression, only the wage level of an establishment is among the ten predictors with the highest absolute t-statistics. If we compare firm financed and worker co-financed training, only three tasks are significant in both specifications: computer usage, organizing and advising resp. counselling.

5.3 Interrelation of firm-financed and co-financed training

Substitution between firm-financed and co-financed training might exist if workers who do not have access to purely firm-financed training compensate this by investing in training (at least partly) themselves. A complementary effect is, however, also likely as the mechanism which increases the incentive to invest in training might be the same for both parties. The interrelation of firm-financed and co-financed training therefore remains an empirical question. Even if we interpret firm-financed training courses as *firm-specific* and worker co-financed training as *general* (or a mixture of firm-specific and general training), we cannot derive clear hypotheses regarding the sign of the interrelation. We would expect that establishment characteristics are important predictors for firm-specific training investments. It is plausible to assume that certain establishments benefit more from such investments than others (e.g. those in which technology evolves rapidly). In imperfect labor markets, this should similarly apply to firms' investments in general training; and should translate to workers' investments. Individual attributes such as age, education, and labor attachment can also be expected to be relevant for the investment made by the business, just as such factors are relevant for the individual worker. However, if the expectation about lifetime returns differs between establishment and worker, we would expect a substation of both types of training.

To investigate the possible interrelation between firm-financed and worker co-financed training, we start with an aggregate analysis and run the previous specification (Equation 1 and Equation 5) for training completely in working time and training overlapping with leisure time. To examine whether workers in establishments which provide a lot of training during working hours also often co-finance training, we correlate the 149 establishment random effects of both regression before and after conditioning on observables. The results indicate a small complementarity between both forms of training on the establishment level before adding establishment, worker and job characteristics. After conditioning on observable heterogeneity, this positive association becomes very small and close to zero, indicating that worker and job heterogeneity confounds this picture.

In order to gain a more detailed view on this, we also ran a multinomial logit regression, which is presented in Table 11. We distinguish not only between workers who participate in training

[28] We refrain from using more advanced statistics because of computational aspects.

entirely during working-time or in training which overlaps with leisure time, but we also observe workers who participate in either of the two (or neither). The reference group in Table 11 are workers who do not participate in any form of training. This analysis reveals some interesting further insights about training.

We start with the establishment level. Establishments in the service sector and high wage establishments provide more training which takes place entirely during working time. Workers in these businesses, however, do not themselves invest more in training than workers in other firms do. Therefore, we do not see any differences regarding worker co-financed training or participation in both forms of training. For firm-financed training, we find some differences regarding the firm size: firms with 200 to 500 employees (the middle category in our dataset) seem to provide more training than smaller or larger firms. We do not have an explanation for this and it is the only specification for which we find a correlation between training investments and firm size. Unsurprisingly, workers in rapidly growing businesses, as measured in rising wages, seem to participate in more training both provided by the firm but simultaneously co-financed by the worker.

With regard to worker heterogeneity, we once again find that women tend to participate in no training, rather than in firm-financed training. Royalty (1996) found a similar result in the United States in the 1990's using a multinomial probit approach.[29] If we differentiate between women who participate only in worker co-financed training (but not in firm-financed training as well), we find higher training rates for women in worker co-financed training (not depicted here). However, women seem to participate more in both forms of training which points to more heterogeneity in participation in training for women (here p-value: 0.14; a fact which we further investigate in Steffes and Warnke, 2016). In terms of workers' level of education, there are only few low-skilled workers in our dataset but they seem to participate less in all forms of training particularly in multiple training courses.[30] Older workers do not necessarily participate less in firm-financed training but they participate much less in training during leisure time, as indicated by a significant negative squared term for age and as seen above.

Looking at job characteristics, we once again see that workers with managerial responsibilities participate more in any form of training whilst workers with a limited contract or part-time workers tend more often not to participate in training which takes place completely in working hours. This suggests that the part-time disadvantage is not driven by our definition of firm-financed and worker co-financed training.

[29]Royalty (1996) assumes independent error for the multinomial probit approach. This has no advantage compared to the multinomial logit approach with respect to the independence of irrelevant alternatives assumptions (see also Footnote 19).

[30]The extremely large coefficient for low-skilled workers participating in both types of training can be explained by sample size issues.

6 Conclusions

A large empirical literature addresses determinants and returns to training, few studies have looked simultaneously at multiple firms and their individual workers. The reason for this is clear. Such an approach requires precise matched-employer-employee data including detailed training information on the individual level which has as yet been barely available. The lack of such empirical studies is clearly a shortcoming in this field as theory predicts that training investment decisions of firms and workers are clearly interlinked.

In this study, we have shed new light on observable and unobservable determinants of job-related training using a new matched dataset of 149 establishments and 7,000 workers. We first analyzed the incidence, and later also the intensity of workers participation in training within and across firms. We investigated which observable characteristics such as firm size, education level or the tasks involved in a given job can be used to predict the participation in job-related training courses. We distinguished between training courses fully financed by firms and those co-financed by workers and took into account the role of mandatory occupational training. Furthermore, we considered how similar workers within an establishment behave with respect to participation in training. Second, we identified the relevance of unobserved heterogeneity amongst firms and workers in order to build a complete picture of the determinants of training investments. Our study thus provides econometricians analyzing returns to training with important indications about the source of endogeneity. Finally, we examined the link between firm- and worker-financed training in order to discover the link between them.

Our empirical findings have shown that that whilst job-related training is unequally distributed across firms, it is even more unequally distributed across workers within firms. Interestingly, most of the variation in training between firms could be explained by observable characteristics such as the sector, the economic performance or the amount of wage compression within the firm. After conditioning for these variables, training differences between firms became almost negligible. This may in particular be due to the inclusion of job characteristics which exhibit the highest degree of observable heterogeneity.

Results were quite different however, at the worker level, which were found to be twice as significant as results at the establishment level in determining the individual's likelihood of participating in training. The majority of the training differences between workers within the same establishment could not be explained by observable attributes of either the worker or the job itself. Interestingly, socio-demographic characteristics had only a small predictive power for training. Whilst job attributes do seem to be important, the reasons for the majority of variation in training participation between workers within the same firm remain unexplained. Unsurprisingly, unobservable heterogeneity between co-workers was a more significant determinant more for worker co-financed training courses (overlapping with leisure time) than for training courses

taking place completely within working hours.

All in all, our findings are important for econometricians wishing to measure *causal* returns to training. Unobserved heterogeneity seems to be a significant cause for differences between workers in terms of whether or not they choose to participate in training. This does not hold, however for establishment differences (or at least to a much lesser extent).

A further interesting finding described in this paper is the considerable importance of job characteristic such as the specific occupation or the tasks involved in a job, in determining workers' participation in training. Many determinants used in the existing literature, both at the firm- and at the worker-level, became insignificant or, at least, decreased in magnitude after job attributes had been controlled for. This was true for example, for the training advantage of workers in the public sector or those who hold a tertiary degree. This finding is extremely significant as many previous studies which analyzed the determinants of training did not include information on job characteristics. These variables explain a substantial proportion of the unobserved heterogeneity between individuals and firms with similar observable characteristics.

Future research should aim to acquire better matched employer-employee data which allows further examination of the assumption that workers and firms are independent when it comes to job-related training. This requires a sufficient number of workers switching between establishments. Furthermore, a more detailed analysis of the interrelation of firm-financed and worker co-financed training, using plausible instrumental variables, could provide valuable information about as to whether firms harness workers' willingness to invest in training. Results presented here show that some workers such as managers participate more in all forms of training, whilst female workers for example, participate less in firm-financed training but more in training that they themselves co-finance.

References

Abowd, J. M., Kramarz, F., 1999. Econometric Analyses of Linked Employer-Employee Data. Labour Economics 6 (1), 53–74.

Acemoglu, D., 1997. Training and Innovation in an Imperfect Labour Market. The Review of Economic Studies 64 (3), 445–464.

Acemoglu, D., Pischke, J.-S., 1998. Why Do Firms Train? Theory and Evidence. The Quarterly Journal of Economics 113 (1), 79–119.

Acemoglu, D., Pischke, J.-S., 1999. Beyond Becker: Training in Imperfect Labour Markets. The Economic Journal 109 (453), 112–142.

Acemoglu, D., Pischke, J.-S., 1999. The Structure of Wages and Investment in General Training. Journal of Political Economy 107 (3), 539–572.

Almeida Santos, F., Mumford, K. A., 2004. Employee Training in Australia: Evidence From AWIRS*. Economic Record 80 (s1), 53–64.

Amemiya, T., 1977. Some Theorems in the Linear Probability Model. International Economic Review, 645–650.

Arulampalam, W., Booth, A. L., Bryan, M. L., 2004. Training in Europe. Journal of the European Economic Association 2 (2-3), 346–360.

Asplund, R., 2005. The Provision and Effects of Company Training: A Brief Review of the Literature. Nordic Journal of Political Economy 31, 47–73.

Autorengruppe Bildungsberichterstattung, 2012. Bildung in Deutschland 2012: Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf. W. Bertelsmann Verlag, Bielefeld.

Backes-Gellner, U., Oswald, Y., Tuor Sartore, S. N., 2011. Part-Time Work and Employer-Provided Training: Boon to Women and Bane to Men? Kyklos 67 (4), 463–481.

Barron, J. M., Black, D. A., Loe, 1993. Gender Differences in Training, Capital, and Wages. The Journal of Human Resources 28 (2), 343–364.

Bassanini, A., Booth, A. L., Brunello, G., De Paola, M., Leuven, E., 2005. Workplace Training in Europe. IZA Discussion Papers 1640 (1640).

Becker, G. S., 1962. Investment in Human Capital: A Theoretical Analysis. The Journal of Political Economy 70 (5), 9–49.

Bishop, J. H., 1996. What We Know About Employer-Provided Training: A Review of Literature. CAHRS Working Paper #96-09.

Carneiro, P., Heckman, J. J., 2003. Human Capital Policy. In: Heckman, J. J., Krueger, A. (Eds.), Inequality in America: What Role for Human Capital Policies. MIT Press.

Chang, C., Wang, Y., 1996. Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revisited. Journal of Labor Economics, 505–519.

Dietz, D., Zwick, T., 2016. The Retention Effect of Training: Portability, Visibility, and Credibility. ZEW Discussion Papers (16-011).

Fitzenberger, B., Muehler, G., 2014. Dips and Floors in Workplace Training: Using Personnel Records to Estimate Gender Differences. Scottish Journal of Political Economy 62 (4), 325–429.

Frazis, H., Gittleman, M., Joyce, M., 1999. Correlates of Training: An Analysis Using Both Employer and Employee Characteristics.

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press.

Gibbons, R., Waldman, M., 2006. Enriching a Theory of Wage and Promotion Dynamics inside Firms. Journal of Labor Economics 24 (1), 59–108.

Goldstein, H., Browne, W., Rasbash, J., 2002. Partitioning Variation in Multilevel Models. Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences 1 (4), 223–231.

Görlitz, K., Tamm, M., 2016. Revisiting the Complementarity between Education and Training – The Role of Job Tasks and Firm Effects. Education Economics 24 (3), 261–279.

Grund, C., Martin, J., 2012. Determinants of Further Training-Evidence for Germany. The International Journal of Human Resource Management 23 (17), 3536–3558.

Hox, J., 2010. Multilevel Analysis: Techniques and Applications, 2nd Edition. Routledge Academic.

Huber, M., Schmucker, A., 2012. Panel WeLL: Arbeitnehmerbefragung für das Projekt Berufliche Weiterbildung als Bestandteil Lebenslangen Lernens Dokumentation für die Originaldaten Wellen 1-4. FDZ Datenreport. Documentation on Labour Market Data (03/2012 DE).

Jacobebbinghaus, P., Seth, S., 2007. The German Integrated Employment Biographies Sample IEBS. Schmollers Jahrbuch 127 (2), 335–342.

Jost, P., Kräkel, M., 2008. Human Capital Investments in Asymmetric Corporate Tournaments. Journal of Economics and Business 60 (4), 312–331.

Katz, E., Ziderman, A., 1990. Investment in General Training: The Role of Information and Labour Mobility.

Keane, M. P., 1992. A Note on Identification in the Multinomial Probit Model. Journal of Business & Economic Statistics 10 (2), 193–200.

Kölling, A., 2000. The IAB-Establishment Panel. Schmollers Jahrbuch 120 (2), 291–300.

Lazear, E. P., 2009. Firm-Specific Human Capital: A Skill-Weights Approach. Journal of Political Economy 117 (5), 914–940.

Lazear, E. P., Rosen, S., 1990. Male-Female Wage Differentials in Job Ladders. Journal of Labor Economics, 106–123.

Lynch, L. M., Black, S. E., 1998. Beyond the Incidence of Training: Evidence From a National Employers' Survey.

Paglin, M., Rufolo, A. M., 1990. Heterogeneous Human Capital, Occupational Choice, and Male-Female Earnings Differences. Journal of Labor Economics, 123–144.

Pfeifer, C., 2015. Intra-Firm Wage Compression and Coverage of Training Costs Evidence From Linked Employer-Employee Data. ILR Review.

Pischke, J.-S., 2001. Continuous Training in Germany. Journal of Population Economics 14 (3), 523–548.

Polachek, S. W., 1981. Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure. The Review of Economics and Statistics, 60–69.

Prendergast, C., 1993. The Role of Promotion in Inducing Specific Human Capital Acquisition. The Quarterly Journal of Economics 108 (2), 523–534.

Rabe Hesketh, S., Skrondal, A., 2012. Multilevel and Longitudinal Modelling Using Stata, third edition Edition. STATA Press.

Royalty, A. B., 1996. The Effects of Job Turnover on the Training of Men and Women. Industrial & Labor Relations Review 49 (3), 506–521.

Schmucker, A., Seth, S., Eberle, J., 2014. WeLL-Befragungsdaten verknüpft mit administrativen Daten des IAB:(WELL-ADIAB) 1975-2011. Tech. rep., Institut für Arbeitsmarkt-und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany].

Spitz Oener, A., 2006. Technical Change, Job Tasks, and Rising Educational Demands: Looking Outside the Wage Structure. Journal of Labor Economics 24 (2), 235–270.

StataCorp, L., 2013. Stata Multilevel Mixed-Effects Reference Manual. College Station, TX: StataCorp LP.

Steffes, S., Warnke, A. J., 2016. Gender Differences in Wages and Training. Working Paper.

Stevens, M., 1994. A Theoretical Model of On-The-Job Training with Imperfect Competition. Oxford Economic Papers, 537–562.

Stryhn, H., Sanchez, J., Morley, P., Booker, C., Dohoo, I., 2006. Interpretation of Variance Parameters in Multilevel Poisson Regression Models. In: Proceedings of the 11th International Symposium on Veterinary Epidemiology and Economics.

Zwick, T., 2005. Continuing Vocational Training Forms and Establishment Productivity in Germany. German Economic Review 6 (2), 155–184.

Zwick, T., 2015. Training Older Employees: What is Effective? International Journal of Manpower 36 (2), 136–150.

Appendix

A Further Analyses

The results we have presented here are remarkably stable, even where different sets of variables or different samples are added or removed. The interpretation of our results is also quite similar whether we look at the incidence of training (Section 4-5) or the intensity of training (shown in Section A.1). Our results are robust to other specifications as shown in Section A.2 and A.3.

A.1 Intensity of Training

In this section, we will illustrate that the basic results are similar if we take into account the number of training courses a worker participated in. As discussed in Section 3, we have so far restricted our analyses to workers who participated in no more than three training courses (see Table 1). For overall participation in training, we also include workers who took part in more than three training courses (the maximum per wave is eighteen).[31] We use a multilevel Poisson regression for the analyses of training intensity (for an introduction see Hox, 2010). Again, we standardize all coefficients.

$$\begin{aligned} \mu_{it} = \mathrm{E}\big[y_{it} \,\big|\, \alpha_{J(i)},\, \theta_i,\, \tau_t,\, \boldsymbol{X}_{it},\, \boldsymbol{Z}_{J(i),t},\, \boldsymbol{W}_{J(i),t}\,\big] = \\ = \exp(\alpha_{J(i)}) \cdot \exp\theta_i \cdot \exp\tau_t + \boldsymbol{X}_{it}\beta + \boldsymbol{Z}_{J(i),t}\,\delta + \boldsymbol{W}_{J(i),t}\,\gamma \end{aligned} \tag{9}$$

For the consistency of this estimator it is sufficient to specify the correct conditional mean function (the random-effects have to be uncorrelated with the other characteristics, but the correct distributional assumption is not required, see Rabe Hesketh and Skrondal, 2012). The Poisson multilevel model is more flexible than a standard Poisson regression by allowing for certain forms of overdispersion (the variance depends on the variation of the random effects and it is not automatically equal to the mean) (Rabe Hesketh and Skrondal, 2012).

Unfortunately, there is no latent-response formulation for the Poisson model available which would enable us to use a simple variance partitioning similar to the logit model. We can, however, still consider a simple form of the intra-class correlation as ratio of the establishment variation compared to the total variance (establishment plus worker, see Stryhn et al., 2006). We compare our new results to the results from Equation (10) excluding the term for the variance of the logistic distribution.

$$\frac{\sigma_\theta^2}{\sigma_\alpha^2 + \sigma_\theta^2} \tag{10}$$

[31] The results here are very similar if we look at the intensity for up to three training courses.

A.2 Decomposition Based On Establishment Fixed-Effects Estimation

In the following, we show that our basic results hold if we use a logit model including a dummy indicator for each establishment (fixed-effects) and (only) worker random-effects. We do not include worker fixed-effects because there are too few workers who switch from one WeLL establishment to another, thereby facilitating the separation of firm and worker fixed-effects.[32] We exclude observable establishment characteristics because the majority of characteristics are time-invariant. We restrict this analyses to those workers in establishments with at least 25 interviews, this leaves us with 12,143 workers in 125 establishments.[33]

The results regarding observable worker and job characteristics are similar to those shown in Table 7.[34] As already seen in Section 5, women, low-skilled and older workers and those with health issues show lower rates of participation in training. The lower training rates for workers with recent unemployment experience, for part-timers and for high-tenure workers becomes significantly negative while the training advantage for recently hired workers becomes significantly positive. In the decomposition, we find almost similar results for the worker random effect variation ($\sigma^1_\theta \approx 1.39$ and $\sigma^5_\theta \approx 0.64$), while the variance of the coefficients of the establishment indicators is considerably larger and remains more important after introducing worker and job characteristics ($\sqrt{\text{Var}(\beta_1^{\text{FE}_1}, \ldots, \beta_{125}^{\text{FE}_1})} \approx 0.98$ in the basic model $\sqrt{\text{Var}(\beta_1^{\text{FE}_5}, \ldots, \beta_{125}^{\text{FE}_5})} \approx 0.58$). The larger variance for the fixed-effects is not surprising, this is down to the fact that random effects components were weighted between the fixed-effects estimates for each establishment and the average over all establishments according to the sample size of an establishment (see Gelman and Hill, 2007). In our view, the lower reduction in the variance at the establishment level is primarily a result of problems associated with the small sample size. If we re-run the same regressions for all 149 establishments, we see that the variance of the establishment indicators doubles (in both specifications). Nonetheless, this could indicate that the role of unobservable establishment heterogeneity is not as low as indicated by the multilevel model.

A.3 Alternative Decomposition: Linear Approximation

Goldstein et al. (2002) suggest a linear approximation if probabilities are sufficiently far away from 0 or 1. Here, we treat participation in training as if it were normally distributed and use the estimated variation of the residuals for the variance decomposition. We therefore estimate a multilevel model of participation in training on the same variables as in Section 3 with an

[32] For the assumptions required to separate identify firm- and worker fixed effects, see Abowd and Kramarz (1999).

[33] Establishment averages estimated with fixed-effects are very noisy for establishments with fewer interviews and artificially inflates the importance of the firm dimension in a variance partitioning.

[34] These results are available upon request.

identity link function. This model is in general not consistent (Amemiya, 1977) but the results are quite similar, so we are confident that we can use it as a cross-check for our logistic model. We obtain the following decomposition for the basic and full model including establishment, worker and job characteristics (Equation 1 and 5):

Table 4: Intra-class Correlation Using a Linear Approximation (job-related training in general)

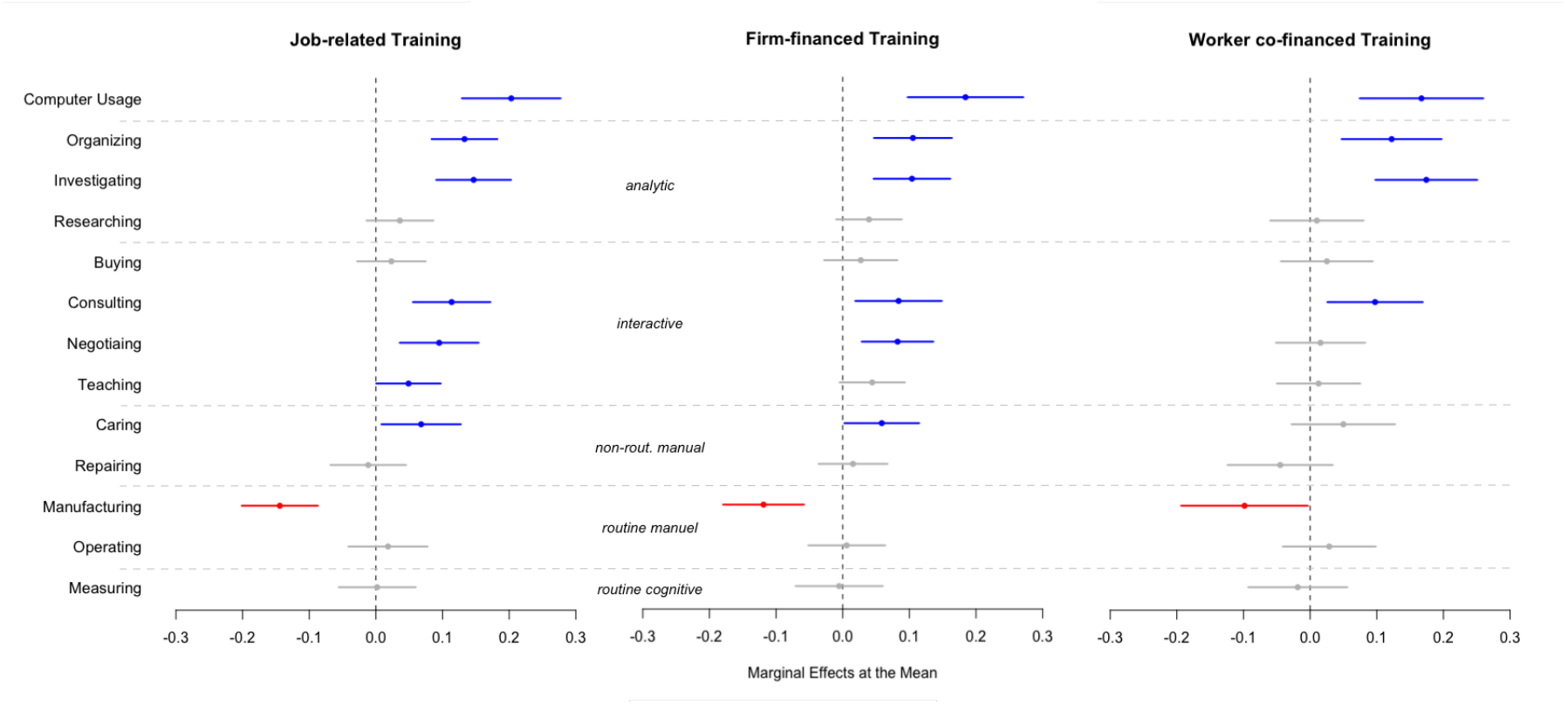
Level	**Null**	**Occ**
Establishment	7.7%	0.7%
Worker	20.2%	13.1%
Residual	72.2%	86.2%

Note: This table reports the estimated intra-class correlation (Equation 6 and 7) for a multi-level model of participation in training with an identity link function.

Unsurprisingly, the results of the variance decomposition differ somewhat from the logit model in relative magnitude because of a different estimate of the residual variation. In general, the anlyses allow a similar interpretation; we are able to explain most of the variation between establishments (ca. 90%), but only one third of the variance between workers within the same establishment.

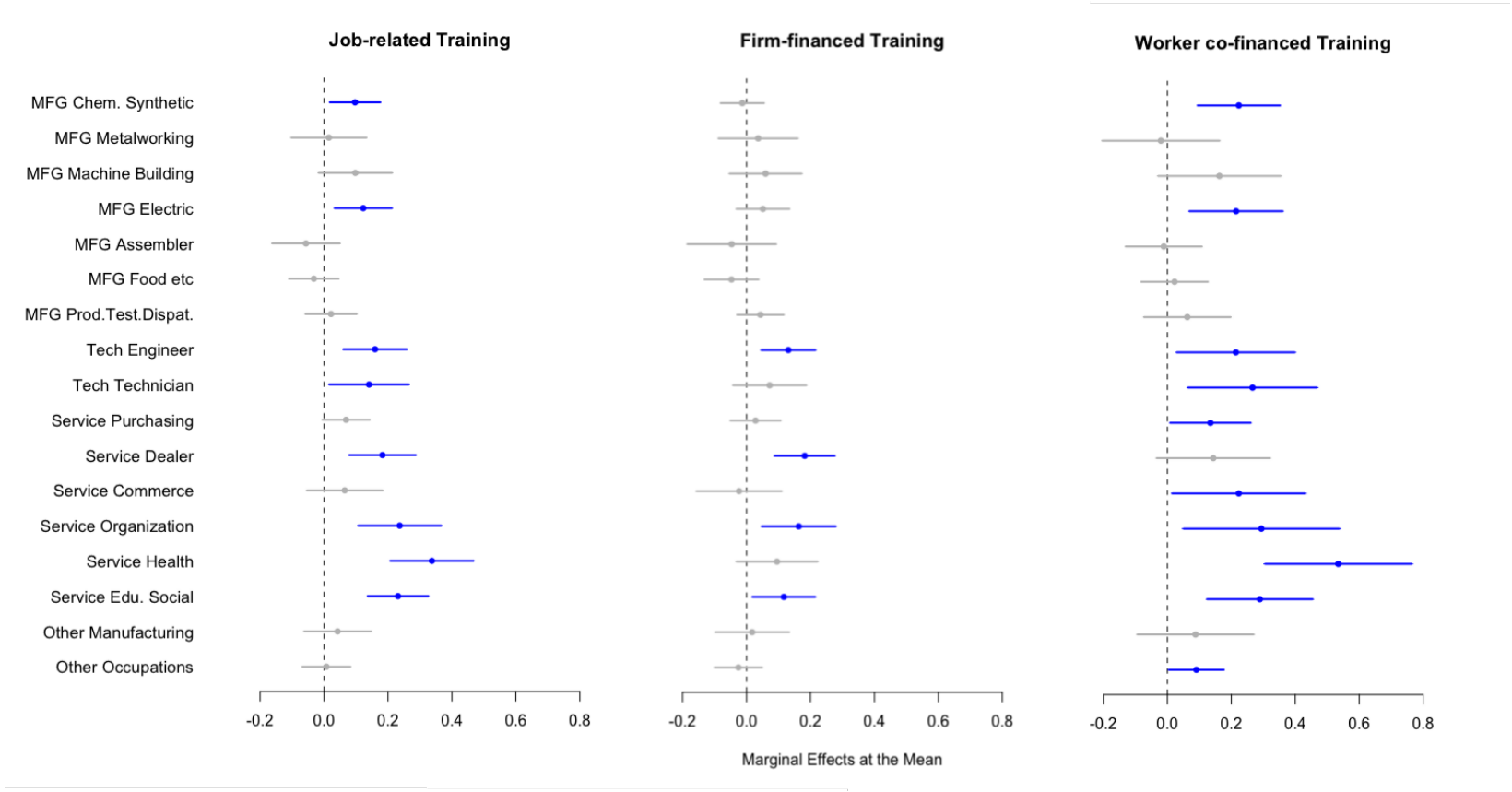
B Figures and Tables

Figure 1: Task Performance and Training Participation



Coefficients are standardized and taken from regression in Column 5 in Table 7 controlling for establishment, worker and job characteristics.

Figure 2: Training Participation Across Occupations



Coefficients are standardized and taken from regression in Column 5 in Table 7 controlling for establishment, worker and job characteristics.

Table 5: Descriptive Statistics Worker and Establishment Characteristics

Variable	Absolute	In Percent	Mean	Std. Dev.
Year 2006	5,618	28.2%		
Year 2007	5,641	28.4%		
Year 2009	4,271	21.5%		
Stat S-H	1,443	7.3%		
State NRW	6,590	33.1%		
State Bavaria	3,984	20.0%		
State MV	1,852	9.3%		
State Saxony	6,027	30.3%		
Sector Industry	5,137	25.8%		
Sector Manuf. / Constr.	5,667	28.5%		
Sector Service	4,681	23.5%		
Sector Health	4,411	22.2%		
Ownership LLC	11,315	56.9%		
Ownership Corporat.	2,997	15.1%		
Ownership Public	4,150	20.9%		
Ownership Partnership	1,434	7.2%		
Founded after 1975	9,641	48.5%		
Employees 100-200	2,752	13.8%		
Employees 200-500	5,731	28.8%		
Employees 500-2000	11,413	57.4%		
Worker Council	18,259	91.8%		
Median Wage (€ per day)			97.2	24.9
Women	7,467	37.5%		
Women x Child < 3	198	1%		
Women x Child 3-5	418	2.1%		
Women x Child 6-9	486	2.4%		
Women x Child 10-18	1,355	6.8%		
Cohabiting Household	15,701	78.9%		
Single Household	3,101	15.6%		
Other Household	1,094	5.5%		
Tenure (in Days)			4,647.5	3,154.0
Tenure Squared			$3.15e^7$	$3.71e^7$
Professional Qualification	12,966	65.2%		
Tertiary Degree	4,494	22.6%		
Adv. Prof. Qualification	2,036	10.2%		
Basic Education	400	2.0%		
Age			45.0	9.7
Age Squared			2,120.0	838.8
Health Very Good	4,871	24.5%		
Health Good	10,720	53.9%		
Health Medium	908	4.6%		
Health Bad	151	0.8%		
Health Very Bad	358	1.8%		
Limited Contract	1,237	6.2%		
Leader	6,026	30.3%		
Practice Job Learnt	13,700	68.9%		
Labor Attachment			8.5	2.4
Full-Time Employed	16,337	82.1%		
Marginally Employed	682	3.4%		
Part-Time Employed	2,448	12.3%		
Working Time 43h+	429	2.2%		

Table 6: Most Relevant Predictors For Training

Rank	**Job-Related Training in General**	**Firm-financed Training**	**Worker co-financed Training**
1	Task Computer Usage	Task Computer Usage	Health Occupations
2	Task Organizing / Making Plans	Managerial Resp.	Task Advising / Informing
3	Task Advising/Counselling	*Task Manufacturing*	*Age Squared*
4	Health Occupations	Dealer Occupation	Task Computer Usage
5	*Task Manufacturing*	Establ. Wage Level	Educational or Social Jobs
6	Educational or Social Jobs	Task Organizing	Technical-Physical Jobs
7	Managerial Resp.	Task Advising/Counselling	Organisational Jobs
8	Establ. Wage Level	Establ. Wage Trend	*Year 2010*
9	Task Gathering information	Engineering Occupation	Occupation Electr. Techn.
10	Organisational Jobs	Task Negotiaing	Age

Note: Table 6 shows the for each regression in Table 7 to 9 the five variables with the largest absolute t-Statistic (excluding the intercept). Variables in black have a positive coefficient while red indicates negative coefficients. All variables are statistically significant on the 1% percent level.

Table 7: Participation in Job-related Training Courses

Dependent Variable:	**Participation in Job-related Training Courses**									
Variable	Null Model		+Firm Variables		+Worker Variables		+Firm & Worker Var.		+Firm & Worker + Occ.	
Year 2008	-0.15	(0.09)	-0.14	(0.09)	-0.13	(0.09)	-0.13	(0.09)	-0.13	(0.08)
Year 2009	-0.01	(0.08)	-0.01	(0.08)	0.01	(0.08)	0.01	(0.08)	-0.02	(0.07)
Year 2010	-0.16**	(0.07)	-0.16**	(0.07)	-0.13*	(0.07)	-0.13*	(0.07)	-0.14**	(0.06)
State Saxony			0.16	(0.12)			0.09	(0.11)	0.11	(0.09)
State Bavaria			-0.01	(0.07)			-0.01	(0.07)	0.03	(0.05)
State NRW			-0.09	(0.08)			-0.06	(0.08)	-0.02	(0.06)
State MV			-0.06	(0.09)			-0.1	(0.09)	-0.02	(0.06)
Service Sector			0.32***	(0.06)			0.31***	(0.06)	0.09*	(0.05)
Public Sector			0.23***	(0.08)			0.19**	(0.07)	0.03	(0.06)
Employees 200-500			-0.05	(0.06)			-0.05	(0.06)	0.00	(0.04)
Employees 500-2000			-0.06	(0.08)			-0.07	(0.07)	-0.04	(0.05)
Founded before 1975			-0.09	(0.08)			-0.07	(0.08)	0.00	(0.06)
Founded after 1991			0.02	(0.06)			0.02	(0.06)	0.03	(0.04)
Worker Council			0.00	(0.04)			0.02	(0.04)	0.05	(0.04)
Median Wage			0.28***	(0.07)			0.2***	(0.06)	0.07*	(0.04)
Median Wage Trend			0.22***	(0.05)			0.19***	(0.05)	0.14***	(0.03)
Wage Compression			0.21***	(0.07)			0.17**	(0.07)	0.05	(0.06)
Women					0.00	(0.04)	-0.02	(0.04)	-0.07**	(0.03)
Cohabiting					0.04*	(0.03)	0.05*	(0.03)	0.01	(0.02)
Tertiary Educ.					0.32***	(0.04)	0.31***	(0.04)	0.03	(0.04)
No Voc. Qualification					-0.16***	(0.03)	-0.15***	(0.03)	-0.06**	(0.03)
Age					-0.02	(0.23)	0.01	(0.23)	0.37	(0.24)
Age Sq.					-0.27	(0.22)	-0.29	(0.22)	-0.59**	(0.23)
Unempl. Exp.					-0.09**	(0.03)	-0.08**	(0.03)	-0.06	(0.04)
Labor Attachment					0.09***	(0.02)	0.08***	(0.02)	0.03	(0.02)
Foreign Citizenship					-0.05*	(0.03)	-0.04	(0.03)	-0.03	(0.02)
Health Status					-0.09***	(0.03)	-0.1***	(0.03)	-0.06**	(0.02)
Limited Contract									-0.04	(0.03)
Managerial Resp.									0.15***	(0.03)
Part-Time									-0.04	(0.03)
Tenure									-0.2	(0.21)
Tenure Squared									0.1	(0.18)
Recently Hired									0.04	(0.04)
Constant	-0.38***	(0.08)	-0.25***	(0.06)	-0.36***	(0.07)	-0.26***	(0.05)	-0.31***	(0.04)
Occupations	No		No		No		No		Yes	
Tasks	No		No		No		No		Yes	
σ^2_{Firm}	0.65	(0.10)	0.22	(0.06)	0.48	(0.08)	0.18	(0.05)	0.05	(0.02)
σ^2_{Worker}	1.5	(0.16)	1.5	(0.16)	1.29	(0.14)	1.3	(0.14)	0.79	(0.1)
N	12560		12560		12560		12560		12560	
Wald χ^2	207		358		606		743		2141	
Log Likelihood	-8058		-8006		-7898		-7852		-7453	

Note: Own calculations; Multilevel logit estimation; cluster-robust standard errors are in parentheses (on the firm-level); Mandatory-job training courses are excluded. Regression includes a variables capturing the number of months a worker has been asked retrospectively. Significance levels: *= significant at 10%-level, **= significant at 5%-level, ***= significant at 1%-level

Table 8: Participation in Training Courses in Working-Time

Dependent Variable:	**Participation in Job-related Training Courses In Working-Time**									
Variable	Null Model		+Firm Variables		+Worker Variables		+Firm & Worker Var.		+Firm & Worker + Occ.	
Year 2008	-0.1	(0.08)	-0.1	(0.08)	-0.09	(0.08)	-0.09	(0.08)	-0.1	(0.08)
Year 2009	0.04	(0.08)	0.05	(0.08)	0.06	(0.07)	0.06	(0.08)	0.03	(0.07)
Year 2010	-0.05	(0.06)	-0.05	(0.06)	-0.03	(0.06)	-0.03	(0.06)	-0.05	(0.06)
State Saxony			0.34**	(0.13)			0.27**	(0.13)	0.23**	(0.11)
State Bavaria			0.1	(0.09)			0.08	(0.09)	0.13*	(0.07)
State NRW			-0.05	(0.11)			-0.03	(0.10)	0.00	(0.08)
State MV			0.06	(0.09)			0.01	(0.09)	0.04	(0.07)
Service Sector			0.14**	(0.07)			0.15**	(0.06)	0.07	(0.06)
Public Sector			0.11	(0.08)			0.1	(0.08)	0.00	(0.07)
Employees 200-500			-0.01	(0.06)			0.00	(0.06)	0.05	(0.05)
Employees 500-2000			-0.02	(0.08)			-0.04	(0.07)	-0.01	(0.05)
Founded before 1975			0.00	(0.09)			0.02	(0.08)	0.04	(0.08)
Founded 1976-1991			0.03	(0.06)			0.03	(0.06)	0.04	(0.05)
Worker Council			0.03	(0.04)			0.03	(0.04)	0.05	(0.04)
Median Wage			0.41***	(0.07)			0.33***	(0.07)	0.2***	(0.06)
Median Wage Trend			0.2***	(0.06)			0.18***	(0.06)	0.14***	(0.05)
Wage Compression			0.25***	(0.06)			0.22***	(0.06)	0.13***	(0.05)
Women					-0.08**	(0.03)	-0.09***	(0.03)	-0.09***	(0.03)
Cohabiting					0.04	(0.03)	0.04	(0.03)	0.02	(0.03)
Tertiary Educ.					0.22***	(0.04)	0.21***	(0.04)	0.00	(0.04)
No Voc. Qualification					-0.14***	(0.04)	-0.13***	(0.04)	-0.06	(0.04)
Age					-0.42*	(0.22)	-0.38*	(0.22)	-0.19	(0.21)
Age Sq.					0.21	(0.21)	0.18	(0.21)	0.00	(0.21)
Unempl. Exp.					-0.1***	(0.03)	-0.08**	(0.03)	-0.03	(0.04)
Labor Attachment					0.08***	(0.02)	0.07***	(0.02)	0.02	(0.02)
Foreign Citizenship					-0.06*	(0.03)	-0.05*	(0.03)	-0.04	(0.03)
Health Status					-0.09***	(0.03)	-0.1***	(0.03)	-0.06**	(0.02)
Limited Contract									-0.09***	(0.03)
Managerial Resp.									0.13***	(0.03)
Part-Time									-0.12***	(0.04)
Tenure									-0.22	(0.21)
Tenure Squared									0.19	(0.18)
Recently Hired									0.02	(0.04)
Constant	-1.26***	(0.08)	-1.11***	(0.06)	-1.24***	(0.08)	-1.12***	(0.05)	-1.14***	(0.05)
Occupations	No		No		No		No		Yes	
Tasks	No		No		No		No		Yes	
σ^2_{Firm}	0.56	(0.10)	0.22	(0.05)	0.46	(0.09)	0.19	(0.05)	0.09	(0.03)
σ^2_{Worker}	1.18	(0.13)	1.18	(0.13)	1.06	(0.12)	1.06	(0.12)	0.74	(0.09)
N	12560		12560		12560		12560		12560	
Wald χ^2	73		181		264		397		1642	
Log Likelihood	-7262		-7216		-7164		-7123		-6877	

Note: Own calculations; Multilevel logit estimation; cluster-robust standard errors are in parentheses (on the firm-level); Mandatory-job training courses are excluded. Regression includes a variables capturing the number of months a worker has been asked retrospectively. Significance levels: *= significant at 10%-level, **= significant at 5%-level, ***= significant at 1%-level

Table 9: Participation in Training Courses Overlapping With Leisure-Time

Dependent Variable:	**Participation in Job-related Training Courses At Least Partly in Leisure-Time**									
Variable	Null Model		+Firm Variables		+Worker Variables		+Firm & Worker Var.		+Firm & Worker + Occ.	
Year 2008	-0.17*	(0.09)	-0.17*	(0.09)	-0.16*	(0.09)	-0.16	(0.09)	-0.16*	(0.09)
Year 2009	-0.13	(0.08)	-0.13	(0.08)	-0.11	(0.08)	-0.11	(0.08)	-0.14	(0.08)
Year 2010	-0.26***	(0.07)	-0.26***	(0.07)	-0.22***	(0.07)	-0.22***	(0.07)	-0.22***	(0.07)
State Saxony			-0.07	(0.16)			-0.11	(0.15)	-0.05	(0.14)
State Bavaria			-0.08	(0.10)			-0.07	(0.10)	-0.09	(0.09)
State NRW			-0.04	(0.11)			0.01	(0.11)	0.03	(0.11)
State MV			-0.11	(0.13)			-0.13	(0.12)	-0.05	(0.11)
Service Sector			0.36***	(0.07)			0.29***	(0.07)	0.03	(0.07)
Public Sector			0.18*	(0.10)			0.13	(0.09)	0.04	(0.07)
Employees 200-500			-0.08	(0.07)			-0.08	(0.06)	-0.08	(0.06)
Employees 500-2000			-0.08	(0.09)			-0.09	(0.09)	-0.06	(0.07)
Founded before 1975			-0.18	(0.12)			-0.16	(0.12)	-0.07	(0.11)
Founded 1976-1991			-0.04	(0.08)			-0.03	(0.07)	-0.01	(0.06)
Worker Council			0.00	(0.05)			0.02	(0.05)	0.06	(0.04)
Median Wage			-0.05	(0.10)			-0.09	(0.09)	-0.15*	(0.08)
Median Wage Trend			0.13**	(0.07)			0.11*	(0.06)	0.07	(0.06)
Wage Compression			0.13	(0.08)			0.07	(0.08)	-0.02	(0.07)
Women					0.21***	(0.05)	0.17***	(0.05)	0.08*	(0.04)
Cohabiting					0.02	(0.03)	0.02	(0.03)	-0.01	(0.04)
Tertiary Educ.					0.28***	(0.04)	0.28***	(0.04)	0.07*	(0.04)
No Voc. Qualification					-0.18***	(0.05)	-0.18***	(0.05)	-0.08*	(0.04)
Age					0.58**	(0.28)	0.59**	(0.28)	0.87***	(0.31)
Age Sq.					-0.85***	(0.28)	-0.86***	(0.28)	-1.06***	(0.30)
Unempl. Exp.					-0.03	(0.04)	-0.03	(0.04)	-0.05	(0.04)
Labor Attachment					0.09***	(0.03)	0.09**	(0.04)	0.04	(0.03)
Foreign Citizenship					-0.02	(0.04)	-0.02	(0.04)	-0.01	(0.03)
Health Status					-0.05*	(0.03)	-0.05*	(0.03)	-0.02	(0.03)
Limited Contract									0.05	(0.03)
Managerial Resp.									0.08**	(0.04)
Part-Time									0.09**	(0.04)
Tenure									0.11	(0.21)
Tenure Squared									-0.19	(0.20)
Recently Hired									0.1**	(0.04)
Constant	-2.02***	(0.10)	-1.97***	(0.10)	-2***	(0.09)	-1.99***	(0.09)	-2.02***	(0.08)
Occupations	No		No		No		No		Yes	
Tasks	No		No		No		No		Yes	
σ^2_{Firm}	0.66	(0.11)	0.4	(0.09)	0.47	(0.09)	0.34	(0.08)	0.18	(0.06)
σ^2_{Worker}	2.1	(0.24)	2.1	(0.24)	1.92	(0.22)	1.91	(0.22)	1.52	(0.19)
N	12560		12560		12560		12560		12560	
Wald χ^2	149		240		394		491		1185	
Log Likelihood	-5849		-5825		-5747		-5731		-5538	

Note: Own calculations; Multilevel logit estimation; cluster-robust standard errors are in parentheses (on the firm-level); Mandatory-job training courses are excluded. Regression includes a variables capturing the number of months a worker has been asked retrospectively. Significance levels: *= significant at 10%-level, **= significant at 5%-level, ***= significant at 1%-level

Table 10: Participation in Training: Intensity

Dependent Variable:	**Number of Training Courses (incl. non-participants)**									
Variable	Null Model		+Firm Variables		+Worker Variables		+Firm & Worker Var.		+Firm & Worker + Occ.	
Year 2008	-0.19**	(0.08)	-0.18**	(0.08)	-0.18**	(0.08)	-0.18**	(0.08)	-0.17**	(0.08)
Year 2009	0.03	(0.08)	0.03	(0.08)	0.05	(0.08)	0.04	(0.08)	0.02	(0.08)
Year 2010	-0.15*	(0.08)	-0.15*	(0.08)	-0.12	(0.08)	-0.12	(0.08)	-0.13	(0.08)
State Saxony			0.3**	(0.15)			0.24*	(0.14)	0.26**	(0.10)
State Bavaria			0.1	(0.12)			0.09	(0.12)	0.15*	(0.09)
State NRW			-0.02	(0.12)			0.02	(0.11)	0.09	(0.08)
State MV			0.05	(0.19)			-0.01	(0.19)	0.11	(0.13)
Service Sector			0.48***	(0.08)			0.45***	(0.08)	0.12**	(0.06)
Public Sector			0.32***	(0.10)			0.26***	(0.10)	0.01	(0.08)
Employees 200-500			-0.04	(0.07)			-0.04	(0.07)	0.01	(0.05)
Employees 500-2000			-0.04	(0.10)			-0.07	(0.09)	-0.05	(0.07)
Founded before 1975			-0.09	(0.10)			-0.06	(0.10)	0.01	(0.07)
Founded 1976-1991			-0.04	(0.10)			-0.04	(0.10)	0.04	(0.07)
Worker Council			0.14	(0.14)			0.16	(0.13)	0.17*	(0.09)
Median Wage			0.01***	(0.00)			0.01**	(0.00)	0.00*	(0.00)
Median Wage Trend			0.11***	(0.02)			0.1***	(0.02)	0.09***	(0.02)
Wage Compression			0.84***	(0.27)			0.68***	(0.24)	0.26	(0.19)
Women					0.07*	(0.04)	0.05	(0.04)	0.05	(0.04)
Cohabiting					0.04	(0.04)	0.04	(0.04)	0.01	(0.03)
ISCED High					0.36***	(0.04)	0.35***	(0.04)	0.08**	(0.04)
ISCED Low					-1.04***	(0.17)	-1.01***	(0.17)	-0.6***	(0.16)
Age					0.24**	(0.12)	0.25**	(0.12)	0.37***	(0.13)
Age Sq.					-0.38***	(0.11)	-0.39***	(0.11)	-0.48***	(0.12)
Unempl. Exp.					-0.03	(0.04)	-0.02	(0.04)	0.00	(0.05)
Labor Attachment					0.07***	(0.01)	0.07***	(0.01)	0.04***	(0.01)
Foreign Citizenship					-0.11	(0.10)	-0.09	(0.10)	-0.04	(0.10)
Health Status					-0.04***	(0.02)	-0.04***	(0.02)	-0.02*	(0.01)
Limited Contract									-0.05	(0.06)
Leader									0.18***	(0.03)
Part-Time									-0.03	(0.04)
Tenure									-0.01	(0.02)
Tenure Squared									0.00	(0.00)
Recently Hired									0.09*	(0.05)
Constant	-0.38***	(0.07)	-1.85***	(0.33)	-0.37***	(0.07)	-1.62***	(0.31)	-1.7***	(0.26)
Occupations	No		No		No		No		Yes	
Tasks	No		No		No		No		Yes	
σ^2_{Firm}	0.31	(0.05)	0.12	(0.03)	0.25	(0.04)	0.11	(0.02)	0.04	(0.01)
σ^2_{Worker}	0.34	(0.02)	0.34	(0.02)	0.29	(0.02)	0.3	(0.02)	0.2	(0.02)
N	13089		13089		13089		13089		13089	
Wald χ^2	396		680		859		1200		3963	
Log Likelihood	-16327		-16270		-16132		-16081		-15590	

Note: Own calculations; Multilevel poisson estimation; cluster-robust standard errors are in parentheses (on the firm-level); Mandatory-job training courses are excluded. Regression includes a variables capturing the number of months a worker has been asked retrospectively. Significance levels: *= significant at 10%-level, **= significant at 5%-level, ***= significant at 1%-level

Table 11: Participation in Training: Multinomial Logit

Dependent Variable:	**Training Participation in ...**		
Variable	Firm-financed Training	Worker co-financed	Both Forms of Training
Year 2008	-0.14 (0.17)	-0.24 (0.20)	-0.9*** (0.34)
Year 2009	0.15 (0.16)	-0.14 (0.19)	-0.61* (0.31)
Year 2010	-0.10 (0.16)	-0.45** (0.19)	-0.8*** (0.30)
State Saxony	0.33* (0.18)	-0.09 (0.24)	0.31 (0.29)
State Bavaria	0.26* (0.15)	-0.16 (0.20)	0.30 (0.25)
State NRW	-0.08 (0.16)	-0.02 (0.19)	0.20 (0.25)
State MV	-0.10 (0.18)	-0.09 (0.26)	0.33 (0.30)
Service Sector	0.27*** (0.10)	0.03 (0.11)	0.07 (0.14)
Public Sector	-0.10 (0.16)	0.01 (0.14)	-0.09 (0.20)
Employees 100-200	-0.17* (0.09)	0.10 (0.10)	0.13 (0.13)
Employees 500-2000	-0.20** (0.09)	-0.01 (0.10)	0.08 (0.12)
Founded before 1975	0.16 (0.14)	-0.10 (0.16)	-0.07 (0.17)
Founded after 1991	0.17 (0.11)	0.02 (0.14)	-0.04 (0.15)
Worker Council	0.14 (0.15)	0.17 (0.13)	0.32 (0.20)
Median Wage	0.01** (0.00)	0.00 (0.00)	0.00 (0.00)
Median Wage Trend	0.10*** (0.03)	0.06 (0.03)	0.13*** (0.03)
Wage Compression	0.06 (0.05)	-0.08 (0.06)	0.06 (0.08)
Women	-0.24*** (0.06)	0.01 (0.08)	0.18 (0.12)
Cohabiting	0.04 (0.06)	0.02 (0.08)	-0.01 (0.10)
ISCED High	-0.03 (0.10)	0.09 (0.09)	0.20* (0.12)
ISCED Low	-0.34 (0.27)	-0.38 (0.30)	-14.03*** (0.29)
Age	0.04 (0.24)	0.92*** (0.32)	0.39 (0.44)
Age Sq.	-0.22 (0.23)	-1.04*** (0.30)	-0.78* (0.42)
Unempl. Exp.	-0.12 (0.10)	-0.25** (0.12)	-0.21 (0.14)
Labor Attachment	0.03 (0.03)	0.03 (0.03)	0.09 (0.06)
Foreign Citizenship	-0.19 (0.21)	0.00 (0.24)	-0.57 (0.65)
Poor Health Status	-0.06 (0.06)	-0.02 (0.07)	-0.30** (0.12)
Limited Contract	-0.36** (0.14)	0.12 (0.13)	-0.02 (0.18)
Leader	0.32*** (0.07)	0.23*** (0.09)	0.42*** (0.11)
Part-Time	-0.27*** (0.10)	0.14 (0.09)	0.00 (0.14)
Tenure	-0.21 (0.21)	0.07 (0.21)	-0.27 (0.32)
Tenure Squared	0.12 (0.19)	-0.21 (0.19)	0.21 (0.28)
Recently Hired	0.04 (0.13)	0.24* (0.13)	0.25 (0.23)
Constant	-2.89*** (0.38)	-2.58*** (0.50)	-5.70*** (0.67)
Occupations	Yes		
Tasks	Yes		
N	12560		
Pseudo R^2	0.102		
Log Likelihood	-12525		

Note: Multinomial logit estimation. cluster-robust standard errors are in parentheses (on the firm-level); Mandatory-job training courses are excluded. Regression includes a variables capturing the number of months a worker has been asked retrospectively. Reference Groups are workers who do not participate in training.
Significance levels: *= significant at 10%-level, **= significant at 5%-level, ***= significant at 1%-level