Working Paper No. 116

The more they spend, the more I earn? Firms' training investments and post-training wages of apprentices

Hans Dietrich, Harald Pfeifer and Felix Wenzelmann

Universität Zürich
IBW – Institut für Betriebswirtschaftslehre

UNIVERSITÄT BERN

Working Paper No. 116

The more they spend, the more I earn? Firms' training investments and post-training wages of apprentices

Hans Dietrich, Harald Pfeifer and Felix Wenzelmann

March 2016

The Swiss Leading House on Economics of Education, Firm Behavior and Training Policies is a Research Program of the Swiss State Secretariat for Education, Research, and Innovation (SERI).

www.economics-of-education.ch

The more they spend, the more I earn? Firms' training investments and post-training wages of apprentices

Hans Dietrich [a]
Harald Pfeifer [b, c]
Felix Wenzelmann [b]

[a] *Institute for Employment Research (IAB), Nuremberg*
[b] *Federal Institute for Vocational Education and Training (BIBB), Bonn*
[c] *Research Centre for Education and the Labour Market (ROA), Maastricht*

Abstract

In this paper, we examine the relation between a firm's training investment and the post-training wages of apprenticeship graduates. For our analysis, we first calculate a training investment indicator using detailed information about firm-level training costs. We then merge the firm-level data with individual-level administrative data on employment and wages of apprenticeship graduates. Using regression models controlling for selection into employment, we find that a firm investment in training relates positively with graduates' post-training wages. Doubling a firm's training investment leads to a wage mark-up of about 2.8%. This result is robust to different specifications. However, we find that especially graduates from low-investment firms benefit from a higher training investment. The wage mark-up for graduates from firms with already high investment levels is small and statistically not significant.

JEL Classification: J24, J31, J62
Keywords: Training investment, post-training wages, apprenticeship system

1 Introduction[1]

Within the German apprenticeship system, firms are a major provider of education and training. Despite the high degree of curricula standardization, firms are fairly flexible in the way to train the skills outlined in the curricula. For example, firms may delegate the training task to regular employees instead of paying professional trainers to organize their training. They may finance a separate training centre, in which apprentices practice their skills or organize additional in-house classroom teaching to extend theoretical knowledge about the occupation. Firms may also train mainly on-the-job, developing the skills of the apprentice through work experience in real production environments.

Because the different training methods applied are directly linked to the firm's costs of training, empirical studies find a large variance of training costs across firms.[2] This raises the question in which way apprentices' labour market outcomes are related to the firm's cost of training. On the one hand, one might expect that apprentices trained in a high-investment environment receive a wage bonus on the labour market due to the potentially higher quantity or quality of human capital that is generated during training. On the other hand, human capital theory stresses that firms need to recoup their training costs through lower wages during the subsequent employment phase. While the former argument implies a positive relation between the size of a firm's training investment and post-training wages, the latter argument implies a negative relationship.

In this paper, we aim to answer this question empirically. We first construct an indicator of firms' training investment (*FTI*), which is calculated on the basis of detailed firm-level survey information on the costs of training. The survey data is then merged with individual-level administrative data by using the firm identification number included in both data sources. The merging of data sets allows us to observe post-training employment characteristics and wages for those apprentices that have previously been trained in the surveyed firms.

[1] We thank Holger Alda, Uschi Backes-Gellner, Christian Grund, Gudrun Schönfeld, Thomas Siedler, Gesine Stephan and the participants of the EHB VET-Congress 2013, POEK Conference 2013 and AIEL Conference 2015 for helpful comments.

[2] See Schönfeld et al. (2010) for a detailed description of firms training costs in Germany.

The main contribution of this paper to the literature is the analysis of the relation between firms' training investment and the post-training wages of apprenticeship graduates. Whereas wage outcomes of apprentices have been studied extensively in the literature, the role of firms' prior training investment has not been analysed empirically.

Our results support the hypothesis that higher firm investments in training are positively related to post-training wages of apprenticeship graduates. An increase in firms' training investments of 100% leads to a wage mark-up of about 2.8%. However, we also find considerable heterogeneity in this relation, as wages increase by around 4% for graduates coming from low-investment firms, while the coefficient for high-investment firms remains insignificant. This result suggests that, beyond a certain threshold of investments, apprenticeship graduates do not benefit from additional resources dedicated to their training. Conversely, graduates that engage in training in the segment of low-investment firms have considerable wage gains should these firms be inclined to invest more in their training.

The remainder of this paper is structured as follows. Section 2 summarizes recent literature and discusses their theoretical implications. Section 3 explains the construction of the data set and the variables used in the regression analysis. Section 4 discusses potential selection issues and develops a corresponding estimation strategy. Section 5 presents the results of the regression analysis and provides alternative estimates and a robustness analysis. Section 6 sums up the results and discusses prospects for further research.

2 Literature

The relation between education and wages is a well-studied topic in economics. Since the seminal works of Becker (1962, 1964), Mincer (1958, 1974) and Schultz (1963, 1971), a series of papers attempts to explain why firms and individuals invest in education and training and who benefits in terms of higher productivity and wages.

With respect to the investment of firms in the training of workers, classical human capital theory (Becker 1964) stresses that firms should not invest in general training at all, because it is unlikely that they recoup the training costs in the aftermath of the training through wages below productivity. In contrast to this notion, Acemoglu and Pischke (1998, 1999a,b) argue that firms may have incentives to pay for the training because the rent received afterwards compensates for costs in the presence of a compressed wage structure. The authors refer to the German apprenticeship system, for which several firm-level surveys provide empirical evidence for substantial investments in essentially general training (von Bardeleben et al. 1995, Beicht et al. 2004, Schönfeld et al. 2010).

Despite the extensive number of studies on the training investment motivation of firms and individuals, little is known about the impact of the *size* of the firm investment on labour market outcomes of trainees. Blundell (1999) gives an overview of estimates on the returns of different education and training qualifications in the UK. Most of the studies find positive wage returns to education, reaching from just about zero percent for a lower level qualification training course to more than 18 percent for a university degree.

With respect to vocational training, several studies have estimated the wage returns to *continuing* vocational training in the form of on-the-job training courses (see Bassanini et al. 2005 for an overview of literature in this field). Some of these studies control for the time the individual worker spends on training. Bartel (1995), for example, uses personnel records of a large manufacturing firm in the US to find a persistent wage effect of the volume (days) of worker training courses. Likewise, Büchel and Pannenberg (2004) find significant effects of the number and duration of continuing training courses in Germany.

Wage return estimates for apprenticeship training are less frequent in the literature. Riphahn and Zibrowius (2015) analyse wage effects of apprenticeship training using instrumental variable regressions and find large wage premiums for obtaining an apprenticeship training qualification. They further discuss previous literature, mostly providing support for positive wage returns to this form of training.

Von Wachter and Bender (2006) use linked employer-employee data to show that apprentices moving to another firm after training experience a (temporary) wage penalty compared to apprentices staying in the training firm.

While the latter studies focus on wage outcomes of apprenticeship graduates, our interest lies in the relation between the *size* of a firm's training investment and the post-training wage of apprentices. To the best of our knowledge, no study so far has used micro data combining firm-level investment information and administrative individual-level wage records to analyse this relationship. The unique data set is described in more detail in the next section.

3 Data sources and variable construction

3.1 Data sources

The information about training investments of firms stems from the BIBB Cost-Benefit Survey 2007 (in the following CBS). About 3,000 training firms report the costs borne for the training of their apprentices. The interviews were conducted by the survey institute *infas* using the computer assisted personal interview (CAPI) method. Interview partner was the person responsible for the organization of training in the firm and/or the firm's human resources manager. In small firms, the interview partner often was the owner or managing director of the firm.

The sample was drawn from the 'Establishment Register' of the Federal Employment Agency in which all firms with at least one employee subject to social security contributions are registered. The 'Establishment Register' allows a representative sampling of German training firms. In case the interviewed firms agreed on the data matching question (which was the case for a representative subsample of 2,000 firms[3]) it was possible to combine firm specific cost data with

[3] A detailed analysis about the potential selectivity of firms and on the match quality of survey and administrative data is provided in Dietrich et al. (2014). The authors find that the subset of firms having given the permission for the match is not different from firms that did not, contingent on a large set of observable characteristics. Further, the authors test for the quality of the matching by comparing the surveyed number of employees and apprentices with the corresponding number of employees and apprentices from the administrative data source. They conclude that the match quality is satisfactory

the employment records of individuals completing apprenticeship training within the interviewed firms. For that group of apprentices, a data excerpt was drawn from the administrative records[4]. The data excerpt includes variables describing the apprenticeship training and the labour market status of the graduates one year after apprenticeship training. More specifically, variables included are wages, employment status and the training history of the apprentices plus a set of demographic variables including the apprentice's age, sex and nationality.

We further generate variables identifying the employment background of the individual before the start of the apprenticeship, such as prior training history, unemployment duration and schooling level (graduation from high-school - *Abitur*). The data also contains information that refers to the training phase itself and that consequently are assumed to be relevant for post-training labour market integration. These variables are e.g. the apprentice's training pay and the exact duration of training until graduation. We further observe whether the apprentice receives complementary scheme-based support during the training phase, which can be applied for in case the apprentice has learning difficulties. Especially the latter variable can be interpreted as an indicator for ability, because the scheme intends to support weak performers during training.

Finally, we draw on administrative data on the apprentice's labour market integration after training. Our strategy is to use the employment and wage information for the apprenticeship graduate one year after the end of the training period as outcome variables for equations discussed in Section 4. One reason for focusing on wages one year after graduation is that data beyond this period is not available at the time of writing the paper. However, observing a wage early in the labour market career also has advantages over later observations because the "pure" effect of training investments could be blurred by other events (e.g. by continuing training, on-the-job learning or career changes), which would need to be accounted for in the regression framework.

---

for over 90% of the firms and comparable to the already existing survey-administrative data matches (e.g. the IAB Linked Employer-Employee Data Set, the LIAB).

[4] For a more detailed description of the administrative data see vom Berge et al. (2013).

For the analysis, we use information of about three quarters of the matched training firms. The reason for the smaller size of the sample is that we do not observe apprentices finishing training for each firm. The reference period of the survey is 2007 and we only observe apprentices having graduated between 2006 and 2008. The missing firms are thus firms where apprentices have been in the first or second training year in 2007 and thus do not enter the labour market until 2009 or later - a date past our observation period. The reason for choosing apprenticeship graduates of the years from 2006 to 2008 is that we assume the training investment indicator to be most valid for the periods around the reference date of 2007.

Because we have no information on the exact weekly working time of the individuals one year after training, we only include full time workers in the group of employed workers. Hence, about 600 part time workers are excluded from the sample. We also exclude those individuals that start another apprenticeship after having graduated from the observed one. From the remaining 11,323 apprenticeship graduates that have been trained in the 1,597 training firms, 2,206 are registered unemployed, in a public financed program or absent from the labour market one year after the end of the training[5], which leaves us with 9,117 employed apprenticeship graduates, for which we obtain information about daily gross wage.

3.2 Variable construction

The main explanatory variable of interest is the indicator on training investments (*FTI*) of firm *j*, which is calculated as a firm-level average per apprentice and year of training. The training investment indicator FTI_j consists of three components. The first component, m_j, includes training material, such as books, clothing and instruments. The second component, k_j, is a measure for the

[5] In the official statistics this status is marked as "gap". Included in these groups are e.g. those individuals attending university, traveling or doing their military service. As a robustness check we excluded those individuals from the sample (see Section 5.4).

costs of training infrastructure, which includes separate internal training centres, in-house classroom teaching and training machines.[6]

The third component takes account of the costs for training personnel, which is calculated by multiplying the average hours of full- and part-time trainers per apprentice (th_j) with their respective hourly wage (tw_j).[7] Whereas full-time trainers dedicate most of their time to training, part-time trainers may instruct apprentices while performing their regular job. Thus, part-time trainers may still be productive during the training.[8] Because we are interested in the *net* investment of training, we need to calculate the costs resulting from the actual productivity gap of part-time trainers. The survey therefore directly inquires about the degree to which full- and part-time trainers are unproductive.[9] Because part-time trainers can have different functions in the firm, they are likely to earn different wages and thus incur different costs when training apprentices. The survey therefore provides wages, training hours and productivity loss separately for different groups of workers (i.e. for unskilled workers, skilled workers and managers). The correction factor for trainers' productivity is denoted by δ_j. The training investment (*FTI*) for an average apprentice in a firm is thus determined by

$$FTI_j = \delta_j \, (th_j * tw_j) + m_j + k_j \qquad (1)$$

Apart from the main explanatory variable *FTI*, we control for a set of individual and firm-level characteristics. Individual characteristics for apprentices include age (one year after the end of apprenticeship), gender, nationality and level of secondary education (high school degree – *Abitur* - or not). We further

[6] Note that we do not include costs for apprentice pay in the indicator, since we do not consider these costs to be relevant for human capital formation

[7] Note that average training hours per apprentice are calculated by dividing the total training hours by the number of apprentices.

[8] To illustrate this argument, consider the example of a painter, who brushes and paints a wall while explaining the techniques to the apprentice. During this activity, the part-time trainer (i.e. the painter) is fully productive although actively training the apprentice.

[9] The respective questions were formulated as: "How many hours in total did the fulltime (part-time) trainer spent on the training in an average week of 2007?" And: "How much lower was the productivity of the full-time (part-time) trainers (in the respective worker groups) during these training hours?"

observe individuals' employment related characteristics for example prior to training unemployment experience, the occupation of training and post-training characteristics, such as employer-change. The data set also provides an exact measure of the training duration, which not only varies by type of training occupation but also by the individual progress and success in the final exam. On the firm level, we control for structural variables, such as firm size, region and economic sector, as the literature documents relatively strong differences of net investments due to these characteristics.

For our wage equations, we use the logarithmic daily gross wage of those apprenticeship graduates that are employed one year after the apprenticeship. The reason for choosing a time gap of one year is that, in several firms, the retention of apprentices is obligatory for at least six months in continuation of the training. Thus, using the first registered employment and wage information would give a distorted picture of post-training employment and wages.

3.3 Description of relevant variables

Table 1 presents summary statistics for the variables of the estimation models. 81% of the apprenticeship graduates are full-time employed one year after the graduation from the apprenticeship. The daily gross wage of employed individuals averages 74 Euros. The kernel density estimate in Figure A1 in the appendix illustrates that daily wage is close to normally distributed across apprenticeship graduates.

A training firm invests on average about 8,900 Euro per apprentice and year of training. This leads up to around 25,500 Euro for the average apprenticeship in our sample of about 2.85 years. The distribution of investments is skewed to the right with a median of 6,766 Euro (see Figure A2 in the appendix). The retention rate, i.e. the share of apprentices working in the training firm one year after the graduation, is 57%. In our sample 35% of the employed individuals work in an occupation that differs from their original training occupation.

Table 1: Summary statistics of individual and firm variables

	Employed		Not employed		Full sample	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Individual-level variables:						
Daily gross wage (one year after training) in €	74.19	(22.97)				
Individual changed firm after apprenticeship	0.43	(0.49)				
Individual changed occupation after apprenticeship	0.35	(0.48)				
Employed one year after training					0.81	(0.40)
Received social benefits before apprenticeship	0.03	(0.16)	0.11	(0.31)	0.04	(0.20)
Years of unemployment before apprenticeship	0.07	(0.28)	0.14	(0.42)	0.08	(0.31)
Received scheme-based training support	0.02	(0.14)	0.06	(0.24)	0.03	(0.17)
Duration of apprenticeship (in years)	2.88	(0.58)	2.70	(0.65)	2.85	(0.60)
Apprentice achieved *Abitur*	0.24	(0.43)	0.13	(0.33)	0.22	(0.42)
Age of apprentice (years)	22.92	(2.60)	22.66	(2.54)	22.87	(2.59)
Male	0.59	(0.49)	0.65	(0.48)	0.60	(0.49)
Non-German	0.06	(0.23)	0.09	(0.28)	0.06	(0.24)
Firm-level variables:						
Training investment FTI (per year of training) in €	8990	(7616)	8597	(7603)	8913	(7615)
Manufacturing	0.38	(0.49)	0.31	(0.46)	0.37	(0.48)
Wholesale and retail trade	0.09	(0.28)	0.10	(0.30)	0.09	(0.29)
Services I*	0.22	(0.41)	0.23	(0.42)	0.22	(0.42)
Services II**	0.12	(0.33)	0.08	(0.28)	0.12	(0.32)
Administration/Education/Health	0.19	(0.39)	0.27	(0.45)	0.20	(0.40)
1-9 employees	0.04	(0.20)	0.07	(0.26)	0.05	(0.21)
10-49 employees	0.12	(0.32)	0.15	(0.36)	0.12	(0.33)
50-499 employees	0.37	(0.48)	0.49	(0.50)	0.39	(0.49)
500 and more employees	0.48	(0.50)	0.29	(0.45)	0.44	(0.50)
Observations	9,117		2,616		11,323	

*Services I: hotels and restaurants, transport and telecommunication, energy and water supply

**Services II: banking and insurance, real estate, renting and business activities.

With respect to the variables relevant for the selection model in section 5.2, about 4% of all apprentices received social benefits before the start of the training. 82% of the ones that did not receive benefits are employed, while of those having received benefits only 52% have a job. The group of the employed graduates differs in some characteristics from the group of the not employed.

Firms' training investment was slightly higher for the now employed than for the now unemployed graduates (9,000 vs. 8,600 Euros). Further, the now employed graduates were more often trained in large firms with 500 or more employees (48% vs. 29%), have shorter unemployment periods before the training (0.07 vs. 0.14 years), received less often scheme-based training support (2% vs. 6%) and more often achieved the high school degree *Abitur* (24% vs. 13%) than those not employed (see Table 1).

Concerning our restriction variable for the selection model, the descriptive statistics show that employed individuals received less often benefits before the apprenticeship (3% compared to 11%).

4 Estimation strategy

Works dealing with impact of education on earnings often face one main issue to be solved: Selectivity into and out of education (for an overview of studies and methods see Card 1999, 2001 or Wolter and Ryan 2011). In the case of German apprenticeship training, the issue is equally important. It is reasonable to assume that the selection into the training firm and, after training, the selection into employment may lead to a bias of post-training wage estimates. The main concern is that apprenticeship graduates may have unobservable characteristics that are correlated both with the choice of (or access to) the training firm and with the post-training wage received on the labour market. In addition, unobserved characteristics may influence the

decision of the firm to retain the apprenticeship graduate instead of releasing him onto the labour market (‘lemons’ problem). Any attempt to estimate the impact of the firm’s training investment on post-training wages needs to take these concerns into account.

In this paper, our strategy is to face these sources of possible selectivity in different steps. We start with a set of OLS regressions including only those individuals being employed one year after the training. In Model A we estimate the correlation between the training investment of the training firm and the daily gross wage one year after the training. In Model B, we add the in Section 3.2 described firm level control variables (such as firm size, economic sector or region) and the year of graduation. To control for the selection of apprentices into the training firm, we add the training vocation on a 1-digit level of the national occupation classification (KldB 2010). We further include the above described individual control variables (for example secondary schooling level or prior unemployment spells) in Model C.

In a last step we address selectivity into regular employment. Not all apprenticeship graduates are employed one year after training. One group is unemployed or out of the labour force for different reasons, such as military service, studying at university or going abroad. Another reason could be that the wage offers on the labour market are too low for the graduate. We thus assume that the apprentices being employed after training differ in their unobserved characteristics from those that are not employed. To address the selectivity problem we employ a two-step Heckman selection model (Heckman 1974, 1979). For the identification of the model, a exclusion restriction is required which is correlated with the probability of being employed one year after training but not with the post-training wage (Wooldridge 2002).

Our choice for an exclusion restriction is a variable indicating the receipt of household based social benefits (*Hartz IV*) prior to the apprenticeship training. We argue that the receipt of social benefits is an important indicator

for the individual employment probability of apprenticeship graduates and interpret this variable as a proxy for social background and social status. Based on a sociological standard theory of rational educational decisions, social background affects educational decisions and labour market placement under control of an individual's performance (Boudon 1974, Breen and Goldthope 1997, Stocké 2007). For our exclusion restriction to hold, we consequently argue that social benefit receipt is important for the selectivity into employment without being correlated with the wage obtained in the labour market, which is mainly related to the skills obtained during training. Thus, once the apprentice has reached employment status, social background (approximated by benefit receipt) has no influence on the wage received from employment.

We formalize the estimation model for the logarithmic wage one year after the apprenticeship (*wij*) of individual *i* in firm *j* as:

$$Ln(w_{ij}) = \alpha + \beta_1 ln(FTI_j) + \beta_2 P_{ij} + \beta_3 C_j + \epsilon_{ij}\,, \qquad (2)$$

whereas *FTI* is the training investment of firm *j* (see Section 3.2). *P* is a vector of person-level variables and *C* a vector for firm-specific variables, as described in this section.

In section 5.3, we discuss results of alternative specifications that aim to tackle the issue of firm retention of apprentices. We include the variables "firm change" and "occupation change" in the model. We follow the argumentation of von Wachter and Bender (2006) by including the differential of the retention rate in the graduation year from the firms' usual retention rate and the interaction term of this variable with the "firm change" variable. We also provide estimates that exclude the group of individuals being absent from the labour market, because this group is very heterogeneous and, in case of the military service, the absence is rather involuntary and not random. We finally

check for effect heterogeneity by splitting the sample into graduates from low- and high-investment firms. In all of the following regression models we use firm-clustered standard errors.

5 Empirical results

5.1 OLS regression models

Table 2 presents the three OLS regression models on the logarithm of the daily wage one year after the training for the 9,117 employed apprenticeship graduates. Model A, i.e., the bivariate model, shows that firms' training investments (as measured by *FTI*) are positively correlated with the wages apprentices receive one year after their training. Because both wage and training investment are introduced into the regression in a logarithmic form, we can interpret the coefficient as elasticity. Thus, according to Model A, a 100% increase of training investment implies a 10% increase of the wage one year after the training. Model B extends Model A by including firm specific controls (economic sector, size and region of the training firm), and training specific controls (year of graduation and training occupation)[10]. The coefficient of firms' training investments *FTI* decreases to one quarter of the value of Model A.

In Model C, we add a set of individual control variables (socio-demographics and variables on training duration and prior unemployment) into the regression. The training investment coefficient decreases further but remains positive and significant with a value of about 0.02.

With respect to our set of control variables, the coefficients of the individual variables are all of the expected sign. Men, natives and apprentices with a higher schooling level receive higher wages after an apprenticeship

[10] The training occupations were grouped on the one digit level according to the occupational classification of the Federal Employment Office.

compared to the respective comparison group. A longer period of unemployment before training is associated with a lower wage afterwards. We find the same association for individuals having received scheme-based training support during the apprenticeship. Concerning the measure of apprenticeship duration, we find a positive and significant coefficient of 0.066. The coefficient indicates that one additional year of apprenticeship training yields a wage mark up of 6.6%, which is a reasonable estimate considering the results found in the literature.

Table 2: OLS regressions on (ln) daily gross wage

	Model A	Model B	Model C
Firm training investment FTI (ln)	0.101***	0.026**	0.019**
	(0.019)	(0.011)	(0.009)
Age (years)			0.011
			(0.010)
Years of unemployment before apprenticeship			-0.070***
			(0.016)
Received scheme-based training support			-0.166***
			(0.030)
Duration of apprenticeship (in years)			0.066***
			(0.009)
High school degree (*Abitur*)			0.116***
			(0.014)
Male			0.070***
			(0.009)
Non-German			-0.065***
			(0.018)
Region	No	Yes	Yes
Training occupation	No	Yes	Yes
Firm size	No	Yes	Yes
Economic sector	No	Yes	Yes
Year of graduation	No	Yes	Yes
Constant	3.365***	3.827***	3.395***
	(0.162)	(0.098)	(0.088)
Observations	9117	9117	9117
Adjusted R^2	0.059	0.330	0.381

Firm-clustered standard errors in parentheses.
Significance levels: *<10%, **<5% and *** <1%.

5.2 Regression model controlling for selectivity

Table 3 provides the results of the regression model taking into account the potential selection of apprentices into employment (Heckman 1979). The highly significant ρ at the bottom of the regression table indicates that selection is an issue that has to be treated in our context. Analogously, the coefficient of our exclusion restriction variable is large and significantly negative, which confirms the assumption that social aid recipients are less likely to be employed even after graduating from an apprenticeship.

Table 3: Heckman selection model

Dependent variable:	Daily gross wage (ln)	Employed (yes/no)
Firm training investment FTI (ln)	0.028***	-0.035
	(0.008)	(0.022)
Age (years)	0.009***	0.030***
	(0.002)	(0.007)
Years of unemployment before apprenticeship	-0.038*	-0.118**
	(0.015)	(0.044)
Received scheme-based training support	-0.099**	-0.246*
	(0.033)	(0.099)
Duration of apprenticeship (in years)	0.031***	0.227***
	(0.009)	(0.031)
High school degree (*Abitur*)	0.091***	0.150***
	(0.013)	(0.044)
Male	0.088***	-0.105**
	(0.010)	(0.040)
Non-German	-0.040*	-0.050
	(0.017)	(0.057)
Received social benefits before apprenticeship		-0.362***
		(0.064)
Region	Yes	Yes
Training occupation	Yes	Yes
Firm size	Yes	Yes
Economic sector	Yes	Yes
Year of graduation	Yes	Yes
Constant	3.585***	-0.005
	(0.084)	(0.268)
ρ	-0.911	(0.015)
σ	0.337	(0.015)
Observations	9117	11323

2206 restricted cases. Firm-clustered standard errors in parentheses.
Significance levels: *<10%, **<5% and *** <1%.

The coefficients of the selection model shown in the second column of Table 3 imply that the individual age, the duration of the apprenticeship as well as the schooling level positively influence the probability of being employed one year after the training. The unemployment record prior to the training is, in contrast, negatively related with employment. In addition, social benefit receipt during training is a negative indicator for employment. Males are less likely to be employed, which might be driven by the specification that those doing military service (which is only obligatory for men) are included in the sample as not employed. For a robustness check we excluded this group (see Section 5.4). Concerning nationality, we find no significant influence on employment probabilities. The coefficient of the training investment is negative but only significant at the 10.8% level. Being trained in a firm that invests more in the training is therefore not an insurance for finding an employment.

The first column of Table 3 provides the coefficients of the variables that are included in the wage equation. Most importantly, we find that, in line with the OLS regressions, the firm's training investment *FTI* impacts positively on the post-training wage of apprenticeship graduates. The coefficient shows that an increase in the firm's yearly training investments by 100% leads to an increase in the apprentice's daily wage of about 2.8%.

Several of the remaining explanatory variables have significant influence on wages. Male workers and older workers receive higher wages than their respective reference group. Furthermore, the coefficient for workers with a higher schooling level (*Abitur*) is positive and significant. Conversely, non-German apprenticeship graduates and those who have received scheme-based support during training show negative coefficients indicating lower wages than their respective comparison groups. With respect to the duration of the apprenticeship, the selection model yields a lower coefficient of one additional year of training (3.1%) as compared to the OLS estimates (6.6%).

Summing up, our regression analysis confirms our main hypothesis in that the firm-level training investments have a positive impact on post-training wages, independently of the duration of the training.

5.3 Alternative specifications and robustness analysis

In this section, we provide a set of alternative specifications addressing further selectivity issues and effect heterogeneity.

With respect to the selectivity of apprenticeship graduates into their post-training employer, the literature pays considerable attention to the question whether graduates that leave the firm after training receive lower wages than graduates that stay their training firm. Several works find at least temporary wage losses for ‘movers’ compared to ‘stayers’ (for recent evidence see von Wachter and Bender 2006 or Wagner and Zwick 2012). Because we have to take into account that firms invest more in apprentices they plan to retain after training, not controlling for a firm-change of the apprentice could introduce an omitted variable bias in our estimates. A similar argument could be applied to the change of occupation after training: Apprentices from firms with high training investments could have a higher probability of finding a job in the trained occupation, even when they are not retained by the training firm.

To address this issue, we construct a variable that is closely related to the instrumental variable used by von Wachter and Bender (2006) to quantify wage losses by “mover” apprentices. Put in simple terms, the argument of von Wachter and Bender is that the deviation from a firm’s usual retention rate at the time of graduation is a valid instrument for the apprentice-specific probability of moving to another firm. The authors consider the deviation from the usual retention rate to be a shock to the retention rate due to unexpected changes in the labour demand.

For our purpose of controlling for post-training mobility in the estimation model, we construct a variable indicating whether a firm’s retention

rate *at the time of graduation* is below, around or above the average retention rate of the past three years. This variable enters the OLS regression model shown in Table 2, together with the indicators of firm change and occupation change of the apprenticeship graduate.[11] The result of this exercise is displayed in Table B1 in the appendix. Column 1 reports the results of including firm change, occupation change and the deviation from the usual retention rate separately. Column 2 additionally includes the interaction term of the deviation from the usual retention rate and firm change of the graduate.

The results indicate that both firm change and occupation change are associated with a wage penalty for the apprenticeship graduate. Also the respective coefficient for the deviation from the firm's average retention rate is negative and significant, indicating that graduates that were retained in firms with an exceptionally high labour demand experience a wage loss. Further, also the interaction term in Column 2 is large and significantly negative, indicating that apprenticeship graduates *leaving* a firm despite a *higher* retention probability earn considerably lower wages than graduates that are *retained* despite a *lower* retention probability.

The aim of this exercise is, however, to provide complementary evidence that the mobility of apprenticeship graduates does not significantly change the relation between our training investment indicator *FTI* and post-training wages of the (moving or staying) graduates. Indeed, the *FTI* coefficient is only slightly lower than the coefficient from the favoured selection model (Table 3) and slightly higher than in the OLS base regression in Table 2.

As a second robustness check, we re-estimate the selection model excluding those individuals that withdraw from the labour market one year

[11] We use the OLS base regression because we cannot integrate the firm- and occupation-change variables into the Heckman selection model due to asymmetry reasons (i.e., having variables in the outcome equation but not in the selection equation imposes further restrictions on the reduced form model).

after the training for various reasons, e.g. military service or further (academic) education (Table B2). Estimates without these 1,078 individuals show that our results reported in Table 3 are robust to the exclusion of this group of graduates. The negative influence of the 'male' variable in the original equation can be explained by the military service that several of the male apprentices have to serve after training.

Finally, we test for effect heterogeneity by splitting the sample into graduates that were trained in a low-investment firm and graduates that were trained in a high investment firm. As a threshold, we use the mean (of the log) of training investment. Running the previous regressions on both samples shows that increasing training investments of apprentices in low-investment firms yields a wage premium of close to 4% (Table XY, Column 1). For apprentices in high-investment firms, the respective estimate is less than 1% and statistically not significant. These results suggest that especially in the low-investment segment of training firms, apprentices would benefit from an increase in the resources dedicated to training. Conversely, apprentices that are already trained in high-investment firms cannot expect a significant wage increase following further investments in training.

Summing up the exercises above, we find support for the positive relation between the training investment and the later wages of apprentices. However, the relation is strongest in the segment of low-investment training firms.

6 Conclusions

Earlier estimates on the impact of education on wages often lack precise information about the specific conditions of the education process. Education is treated as a black box that is approximated by years of training or formal educational degrees. However, we argue in this paper that differences in

the training investment of the firm may be important for explaining labour market outcomes of apprenticeship graduates. Despite the large body of literature on the relation between youth training and post-training wages, no study so far has used detailed data on monetary measures for firms' training investment.

The paper therefore developed a firm training investment indicator (*FTI*) and merged the firm-level survey data with individual-level administrative records. The data merge made it possible to 'follow' apprentices from the surveyed training firms into their first labour market experiences as skilled workers. The paper then applied a regression analysis taking account of the selection into training and employment.

The results suggest that the variation of firm investments in training can explain at least part of the wage differences of apprenticeship graduates. Doubling firms' training investments leads to a wage mark-up of about 2.8%. Several alternative specifications do not fundamentally change this result. However, especially graduates trained in low-investment firms profit from additional training investments, as their wage increases over proportionally (i.e. by 4%). In turn, graduates in firms that already provide for high training investments show no significant wage gains due to additional investments.

Finally, this paper suggests some important further research topics. First, the significant influence of training investment on post-training wages shows that human capital formation within the dual system of training is not homogeneous but at least party depends on the resources that firms dedicate to their training. While this paper provides evidence for the importance of monetary resources dedicated to training, it does not discuss outcomes of different learning modes that are equally costly to the firm. Opening up the black box of training to understand, which training modes lead to superior labour market outcomes for graduates would be important for all actors in the field, i.e. for individuals, firms and policy makers. Second, this paper focuses

on the wage outcome measured one year after graduation. The question whether the positive relation between training investments and wages persists over time remains a task for future research.

References

Acemoglu, Daron and Joern-Steffen Pischke (1998). Why Do Firms Train? Theory and Evidence. *Quarterly Journal of Economics 113* (1), 79–119.

Acemoglu, Daron and Joern-Steffen Pischke (1999a). Beyond Becker: Training in Imperfect Labour Markets. *Economic Journal 109*, 112–142.

Acemoglu, Daron and Joern-Steffen Pischke (1999b). The Structure of Wages and Investment in General Training. *Journal of Political Economy 107* (3),539–72.

Bartel, Ann P. (1995). Training, Wage Growth and Job Performance: Evidence From a Company Database. *Journal of Labor Economics 13* (3),401–425.

Bassanini, Andrea, Alison L. Booth, Giorgio Brunello, Maria de Paola, and Edwin Leuven (2007). Workplace training in Europe. In: Brunello, Garibaldi and Wasmer (eds.), Education and Training in Europe, Ch. 8 – 13. Oxford University Press.

Boudon, Raymond (1974). Education, Opportunity, and Social Inequality: Changing Prospects in Western Society. New York: Wiley.

Breen, Richard and John H. Goldthorpe. (1997): Explaining educational differentials. Towards a formal rational action theory. In: Rationality and Society 9, 3, S. 275-305.

Büchel, Felix and Markus Pannenberg (2004). Berufliche Weiterbildung in West- und Ostdeutschland - Teilnehmer, Struktur und individueller Ertrag. Zeitschrift für Arbeitsmarkt Forschung 37 (2), 73–126.

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

Becker, Gary S. (1964). Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education. New York; London: Columbia University Press.

Beicht, Ursula, Günter Walden, and Hermann Herget (2004). Kosten und Nutzen der betrieblichen Berufsausbildung in Deutschland. Bundesinstitut für Berufsbildung. Berichte zur beruflichen Bildung, Heft 264. Bertelsmann: Bielefeld.

Blundell, Richard (1999). Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy. Fiscal Studies 20 (1), 1–23.

Card, David (1999). The Causal Effect of Education on Earnings. *Handbook of Labor economics. Volume 3A 3*, 1801– 1863.

Card, David (2001). Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica 69* (5), 1127–60.

Dietrich, Hans, Holger Alda, Harald Pfeifer, Felix Wenzelmann, Gudrun Schönfeld, Stefan Schiel, and Stefan Seth (2014). Linked Employer-Employee data on firms' training costs: Enriching register based LEE data with firm level data on apprenticeship training. *Schmollers Jahrbuch - Journal of Applied Social Science Studies 134 (2), 237-248.*

Heckman, James J (1974). Shadow Prices, Market Wages, and Labor Supply. *Econometrica 42* (4), 679–94.

Heckman, James J. (1979). Sample Selection Bias as a Specification Error. *Econometrica 47* (1), 153–61.

Mincer, Jacob A. (1974). Schooling, Experience, and Earnings. NBER Books National Bureau of Economic Research. Cambridge.

Mincer, Jakob (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy 66(4)*, 281–302.

Muehlemann, Samuel, Harald Pfeifer, Günter Walden, Felix Wenzelmann, and Stefan C. Wolter (2010). The Financing of Apprenticeship Training in the Light of Labor Market Regulations. *Labour Economics 17* (5), 799–809.

Riphahn, Regina T. and Michael Zibrowius (2015). Apprenticeship, Vocational Training and Early Labor Market Outcomes - Evidence from East and West Germany. *Education Economics* (forthcoming).

Schönfeld, Gudrun, Felix Wenzelmann, Regina Dionisius, Harald Pfeifer, and Günter Walden (2010). *Kosten und Nutzen der dualen Ausbildung aus Sicht der Betriebe*. Bielefeld (Bertelsmann).

Schultz, Theodore W. (1963). *The Economic Value of Education*. Columbia University Press.

Schultz, Theodore W. (1971). *Investment in Human Capital. The Role of Education and of Research*. New York (The Free Press).

Stevens, Margaret (1996). Transferable Training and Poaching Externalities. In Alison L.Booth and Dennis J.Snower (Eds.), *Acquiring skills: Market failures, their symptoms and policy responses*, Chapter 2, pp. 21–40. Cambridge University Press.

Stocké, Volker (2007). Explaining Educational Decision and Effects of Families´ Social Class Position: An Empirical Test of the Breen-Goldthorpe Model of Educational Attainment. *European Sociological Review* 23(4), 505-519.

vom Berge, Philipp; König, Marion; Seth, Stefan (2013): Sample of Integrated Labour Market Biographies (SIAB) 1975-2010. (FDZ-Datenreport, 01/2013 (en)), Nürnberg (IAB). (http://doku.iab.de/fdz/reporte/2013/DR_01-13_EN.pdf)

von Bardeleben, Richard, Ursula Beicht, and Kalman Fehér (1995). Betriebliche Kosten und Nutzen der Ausbildung: repräsentative Ergebnisse aus Industrie, Handel und Handwerk. Bielefeld (Bertelsmann).

von Wachter, Till and Stefan Bender (2006). In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers. American Economic Review 96 (5), 1679–1705.

Wagner, Robert and Thomas Zwick (2012). How Acid are Lemons? Adverse Selection and Signalling for Skilled Labour Market Entrants. ZEW Discussion Paper No. 12-014.

Wolter, Stefan C. and Paul Ryan (2011). Apprenticeship. Handbook of the Economics of Education 3, 521‑576.

Wooldridge, Jeffrey M. (2002). *Econometric analysis of cross section and panel data*. MIT Press.

A Figures

Figure A1: Daily gross wage in Euro

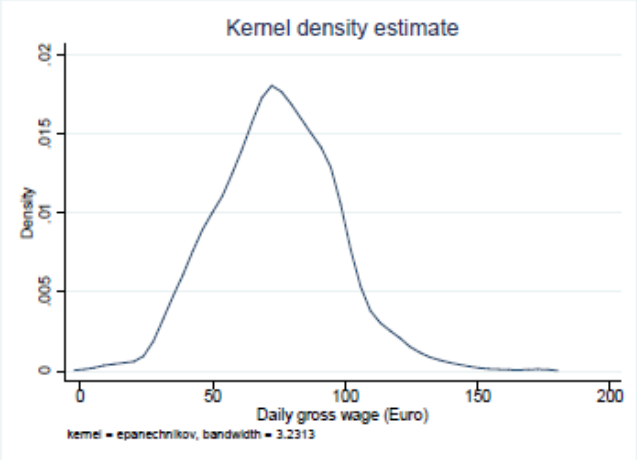
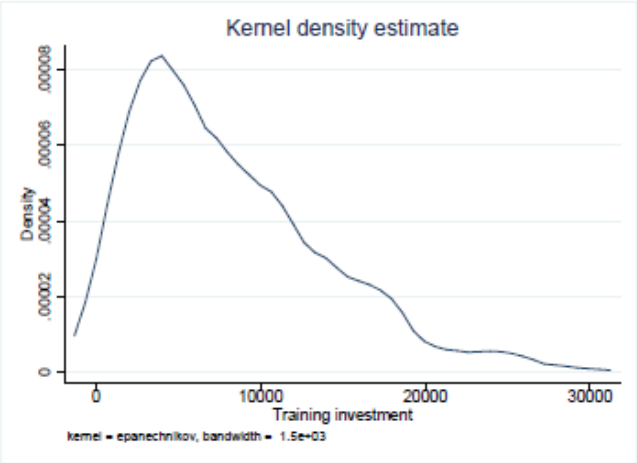


Figure A2: Training investment per year in Euro



B Tables

Table B1: OLS regression: Controlling for mobility after training

Dependent variable:	Daily gross wage (ln)	Daily gross wage (ln)
Firm training investment FTI (ln)	0.023***	0.023***
	(0.008)	(0.008)
Changed firm after apprenticeship	-0.113***	-0.111***
	(0.013)	(0.014)
Difference from average firm retention rate	-0.038***	-0.036***
	(0.012)	(0.013)
Changed firm*Difference retention		-0.153***
		(0.021)
Changed occupation after apprenticeship	-0.031***	-0.031***
	(0.012)	(0.012)
Age (years)	0.012***	0.012***
	(0.002)	(0.002)
Years of unemployment before apprenticeship	-0.067***	-0.067***
	(0.016)	(0.016)
Received scheme-based training support	-0.146***	-0.146***
	(0.028)	(0.028)
Duration of apprenticeship (in years)	0.064***	0.064***
	(0.009)	(0.009)
High school degree (*Abitur*)	0.115***	0.115***
	(0.013)	(0.013)
Male	0.068***	0.068***
	(0.009)	(0.009)
Non-German	-0.057***	-0.057***
	(0.015)	(0.015)
Region	Yes	Yes
Training vocation	Yes	Yes
Firm size	Yes	Yes
Economic sector	Yes	Yes
Year of graduation	Yes	Yes
Constant	3.416***	3.415***
	(0.083)	(0.084)
Observations	9117	9117

Firm-clustered standard errors in parentheses.
Significance levels: *<10%, **<5% and *** <1%.

Table B2: Heckman selection model: Excluding inactive graduates

Dependent variable:	Daily gross wage (ln)	Employed (yes/no)
Firm training investment FTI (ln)	0.027**	-0.057*
	(0.008)	(0.029)
Age (years)	0.010***	0.025**
	(0.002)	(0.008)
Years of unemployment before apprenticeship	-0.030*	-0.214***
	(0.015)	(0.049)
Received scheme-based training support	-0.103**	-0.315**
	(0.033)	(0.108)
Duration of apprenticeship (in years)	0.045***	0.203***
	(0.009)	(0.041)
High school degree (*Abitur*)	0.094***	0.250***
	(0.013)	(0.061)
Male	0.073***	-0.011
	(0.010)	(0.047)
Non-German	-0.039*	-0.107
	(0.017)	(0.069)
Received social benefits before apprenticeship		-0.524***
		(0.070)
Region	Yes	Yes
Training vocation	Yes	Yes
Firm size	Yes	Yes
Economic sector	Yes	Yes
Year of graduation	Yes	Yes
Constant	3.475***	0.656**
	(0.083)	(0.335)
ρ	-0.886	(0.022)
σ	0.315	(0.014)
Observations	9117	10245

1128 restricted cases. Firm-clustered standard errors in parentheses.
Significance levels: *<10%, **<5% and *** <1%.

Table B3: Heckman selection model: Split-sample regressions

	Graduates from firms *above* mean investment	Graduates from firms *below* mean investment
Dependent variable:	Daily gross wage (ln)	Daily gross wage (ln)
Firm training investment FTI (ln)	0.009	0.040***
	(0.021)	(0.013)
ρ	-0.926	-0.892
	(0.017)	(0.025)
σ	0.339	0.318
	(0.022)	(0.019)
Observations	6431	4892

The regressions include the same control variables as in main regression (Table 3).
Significance levels: *<10%, **<5% and *** <1%.