

Pfeifer, Christian

Conference Paper

Intra-firm Wage Compression and Cost Coverage of Training: Evidence from Linked Employer-Employee Data

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Training and Apprenticeship, No. D07-V2

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Pfeifer, Christian (2013) : Intra-firm Wage Compression and Cost Coverage of Training: Evidence from Linked Employer-Employee Data, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Training and Apprenticeship, No. D07-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/80030>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Intra-firm Wage Compression and Cost Coverage of Training: Evidence from Linked Employer-Employee Data

Christian Pfeifer^{a) b)}

^{a)} Institute of Economics, Leuphana University Lüneburg, Scharnhorststr. 1, 21335 Lüneburg, Germany; phone: +49-4131-6772301; e-mail: pfeifer@leuphana.de.

^{b)} Forschungsinstitut zur Zukunft der Arbeit (IZA), Germany.

(work in progress; Jahrestagung Verein für Socialpolitik; 26.2.2013)

Acknowledgements: I thank Michael Beckmann, Lutz Bellmann, Knut Gerlach, Christian Grund, Olaf Hübler, Matthias Kräkel, Markus Leibrecht, Jens Mohrenweiser, participants of the IAB Colloquium in Nürnberg 2011, at the 15th Colloquium on Personnel Economics in Paderborn 2012, at the 26th Annual Congress of the European Society for Population Economics (ESPE) in Bern 2012, at the IAB Establishment Panel User Conference in Nürnberg 2012, and at seminars in Lüneburg for helpful comments. This study uses the Cross-sectional model of the Linked Employer-Employee Data (LIAB) (Years 2005 and 2007) from the IAB. Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and remote data access. Upon request, I can provide my program codes.

Intra-firm Wage Compression and Cost Coverage of Training: Evidence from Linked Employer-Employee Data

Abstract

This paper uses German linked employer-employee data in order to estimate the impact of intra-firm wage dispersion on the probability that firms pay for continuous training. About half of all firms in the estimation sample cover all direct and indirect training costs, which contradicts the standard human capital approach with perfect labor markets. The main finding of my cross-section, panel, and instrumental variable Probit estimations is that firms with larger intra-firm wage compression are indeed more likely to cover all direct and indirect training costs, which is consistent with theoretical considerations of the "new training literature" about imperfect labor markets.

Keywords: Firm-sponsored training; Human capital; Linked employer-employee data; Wage compression

JEL Classification: J24; J31; M53

1. Introduction

Employer-provided continuous training has received increasing attention in economics during the past decades. One reason is its importance for productivity and economic growth. Another reason is the stimulating theoretical work of the "new training literature" that has developed the standard human capital framework by Becker (1962) further. Becker modeled decisions to invest in on-the-job training in an economy with perfect labor markets (e.g., wages equal productivity in all firms, no mobility costs, complete information, no union bargained collective contracts). His main result was that firms do not cover the costs for general training and that firms and workers share the costs for firm-specific training. Workers can keep all returns to training in the former case, whereas workers and firms share the returns in the latter case. Since empirical observations suggest however that firms are highly involved in training and even pay for general training (e.g., the German apprenticeship system (Acemoglu and Pischke, 1998; Mohrenweiser and Zwick, 2009)), the "new training literature" has challenged the assumption of perfect labor markets and the results for cost coverage of training by Becker (for extensive reviews of the theoretical and empirical training literature see for example Asplund, 2005, and Leuven, 2005).

Eckaus (1963) has already stated that firms in imperfect labor markets are likely to pay for more training than Becker's model would predict. For example, firms cannot so easily let workers pay for their training, if training and regular output are jointly produced and training costs cannot be perfectly identified. More influential is Eckaus' notion that firms would have incentives to pay for training, if they can capture rents from training which would be the case for not perfectly mobile workers. Katz and Ziderman (1990) and Chang and Wang (1996) have emphasized information

asymmetries from which imperfect labor mobility arises. They assume that current firms have private information about the productivity of a worker after training. Because other firms do not have these information, they cannot pay the same wages as the current firm. Consequently, the current firm has at least to some degree the opportunity to pay wages below the trained worker's marginal product and to capture rents from training.

A series of prominent papers by Acemoglu and Pischke (1998, 1999a, 1999b) has also analyzed the cost coverage of training in imperfect labor markets. The basic rationale is that firms bear training costs, if they have monopsony power and can capture rents from training due to wage compression (wages relatively more compressed than productivity, wage increases smaller than productivity increases after training) (Acemoglu and Pischke, 1999a). Examples are information asymmetries with respect to a worker's training, ability and motivation (Acemoglu and Pischke, 1998) as well as labor market institutions that affect firms' wage structures such as employment protection, minimum wages, collective contracts, and co-determination (Acemoglu and Pischke, 1999b). Dustmann and Schönberg (2009) focus in their model on unions that increase wage compression due to minimum wages bargained in collective contracts, which in turn increases firm-financed training. Empirical support is presented for apprenticeship training in German firms.

To sum up, one core element in theoretical models of the "new training literature" is that firms with more compressed wage structures (lower intra-firm wage dispersion) should have larger incentives to pay for training, because they are better able to capture rents from training. Pischke (2005, p. 51) concludes: *"Strictly speaking, labor market institutions are not really necessary for this argument, although the example of a minimum wage highlights the workings of the model nicely. However, what is necessary*

for firms to invest is simply that the wage structure $w(t)$ is compressed, i.e. that $w(t)$ is flatter than $f(t)$ [wages w and productivity f are functions of the level of training t]. If this is the case, then the rents the firm can earn from more skilled workers will be greater than the rents earned from less skilled workers. Hence, it may invest in training." Consequently, firms with more compressed wage structures should also be more likely to cover all training costs. This hypothesis is tested in the subsequent paper using linked employer-employee data for large profit maximizing firms in Germany, which allow to generate conditional intra-firm wage dispersion measures as proxies for firms' wage compression. In doing this, I follow partly the call by Acemoglu and Pischke (1999b, p. 567) that "[f]uture empirical work should test the more micro-level implications that follow from our analysis and contrast them with those of the standard theory."

Although a large number of empirical studies on firms' determinants of training exists already for Germany (e.g., Düll and Bellmann, 1998; Düll and Bellmann, 1999; Gerlach and Jirjahn, 2001; Gerlach et al., 2002; Allaart et al., 2009; Stegmaier, 2010; Bellmann et al., 2010; Görlitz, 2010; Görlitz and Stiebale, 2011) and other countries (for literature reviews see for example Asplund, 2005, and Leuven, 2005), only few studies have explicitly examined firms' determinants of training cost coverage (Leber, 2000; Bellmann and Düll, 2001). From several studies, which have used household survey data, we know however that firms bear most of the direct training costs (e.g., Loewenstein and Spletzer, 1998, 1999, for the US; Pischke, 2001, for Germany; Booth and Bryan, 2005, for the UK). In my estimation sample, about half of all training firms cover even all indirect and direct training costs, i.e., the training takes place during paid working time and the firm pays for all outlays such as course fees and travel costs. Such

a complete cost coverage of training is of course largely inconsistent with Becker's model, because firms would not pay at all for general training and only partly for firm-specific training.

As far as I know, no econometric study has yet explicitly tested if a positive correlation between intra-firm wage compression and cost coverage of employer-provided continuous training exists. Though, two recent studies by Almeida-Santos and Mumford (2005) and by Ericson (2008) look at the relationship between wage compression within occupations and individual worker's training participation. Almeida-Santos and Mumford (2005) find with British linked employer-employee data a negative correlation between wage dispersion and training incidence and duration, i.e., more compressed wages lead to more training. Ericson (2008) finds with data from the Swedish Labour Force Survey that general training duration is positively correlated with wage dispersion, whereas duration of firm-specific and mixed training is not significantly affected by the wage dispersion measures. In both studies, the wage compression proxies measure however not the intra-firm wage dispersion but the wage dispersion within occupations and across firms. For Germany, Beckmann (2002a, 2002b) analyses indirectly the effect of wage compression on apprenticeship training by using proxies such as collective contract coverage, which positively affects the probability and intensity of apprenticeship training.

The remainder of the paper is structured as follows. The next section informs about the data and the estimation strategy. Section 3 presents the results of the regression analysis. The paper concludes with a short summary in Section 4.

2. Data and Estimation Strategy

2.1. Estimation Sample

The used data are the cross-sectional models of the German linked employer-employee data set of the Institute for Employment Research (LIAB) (Alda et al., 2005).¹ The LIAB links employer-side information from the IAB Establishment Panel with employee information from process-produced data. The process-produced employee data stem basically from the notification procedure for unemployment, pension, and health insurances. Employers must notify the social security agencies about all employees who are covered by social security at the start and at the end of an employment relationship as well as on the last day of each year. These process-produced employee data include socio-demographic characteristics and individual daily gross wages of workers (in Euros), which are used to generate variables for the conditional intra-firm wage dispersion as an inverse measure for wage compression. Disadvantages of the data are that no information about working hours are available and that wages are censored at the upper earnings limit for social security contributions. Due to the absence of working hours in the data, meaningful aggregate wage variables at the firm-level can only be computed for full-time workers (with the exclusion of apprentices, trainees etc.). The wage censoring leads to a downward bias when proxies for intra-firm wage dispersion are generated, because we observe too low wages (wages equal the social security contribution limit) for high wage workers (wages above the social security contribution limit). This bias should however be much smaller for

¹ For more details see: http://fdz.iab.de/en/Integrated_Establishment_and_Individual_Data/LIAB.aspx (31.1.2012).

conditional than for unconditional wage dispersion measures and can be dealt with by applying censored regression techniques such as Tobit regressions.

As the focus is on firms' determinants of complete cost coverage of continuous training, the IAB Establishment Panel is the main data source for the subsequent analysis. The IAB Establishment Panel contains data on establishments from all sixteen German federal states ("Bundesländer") and all industries. Every year more than 15,000 firms with at least one employee covered by social security are interviewed in an unbalanced panel design survey. The sample is stratified according to ten establishment sizes and sixteen industries in each federal state, with oversampling of larger firms. The observational unit is the establishment, i.e., the local unit in which major activities of an enterprise are carried out.² Main concern of the survey is to gain insights into the firm's most important parts of operation, decision-making, and more specifically employment.

For the purpose of this study, I use the waves 2005 and 2007, because they contain questions about coverage of direct and indirect training costs. Note that the data do not include the years of the big economic crisis starting in 2008, during which many German firms used short-term work ("Kurzarbeit") with training courses at least partly paid for by the Federal Employment Agencies, because this might bias the results. Due to the interest in firms' profit maximizing rationales for training cost coverage, the sample is restricted to profit-maximizing firms from the private sector that have trained at least one worker in the first half of a survey year.³ As training is likely to occur not

² In this paper, the terms establishment and firm are used interchangeably.

³ Note that less than ten percent of large profit-maximizing firms are excluded from the sample because of not having trained any worker.

continuously (i.e., most workers are likely to receive their training once in a while and not the same amount of training in every time period or always in the first half of a survey year), the sample is further restricted to firms with at least 100 workers to mitigate this problem. The sample restriction to larger firms is also preferable in order to make the wage dispersion measures at the firm-level meaningful. Because only full-time workers are considered for the generation of wage variables at the firm-level, the additional restriction is imposed that firms have at least 10 workers in the data from whom the firm-level wage information are generated. At last, only firms without missing values in the used variables are considered. In total, 2,118 firms for the year 2005 and 2,011 firms for the year 2007 remain in the sample for the subsequent empirical analysis. 1,136 of these firms are contained in both years, i.e., in 2005 as well as in 2007 (balanced panel).

2.2. Estimation Strategy and Variables

In order to analyze firms' determinants of complete training cost coverage, a binary variable has been generated (*COSTCOV*), which takes the value one, if a firm states that it usually pays for all direct costs (e.g., course fees, travel costs) and bears also the indirect costs (i.e., the training takes place during paid working time). Remarkably, about half of the firms in the sample completely cover all training costs. Because of the binary dependent variable, binary Probit models are estimated as specified in equation (1), in which Φ denotes the standard normal cumulative density function. Greek letters denote the parameters to be estimated. j is a firm index and t is a time index.

$$\Pr(COSTCOV_{jt} = 1) = \Phi(\alpha + \beta_1 \log WSERT_{jt} + \beta_2 \log WMEAN_{jt} + \gamma X_{jt}) \quad (1)$$

The explanatory variable of main interest is the intra-firm wage compression, for which a proxy can be generated from the process-produced employee data. The most simple approach would be to use the standard deviation of full-time workers' daily wages in a given firm, which would measure the unconditional wage dispersion. This unconditional wage dispersion has however the disadvantage that it does not account for differences in worker characteristics such as qualifications, which affect productivity and wage classifications. Therefore, a conditional wage dispersion measure is a much better proxy for wage compression. I follow the approach of Winter-Ebmer and Zweimüller (1999), who analyzed the effect of intra-firm wage dispersion on firm performance.⁴ Exploiting the nature of the linked employer-employee data set, log-linear Mincer earnings functions for full-time workers (i) are estimated separately for every firm (j) in a given year (t), as specified in equation (2).

$$\begin{aligned} \log WAGE_{ijt} = & \alpha_j + \beta_1 AGE_{ijt} + \beta_2 AGE_{ijt}^2 + \beta_3 TENURE_{ijt} + \beta_4 TENURE_{ijt}^2 \\ & + \beta_5 APPRENTICE_{ijt} + \beta_6 UNIVERSITY_{ijt} + \beta_7 FEMALE_{ijt} + \varepsilon_{ijt} \end{aligned} \quad (2)$$

The dependent variable is the log of workers' individual daily wages. The explanatory variables include the usual productivity related individual worker characteristics such as age, squared age, tenure, squared tenure, highest qualification categories (no job qualification as reference group, apprenticeship degree, university degree), and a female dummy. In order to account for censored wages in the data, I estimate Tobit regressions with different upper earnings limits for East and West Germany as well as for the year

⁴ This approach has been widely used with linked employer-employee data in order to study the effects of wage inequality on firm performance measures such as productivity and profits. For a literature review see Mahy et al. (2011, Appendix Table A1).

2005 and the year 2007.⁵ Based on the results for a firm's earnings function, the standard error of the Tobit regression is then generated as a proxy for the intra-firm wage compression (*logWSERT*). The standard error of the regression in firm *j* can be interpreted as the standard deviation of workers' individual error terms in an estimated earnings function for firm *j* in year *t*, as defined in equation (3). A larger standard error of the regression indicates a larger conditional intra-firm wage dispersion and consequently lower intra-firm wage compression.

$$\log WSERT_{jt} = \text{SER}_{jt}^{\text{TOBIT}} = \hat{\sigma}_{jt} = \sqrt{\text{sum of squared residuals} / \text{number of observations}} \quad (3)$$

Descriptive statistics for the intra-firm wage compression proxy (*logWSERT*) are displayed in Table 1. Mean standard errors of the regressions are on average approximately 0.22 with a standard deviation of 0.07. When comparing the means and standard deviations of my estimated standard errors of the regressions with the results of Winter-Ebmer and Zweimüller (1999), only small differences are evident. Winter-Ebmer and Zweimüller (1999) used data of workers in 130 firms, which have been obtained from Austrian social security records for the years 1975-1991. Their estimated standard errors of Tobit regressions for the log of monthly income have a mean of 0.205 with a standard deviation of 0.074.

- Insert Table 1 about here.

⁵ The corresponding censoring values for the upper earnings limits for social security contributions with respect to daily wages in Euros have been set according to the statutory pension fund (http://fdz.iab.de/en/Integrated_Establishment_and_Individual_Data/LIAB/Working_Tools.aspx (31.1.2012)): WEST2005=170.96, EAST2005=144.66, WEST2007=172.60, EAST2007=149.59.

Because a larger standard error of the wage regression ($\log WSERT$) might be the result of larger wage levels in a firm, the Probit regressions in equation (1) include also the mean of log daily wages of full-time workers in a firm in a given year ($\log WMEAN$). The Probit regressions further control for important differences (X) between firms that might affect training as well as wage structures. Industrial relations are important in this context, because unions and works councils are often associated with more compressed wage structures and with interests in more training for workers (e.g., Acemoglu and Pischke, 1999b; Dustmann and Schönberg, 2009). Moreover, the regressions include variables for the number of layoffs and quits, the number of workers, three firm age categories, state of the art of production technology, profit situation, share of women, share of part-time workers, share of qualified workers, 16 federal state dummies, and 15 sector dummies. Table 1 presents complete variable definitions and descriptive statistics.

Equation (1) for the determinants of cost coverage is estimated using binary Probit models for the separate cross-sections 2005 and 2007 as well as random effects Probit models for a balanced panel. Regression results for separate cross-sections give the overall findings more persuasive power than for a pooled sample, in which the results might be driven by certain firms only observed in one year. The random effects models serve mainly as a robustness check in order to account for within-firm variance, because likelihood ratio tests reject in all regressions the hypothesis that the within-firm variance does not significantly contribute to the total variance. The random effects models are chosen over fixed effects models for several reasons. At first, no consistent fixed effects estimators exist for Probit or Logit models in short panels due to the incidental parameter problem. Fixed effects linear probability models are also no feasible

estimation strategy, because training cost coverage, wage structures, and industrial relations are structural firm characteristics based on strategic decisions so that changes are not common and are unlikely to be in effect rapidly. Accordingly, within-firm variance is very low for most variables of interest in my data.

3. Estimation Results

3.1. Main Findings

Table 2 presents the results of the binary Probit regressions for the probability that a firm covers completely all direct and indirect training costs (*COSTCOV*). The first column contains the results of the cross-section Probit for the year 2005 and the second column the results for the year 2007. In the third column, the results of the random effects Probit model for the balanced panel are presented. Marginal effects at the means of all covariates in each estimation sample have been computed in order to facilitate the quantitative interpretation.

- Insert Table 2 about here.

The main finding can be seen from the first row of marginal effects in Table 2. A larger standard error of a firm's workforce wage regression (*logWSERT*) is significantly negatively correlated with the probability that a firm pays all direct and indirect training costs throughout all three regressions, i.e., firms with lower intra-firm conditional wage dispersion (more compressed wage structures) are on average more likely to cover all training costs. A 0.1 log point higher standard error of the wage regression decreases the cost coverage probability in the year 2005 by 4.9 percentage points ($p=0.004$) and in the

year 2007 by 3.6 percentage points ($p=0.035$).⁶ The random effects Probit regression yields a comparable marginal effect of minus 3.7 percentage points ($p=0.075$) per 0.1 log point higher standard error of the regression. The findings are consistent with the theoretical consideration that firms can capture rents from training due to intra-firm wage compression, which provides incentives for firms to pay for continuous training.

Moreover, the results in Table 2 indicate that firms with higher average wages ($\log WMEAN$) have on average a higher probability of covering all training costs. It can further be seen that unions have no significant effects throughout all regressions, whereas works councils increase more strongly the probability that a firm completely covers training costs. Hence, it seems as if firm-level codetermination is more influential in this context than union bargaining. Only few other control variables significantly affect the cost coverage probability in a consistent pattern across the regressions. Firms with more layoffs have a lower probability to completely cover all training costs, which might be explained by amortization aspects and a loss in employment flexibility if adjustment costs increase after the firm has paid for training. Furthermore, the share of part-time workers indicates a positive correlation with the training cost coverage probability. This finding might be surprising at first glance, if amortization aspects are taken into account. Because part-time workers are often associated with the flexible part of a firm's workforce (periphery), one would expect firms to invest less in their human capital. The findings might however be explained by

⁶ For the interpretation of the economic significance of the effect size recall that $\log WSERT$ has a mean of 0.22 with a standard deviation of 0.07 (see descriptive statistics in Table 1). Thus, an increase by one standard deviation of $\log WSERT$ decreases the cost coverage probability by approximately three percentage points.

the subsequent reasons. First, I use aggregate firm-level data so that the share of part-time workers, as part of the peripheral workforce, might be an indicator for the existence of dual internal labor markets, in which firms rely on stable employment relationships with their core workforce who receives training. Second, workers of the peripheral workforce are by definition more often newly employed by a firm and might need work instructions that are paid by the firm. Third, part-time workers have on average lower income, which might lead to credit constraints so that the firm might have to pay for the training. These interpretations are however only speculations which cannot be tested with the used data.

3.2. Robustness Checks

Several robustness checks have been performed that are summarized subsequently.⁷ I have used alternative proxies for the intra-firm wage compression (dispersion) variable. First, I have used the standard deviation of fulltime workers' daily wages in a firm (unconditional wage dispersion). Second, I have used simple linear regressions instead of Tobit regressions to generate the standard errors of the wage regressions for each firm. Both alternative variables are even at higher significance levels negatively correlated with the probability that a firm pays all training costs than the standard errors of Tobit regressions (*logWSERT*)

The next robustness checks deals explicitly with unobserved firm heterogeneity. In order to do so, I have applied two panel estimation techniques that are both problematic

⁷ The complete results of the robustness checks can be requested from the author.

with my data due to very low within-firm variance for most variables of interest. Nevertheless, they should be mentioned. At first, I have estimated a fixed-effects linear probability model. The estimated marginal effects for the intra-firm wage dispersion variable have the negative sign known from the previous Probit models. Due to the low within variance, the effects are however not statistically significant. Still, the results indicate rather a negative than a positive correlation between intra-firm wage dispersion and training cost coverage if time-invariant unobserved firm heterogeneity is taken into account in a fixed-effects linear probability model. Moreover, I have re-estimated the random effects Probit model with additional variables that contain the means of each observed firm characteristic over time, which is known as Mundlak's approach (Mundlak, 1978). The inclusion of group means in random effects models controls intuitively for unobserved heterogeneity and allows dependence between the random effects and the regressors. This approach is a widespread method in econometrics and can also be applied for Probit models (Chamberlain, 1980), which are sometimes called correlated random effects Probit models (for a detailed textbook discussion see for example Wooldridge, 2010, pp. 610-619). The results of the correlated random effects models indicate again rather a negative than a positive correlation between intra-firm wage dispersion and training cost coverage, even though the effects are not statistically significant due to the low within variance.

Another source of endogeneity might be reverse causality, i.e., the causal link might not go from wage compression to training cost coverage but the other way around. If firms pay for training, workers might receive lower returns to training which decreases wage differentials between trained and untrained workers and consequently increases wage compression. In order to deal with this endogeneity problem, I have estimated

instrumental variable (IV) Probit regressions (for detailed discussions see Rivers and Vuong, 1988, and Wooldridge, 2010, pp. 585-594). Note that IV estimation strategies are also suitable to deal with potential omitted variable biases.

As instruments, which affect the intra-firm wage compression, I use the lowest observed wage of a worker in a firm and the mean of the intra-firm standard errors of log daily wage regressions within industry and federal state cells. Previous studies about training have often emphasized institutional minimum wages, which are however not that common in Germany and not observed in the data. Whereas institutional minimum wages can be seen as exogenous to firms, the lowest observed wage in a firm is a rather technical instrument that has the advantage of exploiting large between-firm variance. The rationales for using the mean of the intra-firm standard errors of log daily wage regressions within industry and federal state cells as a second instrument are norms and spillover effects in regional labor markets (e.g., a firm's wage structure is affected by institutional developments in the past and by wage structures of other firms in the same industry and region). Hence, the first stage is estimated with linear regressions that use the lowest observed log daily wage in a firm (*logWMINIMUM*) and the mean of the intra-firm standard errors of log daily wage regressions within industry (15) and federal state (16) cells (*logWSERTis*) as instruments for a firm's intra-firm standard error of log daily wage regressions (*logWSERT*). It can be seen from Table 3 that *logWMINIMUM* is indeed negatively correlated and *logWSERTis* is positively correlated with *logWSERT* at high statistical significance levels in the first stage regressions. The second stage is then estimated with binary Probit regressions that include the standardized predicted error terms for every firm from the first stage regressions ($\hat{\varepsilon}_j^{first} / \hat{\sigma}_{\varepsilon,first}$).

- Insert Table 3 about here.

The coefficients for the standardized predicted error terms are neither in the year 2005 nor in the year 2007 significantly different from zero and the Wald test of exogeneity cannot be rejected. Therefore, endogeneity seems not to be an important issue in my application. Marginal effects on the probability of complete training cost coverage by the firm (*COSTCOV*) are presented in the third column for every year. As I have used the same estimation samples and compute comparable marginal effects at the means of all covariates in each estimation sample, the IV Probit results can be compared in size with the Probit results in Table 2. The results in Table 3 reveal marginal effects of minus 4.6 percentage points in the year 2005 and minus 4.2 percentage points in the year 2007 per 0.1 log point higher standard error of the wage regression. These marginal effects are comparable in size with the results in Table 2. The statistical significance levels are however lower in the IV Probit regressions due to larger standard errors ($p=0.078$ in the year 2005, $p=0.111$ in the year 2007).

Table 4 presents IV Probit results for the use of the lowest observed log daily wage in a firm (*logWMINIMUM*) as only instrument in order to check the sensitivity of the above IV Probit regressions with two instruments. The results do not change noteworthy. Again, the coefficients for the standardized predicted error terms are neither in the year 2005 nor in the year 2007 significantly different from zero and the Wald test of exogeneity cannot be rejected. A 0.1 log point increase of the standard error of the wage regression decreases the probability of complete cost coverage by 4.6 percentage points ($p=0.099$) in the year 2005 and by 3.8 percentage points ($p=0.164$) in the year 2007, which is comparable in size with the previous results.

- Insert Table 4 about here.

The last robustness check is concerned with the firm sample which has been chosen very conservative with respect to firm size, because only firms with at least 100 workers have been included. The preference for this conservative sample restriction was driven by potential sample selectivity and measurement errors with respect to training and the intra-firm wage dispersion variables in smaller firms. Despite these potential problems, I have relaxed the sample restriction and re-estimated all regressions for a sample of firms with at least 10 workers. The overall results do not change noteworthy. The estimated marginal effects for the wage compression (dispersion) variable are even statistically significant at higher levels than in the sample of larger firms, which can at least partly be attributed to the larger sample size that has increased to more than 4,000 firms in each year and to more than 2,000 firms in the balanced panel.

4. Conclusion

In this empirical paper, I have used German linked employer-employee data, which contain information about firms' cost coverage of training and allow to generate the conditional intra-firm wage dispersion as proxy for a firm's wage compression. The main finding of my econometric analysis is that firms with more compressed wage structures are more likely to cover all direct and indirect training costs. These findings are inconsistent with Becker's model of on-the-job training in perfect labor markets; but they are consistent with theoretical considerations of the "new training literature" that firms can capture rents from training due to wage compression in imperfect labor markets, which provides incentives for firms to pay for training. Moreover, it seems as

if union bargained collective contracts have no significant direct effects on training cost coverage that go beyond the effects of unions on general wage compression, whereas the existence of a works council is rather positively correlated with complete cost coverage, even after controlling for differences in firms' wage structures. Thus, codetermination at the firm-level seems to be more important than union bargaining when it comes to strategic training decisions in firms, which accords with the explicit role of works councils in firms' training practices stated in the German Works Constitution Act ("Betriebsverfassungsgesetz").

Three caveats are in order with respect to my empirical analysis, which leave scope for future research. First, the presented results might still suffer from omitted variable bias and reverse causality issues. In order to deal with those endogeneity problems, longer panel data sets and better instrumental variables would be helpful. The applied IV approach in this paper did however not indicate problems of endogeneity. Second, although I use a linked employer-employee data set to compute variables for the intra-firm wage compression, the data comprise training information only at the aggregated firm-level and not for individual workers. Therefore, my analysis could not account for worker heterogeneity with respect to differences in training cost coverage. Third, the focus of my analysis is on testing one core element of the "new training literature", namely the positive effect of wage compression on training cost coverage by firms. In order to provide concrete policy recommendations for stimulating human capital investments, *"in future work, the link between these stories and training can be more carefully derived, yielding empirical predictions to determine which sources of wage compression, if any, are important in encouraging firm-sponsored training."* (Acemoglu and Pischke, 1999b, p. 567). My finding that firms with union bargained

collective contracts have significantly lower wage dispersion (see first stage regressions in Table 3 and Table 4) shows that unions influence firms' wage structures. Together with previous empirical findings by Beckmann (2002a, 2002b) and Dustmann and Schönberg (2009) for apprenticeship training in German firms, unions are likely to be one important factor in the context of stimulating human capital investments, even if their effect might run through the indirect channel of compressed wage structures.

References

- Acemoglu, D., Pischke, J.-S. (1998). Why do firms train? Theory and evidence. *Quarterly Journal of Economics* 113(1), 79-119.
- Acemoglu, D., Pischke, J.-S. (1999a). Beyond Becker: training in imperfect labour markets. *Economic Journal* 109, F112-F142.
- Acemoglu, D., Pischke, J.-S. (1999b). The structure of wages and investment in general training. *Journal of Political Economy* 107(3), 539-572.
- Alda, H., Bender, S., Gartner, H. (2005). The linked employer-employee dataset created from the IAB establishment panel and the process-produced data of the IAB (LIAB). *Schmollers Jahrbuch (Journal of Applied Social Science Studies)* 125(2), 327–336.
- Allaart, P., Bellmann, L., Leber, U. (2009). Company-provided further training in Germany and the Netherlands. *Empirical Research in Vocational Education and Training* 1(2), 103-121
- Almeida-Santos, F., Mumford, K. (2005). Employee training and wage compression in Britain. *Manchester School* 73(3), 321-342.
- Asplund, R. (2005). The provision and effects of company training: a brief review of the literature. *Nordic Journal of Political Economy* 31, 47-73.
- Becker, G.S. (1962). Investment in human capital: a theoretical analysis. *Journal of Political Economy* 70(5, part 2), 9-49.
- Beckmann, M. (2002a). Lohnstrukturverzerrung und betriebliche Ausbildung: Empirische Analyse des Acemoglu-Pischke-Modells mit Daten des IAB-Betriebspanels. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 35(2), 189-204.
- Beckmann, M. (2002b). Wage compression and firm-sponsored training in Germany: empirical evidence for the Acemoglu-Pischke Model from a zero-inflated count data model. *Konjunkturpolitik* 48(3/4), 368-389.
- Bellmann, L., Hohendanner, C., Hujer, R. (2010). Determinants of employer-provided further training: a multi-level approach. *IZA Discussion Paper No. 5257*.
- Bellmann, L., Düll, H. (2001). Die zeitliche Lage und Kostenaufteilung von Weiterbildungsmaßnahmen: Empirische Ergebnisse auf der Grundlage des IAB-Betriebspanels. In: Dobischat, R., Seifert, H. (Eds.), *Lernzeiten neu organisieren: lebenslanges Lernen durch Integration von Bildung und Arbeit*, Forschung aus der Hans-Böckler-Stiftung 32, pp. 81-128.
- Booth, A.L., Bryan, M.L. (2005). Testing some predictions of human capital theory: new training evidence from Britain. *Review of Economics and Statistics* 87(2), 391-394.

- Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies* 47(1), 225-238.
- Chang, C., Wang, Y. (1996). Human capital investment under asymmetric information: the Pigovian conjecture revisited. *Journal of Labor Economics* 14(3), 505-519.
- Düll, H., Bellmann, L. (1998). Betriebliche Weiterbildungsaktivitäten in West- und Ostdeutschland: Eine theoretische und empirische Analyse mit den Daten des IAB-Betriebspanels 1997. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 31(2), 205-225.
- Düll, H., Bellmann, L. (1999). Der unterschiedliche Zugang zur betrieblichen Weiterbildung nach Qualifikation und Berufsstatus: Eine Analyse auf der Basis des IAB-Betriebspanels 1997 für West- und Ostdeutschland. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 32(1), 70-84.
- Dustmann, C., Schönberg, U. (2009). Training and union wages. *Review of Economics and Statistics* 91(2), 363-376.
- Eckaus, R.S. (1963). Investment in human capital: a comment. *Journal of Political Economy* 71(5), 501-504.
- Ericson, T. (2008). The effects of wage compression on general and firm-specific training. *Applied Economics Letters* 15(3), 165-169.
- Gerlach, K., Hübler, O., Meyer, W. (2002). Investitionen, Weiterbildung und betriebliche Reorganisation. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 35, 546-565.
- Gerlach, K., Jirjahn, U. (2001). Employer provided further training: evidence from German establishment data. *Schmollers Jahrbuch* 121, 139-164.
- Goerlitz, K. (2010). The development of employers' training investments over time: a decomposition analysis using German establishment data. *Journal of Economics and Statistics* 230(2), 186-207.
- Goerlitz, K., Stiebale, J. (2011). The impact of product market competition on employers' training investments: evidence from German establishment panel data. *De Economist* 159, 1-23.
- Katz, E., Ziderman, A. (1990). Investment in general training: the role of information and labour mobility. *Economic Journal* 100(403), 1147-1158.
- Leber, U. (2000). Finanzierung der betrieblichen Weiterbildung und die Absicherung ihrer Erträge: Eine theoretische und empirische Analyse mit Daten des IAB-Betriebspanels 1999. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 33(2), 229-241.
- Leuven, E. (2005). The economics of private sector training: a survey of the literature. *Journal of Economic Surveys* 19(1), 91-111.
- Loewenstein, M.A., Spletzer, J.R. (1998). Dividing the costs and returns to general training. *Journal of Labor Economics* 16(1), 142-171.

- Loewenstein, M.A., Spletzer, J.R. (1999). General and specific training: evidence and implications. *Journal of Human Resources* 34(4), 710-733.
- Mahy, B., Rycx, F., Volral, M. (2011). Does wage dispersion make all firms productive? *Scottish Journal of Political Economy* 58(4), 455-489.
- Mohrenweiser, J., Zwick, T. (2009). Why do firms train apprentices? The net cost puzzle reconsidered. *Labour Economics* 16(6), 631-637.
- Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica* 46(1), 69-85.
- Pischke, J.-S. (2001). Continuous training in Germany. *Journal of Population Economics* 14(3), 523-548.
- Pischke, J.-S. (2005). Labor market institutions, wages, and investment: review and implications. *CESifo Economic Studies* 51(1), 47-75.
- Rivers, D., Vuong, Q.H. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics* 39(3), 347-366.
- Stegmaier, J. (2010). Empirische Analysen zur betrieblichen Weiterbildung unter besonderer Berücksichtigung der Betriebsgröße. PhD Thesis.
- Winter-Ebmer, R., Zweimüller, J. (1999). Intra-firm wage dispersion and firm performance. *Kyklos* 52(4), 555-572.
- Wooldridge, J.M. (2010). *Econometric analysis of cross section and panel data*. 2nd edition. MIT Press: Cambridge, London.

Tables to be included in text

Table 1: Variable definitions and descriptive statistics for firm characteristics

Variables	Definitions	<u>Year 2005</u>		<u>Year 2007</u>		<u>Balanced panel</u>	
		<u>(n=2118)</u>		<u>(n=2011)</u>		<u>(n=2*1136=2272)</u>	
		Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Dependent variable: <i>COSTCOV</i>	Firm covers completely all direct and indirect training costs (dummy)	0.5184		0.4843		0.5158	
Wage variables: <i>logWSERT</i>	Intra-firm standard error of log daily gross wage regressions for full-time workers obtained from Tobit	0.2187	0.0721	0.2215	0.0737	0.2205	0.0731
<i>logWMEAN</i>	Intra-firm mean of log daily gross wages of full-time workers	4.5156	0.2933	4.5119	0.3229	4.5368	0.2933
Control variables (X): Union	Firm is bounded to union bargained collective contract (dummy)	0.7908		0.7454		0.7879	
Works council	Firm has a works council (dummy)	0.8428		0.8091		0.8556	
Number of quits	Number of quits during first half of survey year	3.8381	13.3393	5.0189	16.2395	4.0603	13.1641
Number of layoffs	Number of layoffs during first half of survey year	3.9835	16.7185	3.3819	14.0750	3.6932	16.6723
Number of workers /1000	Number of workers at June 30 th / 1000	0.6496	1.8764	0.5908	1.8959	0.6514	2.0746
Firm age <6 years	Firm younger than 6 years (dummy, reference)	0.0548		0.0542		0.0370	
Firm age 6-15 years	Firm age between 6 and 15 years (dummy)	0.2441		0.1631		0.1888	
Firm age >15 years	Firm older than 15 years (dummy)	0.7011		0.7827		0.7742	
Production technology	State of the art of the production technology (0-5; 0: newest, 5: outdated)	2.0132	0.7095	1.9866	0.7020	1.9859	0.6920
Profit situation	At least good profit situation (subjective perception) in last business year (dummy)	0.4164		0.5763		0.5040	
Share women	Share of female workers	0.3460	0.2452	0.3549	0.2532	0.3336	0.2377
Share part-time	Share of part-time workers	0.1294	0.1770	0.1405	0.1913	0.1226	0.1695
Share qualified	Share of qualified workers (at least apprenticeship or college degree)	0.7260	0.2479	0.7460	0.2453	0.7404	0.2370

Table 2: Probit regressions for complete training cost coverage by firms

	Year 2005: Cross-section Probit	Year 2007: Cross-section Probit	Balanced panel: Random effects Probit
<i>logWSERT</i>	-0.4939*** (0.1712) [p=0.004]	-0.3615** (0.1712) [p=0.035]	-0.3744* (0.2104) [p=0.075]
<i>logWMEAN</i>	0.1129* (0.0586)	0.1026* (0.0561)	0.1555** (0.0764)
Union	0.0150 (0.0309)	-0.0166 (0.0295)	0.0181 (0.0386)
Works council	0.0538 (0.0363)	0.0856** (0.0356)	0.0390 (0.0487)
Number of quits	-0.0014 (0.0011)	-0.0002 (0.0009)	0.0007 (0.0012)
Number of layoffs	-0.0020** (0.0009)	-0.0010 (0.0010)	-0.0019* (0.0011)
Number of workers /1000	-0.0016 (0.0071)	0.0003 (0.0067)	-0.0008 (0.0075)
Firm age 6-15 years	-0.0711 (0.0558)	0.0246 (0.0596)	-0.0165 (0.0770)
Firm age >15 years	-0.1290** (0.0510)	-0.0115 (0.0540)	-0.0458 (0.0731)
Production technology	0.0188 (0.0165)	0.0251 (0.0173)	0.0400* (0.0204)
Profit situation	0.0150 (0.0236)	-0.0246 (0.0243)	0.0046 (0.0280)
Share women	0.1053 (0.0763)	0.0243 (0.0784)	0.1305 (0.1029)
Share part-time	0.1795* (0.0956)	0.1831** (0.0897)	0.2457** (0.1238)
Share qualified	-0.0833 (0.0560)	-0.0475 (0.0575)	-0.1037 (0.0744)
Controls for federal states (16) and industries (15)	Yes	Yes	Yes
Pseudo R ² (McFadden)	0.0628	0.0635	
Number of observations	2118	2011	2272
Mean dependent variable	0.5184	0.4843	0.5158

Notes: Marginal effects at the means of all covariates in each estimation sample on the probability of complete training cost coverage by the firm (*COSTCOV*); binary Probit regressions for 2005 and 2007; random effects Probit regressions for balanced panel. All regressions include control variables (*X*) as described in Table 1, 16 federal state and 15 industry dummies. The random effects Probit regression further includes a dummy variable for the year 2007. Standard errors (robust for cross-section Probits) in parentheses. Significant at * p<0.10, ** p<0.05, *** p<0.01.

Table 3: IV Probit regressions for complete training cost coverage by firms (IV: $\log WMINIMUM$ and $\log WSERTis$)

	Year 2005: IV Probit			Year 2007: IV Probit		
	1st stage	2nd stage	mfX	1st stage	2nd stage	mfX
$\log WMINIMUM$	-0.0467*** (0.0018)			-0.0471*** (0.0020)		
$\log WSERTis$	0.6601*** (0.0592)			0.6543*** (0.0611)		
$\log WSERT$		-1.1486* (0.6508)	-0.4577* (0.2593)		-1.0598 (0.6641)	-0.4225 (0.2647)
$\log WMEAN$	0.0152* (0.0083)	0.2864* (0.1483)	0.1141* (0.0591)	0.0093 (0.0062)	0.2514* (0.1420)	0.1002* (0.0566)
Union	-0.0206*** (0.0030)	0.0399 (0.0784)	0.0159 (0.0313)	-0.0202*** (0.0030)	-0.0449 (0.0750)	-0.0179 (0.0299)
Works council	-0.0023 (0.0038)	0.1340 (0.0914)	0.0534 (0.0364)	-0.0039 (0.0039)	0.2170** (0.0911)	0.0865** (0.0363)
Number of quits	0.0002** (0.0001)	-0.0035 (0.0027)	-0.0014 (0.0011)	-0.00001 (0.0001)	-0.0005 (0.0021)	-0.0002 (0.0008)
Number of layoffs	0.0002*** (0.00005)	-0.0050** (0.0022)	-0.0020** (0.0009)	-0.0001 (0.0001)	-0.0025 (0.0024)	-0.0010 (0.0010)
Number of workers /1000	-0.0055*** (0.0012)	-0.0039 (0.0178)	-0.0016 (0.0071)	-0.0049*** (0.0008)	0.0008 (0.0168)	0.0003 (0.0067)
Firm age 6-15 years	0.0008 (0.0052)	-0.1779 (0.1403)	-0.0709 (0.0559)	0.0004 (0.0059)	0.0621 (0.1494)	0.0247 (0.0596)
Firm age >15 years	-0.0012 (0.0049)	-0.3264** (0.1314)	-0.1301** (0.0524)	0.0080 (0.0052)	-0.0269 (0.1357)	-0.0107 (0.0541)
Production technology	-0.0013 (0.0015)	0.0478 (0.0415)	0.0190 (0.0165)	0.0010 (0.0018)	0.0624 (0.0433)	0.0249 (0.0173)
Profit situation	-0.0005 (0.0023)	0.0370 (0.0591)	0.0148 (0.0236)	-0.0034 (0.0025)	-0.0621 (0.0610)	-0.0247 (0.0243)
Share women	0.0360*** (0.0089)	0.2596 (0.1931)	0.1034 (0.0769)	0.0414*** (0.0081)	0.0693 (0.1989)	0.0276 (0.0793)
Share part-time	0.0192 (0.0122)	0.4499* (0.2399)	0.1793* (0.0956)	0.0165 (0.0115)	0.4586** (0.2251)	0.1828** (0.0897)
Share qualified	0.0083 (0.0058)	-0.2107 (0.1409)	-0.0840 (0.0561)	0.0012 (0.0063)	-0.1167 (0.1444)	-0.0465 (0.0576)
Controls for federal states (16) and industries (15)	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.1519*** (0.0386)	-0.0842 (0.7765)		0.1725*** (0.0306)	-0.3225 (0.7675)	
$\hat{\epsilon}_j^{first} / \hat{\sigma}_{\epsilon}^{first}$		-0.0080 (0.0419)			0.0137 (0.0442)	
Number of observations	2118	2118	2118	2011	2011	2011

Notes: The first stage is estimated with linear regressions that use the lowest observed daily wage in a firm ($\log WMINIMUM$) and the mean of $\log WSERT$ within industry and federal state cells ($\log WSERTis$) as instruments for firms' intra-firm standard error of log daily wage regressions ($\log WSERT$). The second stage is estimated with binary Probit regressions that include the standardized predicted error terms for every firm from the first stage regressions ($\hat{\epsilon}_j^{first} / \hat{\sigma}_{\epsilon}^{first}$). Marginal effects at the means of all covariates in each estimation sample on the probability of complete training cost coverage by the firm ($COSTCOV$) are presented in the third column for every year. All regressions include control variables (X) as described in Table 1, 16 federal state and 15 industry dummies. Robust standard errors in parentheses. Significant at * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: IV Probit regressions for complete training cost coverage by firms (IV: $\log WMINIMUM$)

	Year 2005: IV Probit			Year 2007: IV Probit		
	1st stage	2nd stage	mfX	1st stage	2nd stage	mfX
$\log WMINIMUM$	-0.0500*** (0.0019)			-0.0494*** (0.0020)		
$\log WSERT$		-1.1420* (0.6928)	-0.4551* (0.2761)		-0.9608 (0.6910)	-0.3830 (0.2754)
$\log WMEAN$	0.0180** (0.0088)	0.2867* (0.1485)	0.1143* (0.0592)	0.0096 (0.0064)	0.2553* (0.1422)	0.1018* (0.0567)
Union	-0.0237*** (0.0032)	0.0401 (0.0787)	0.0160 (0.0314)	-0.0221*** (0.0031)	-0.0429 (0.0750)	-0.0171 (0.0299)
Works council	-0.0032 (0.0041)	0.1339 (0.0915)	0.0534 (0.0364)	-0.0035 (0.0041)	0.2165** (0.0911)	0.0863** (0.0363)
Number of quits	0.0002** (0.0001)	-0.0035 (0.0027)	-0.0014 (0.0011)	-0.00001 (0.0001)	-0.0005 (0.0021)	-0.0002 (0.0009)
Number of layoffs	0.0002*** (0.0001)	-0.0050** (0.0022)	-0.0020** (0.0009)	-0.00004 (0.0001)	-0.0025 (0.0024)	-0.0010 (0.0010)
Number of workers /1000	-0.0058*** (0.0013)	-0.0039 (0.0178)	-0.0016 (0.0071)	-0.0052*** (0.0009)	0.0007 (0.0168)	0.0003 (0.0067)
Firm age 6-15 years	-0.0010 (0.0057)	-0.1779 (0.1403)	-0.0709 (0.0559)	-0.0014 (0.0062)	0.0619 (0.1494)	0.0247 (0.0595)
Firm age >15 years	-0.0018 (0.0055)	-0.3264** (0.1314)	-0.1300** (0.0524)	0.0064 (0.0055)	-0.0281 (0.1357)	-0.0112 (0.0541)
Production technology	-0.0016 (0.0016)	0.0478 (0.0416)	0.0191 (0.0166)	0.0012 (0.0019)	0.0628 (0.0433)	0.0250 (0.0173)
Profit situation	-0.0013 (0.0024)	0.0370 (0.0591)	0.0147 (0.0236)	-0.0037 (0.0026)	-0.0619 (0.0610)	-0.0247 (0.0243)
Share women	0.0408*** (0.0094)	0.2592 (0.1934)	0.1033 (0.0771)	0.0429*** (0.0084)	0.0639 (0.1990)	0.0255 (0.0793)
Share part-time	0.0176 (0.0132)	0.4500* (0.2399)	0.1793* (0.0956)	0.0173 (0.0121)	0.4590** (0.2251)	0.1830** (0.0897)
Share qualified	0.0081 (0.0063)	-0.2109 (0.1409)	-0.0840 (0.0561)	0.0039 (0.0065)	-0.1183 (0.1445)	-0.0471 (0.0576)
Controls for federal states (16) and industries (15)	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.3068*** (0.0376)	-0.0868 (0.7819)		0.3071*** (0.0301)	-0.3569 (0.7725)	
$\hat{\epsilon}_j^{first} / \hat{\sigma}_{\epsilon^{first}}$		-0.0082 (0.0446)			0.0047 (0.0461)	
Number of observations	2118	2118	2118	2011	2011	2011

Notes: The first stage is estimated with linear regressions that use the lowest observed daily wage in a firm ($\log WMINIMUM$) as an instrument for the intra-firm standard error of log daily wage regressions ($\log WSERT$). The second stage is estimated with binary Probit regressions that include the standardized predicted error terms for every firm from the first stage regressions ($\hat{\epsilon}_j^{first} / \hat{\sigma}_{\epsilon^{first}}$). Marginal effects at the means of all covariates in each estimation sample on the probability of complete training cost coverage by the firm ($COSTCOV$) are presented in the third column for every year. All regressions include control variables (X) as described in Table 1, 16 federal state and 15 industry dummies. Robust standard errors in parentheses. Significant at * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.