

Information Advantages of Training Employers Despite Credible Training Certificates*

Jens Mohrenweiser (Bournemouth University)

Gabriele Wydra-Somaggio (IAB)

Thomas Zwick (University of Würzburg and ZEW Mannheim)

ABSTRACT:

Employers sponsor general training if they can get an information advantage about the skills of trainees. This advantage might vanish if credible external institutions grade successful training completion. We argue that despite a certificate, training firms can retain an information advantage and can positively select training completers because the training certificate does not capture all relevant skill dimensions. Training firms can get an information advantage on the non-certified skills such as soft skills. We show that the information advantage about non-certified skills can explain why training firms pay higher wages for training completers than outside firms and why employers are able to keep a positive selection of training participants.

JEL Codes: J24, J31, J62, J63, M52, M53.

Key words: training, employer change, adverse selection, asymmetric information.

*Corresponding author: Thomas Zwick, University of Würzburg, thomas.zwick@uni-wuerzburg.de. We would like to thank Decio Coviello, Benoît Dostie, Fabian Lange, Daniel Parent, Christian Pfeifer, Paul Ryan, Uta Schönberg, and Michael Waldman for helpful comments.

1 Introduction

Firms can invest in transferable skills of their employees because training generates an information advantage about training content (Katz and Ziderman, 1990) or employee ability (Chang and Wang, 1995; Acemoglu and Pischke, 1998; Autor, 2001). The information advantage drives a wedge between productivity and 'marketable' skills or outside earnings. The information advantage allows training firms pay trained employees below productivity but above market earnings because trained employees cannot credibly signal the higher skills to the outside market. Hence, training firms can keep trained employees and get a return on investment. The wedge increases in skills (Acemoglu and Pischke, 1999; Booth and Zoega, 2004; Leuven, 2005) leading to a positive selection of employees who stay in the training firm (Schönberg, 2007).¹

The information advantage of training firms might vanish, however, if credible external institutions reveal the individual performance and the training content with graded training certificates (Arcidiacono et al., 2010). Receiving a credible certificate boosts individual employability and outside earnings options. Certificates and final exams after training are widespread. Examples include (upgrading) training programmes accredited by chartered institutes², employer-sponsored Bachelor, Master and MBA courses (Cappelli, 2004; Benson et al., 2004; Pattie et al., 2006; Arcidiacono et al., 2010; Manchester, 2010, 2012; Benson, 2013), recognised language certificates and IT courses (de Grip and Sauermann, 2012; Sauermann, 2015), mandatory health and safety training, and apprenticeship training (Mohrenweiser and Zwick, 2009; Schönfeld et al., 2010; Ryan et al., 2013; Mühlemann et al., 2013; Kriechel et al., 2014). In recent years, also IT providers such as Microsoft, SAP or Oracle increasingly certify proficiency in their software. All of these certificates overcome information asymmetries about training content and should reduce firms' incentives to invest in training (Katz and Ziderman, 1990). We therefore ask how training firms can gain an information advantage if training courses are credibly certified by external institutions.

¹ Company-sponsored training has a larger effect on productivity than on wages (Hinz and Mohrenweiser, 2017; Konings and Vanormelingen, 2015; Jones et al., 2012; Almeida and Carneiro, 2009; Dearden et al., 2006; Conti, 2005) and a stronger wage-compression on the firm-level leads to more company-sponsored training (Pfeifer, 2016)

² Chartered institutes such as the CMI (Chartered Management Institute) or the CIPD (Chartered Institute of Professional Development) among others accredit a variety of training courses from level 2 apprenticeship courses to level 7 master courses in the UK.

We show that even accredited courses with graded certificates do not reveal all relevant skill dimensions. Training certificates typically grade only quantifiable skills for which the level of proficiency can be measured in standard tests (Nickson et al., 2012)³. Such skills are usually referred to as hard skills, i.e. technical skills related to specific work tasks. Certificates and exams allow training and outside firms to assess relative performance of trainees in hard skills in comparison to other training participants. Certificates turn hard skills into public signals. Nonetheless, hard skills do not cover all relevant skill dimensions that are valuable on the labour market. Most notably, soft skills such as communication, courtesy, integrity, responsibility, teamwork, a positive and flexible attitude, and work ethics (Laker and Powell, 2011; Robles, 2012) are generally considered to be more relevant than hard skills for employability, performance, and learning – particularly for entry-level jobs (Andrews and Higson, 2008; Nickson et al., 2012; Heckman and Kautz, 2012)⁴. Nevertheless, because soft skills are typically hard to quantify and measure, they are typically not examined in traditional training certificates (Heckman and Kautz, 2012). Hence, soft skills may remain private information to the training firm even if training is certified and graded by external institutions. In other words, information asymmetry on skills that are not certified might generate the information advantage that creates the earnings difference between training and outside firms.

To analyse information advantages of training firms after certified general training, we first use the model developed by Schönberg (2007) and slightly adapt it to the German apprenticeship system. The model nests public and private information on skills. In the public information case, training and outside firms can equally assess all skill dimensions. In the private information case, the training firm knows more about at least one skill dimension of their employees than the outside firm. The model allows us to derive hypotheses that identify information asymmetries between training and outside firms based on earnings offers and job mobility after training.

³ The term “skills“ does not exclude abilities in the sense that some productive assets of the training participant may have been present already before the training started or improved the skill increase during training.

⁴ We use the hard/ soft skill dichotomy instead of the cognitive/ non-cognitive skill dichotomy that is frequently used in the literature (Heckman et al., 2006). Hard skills better describe skills apprentices (our sample population) learn compared to cognitive skills that might be better suited for college/ high-school comparisons.

Second, we empirically test these hypotheses using particularities of the German apprenticeship training system. Apprenticeship training is a prime example of employer-sponsored training in skills that can be publicly signalled. Skills provided during apprenticeship are transferable, visible and certified with marked final exams by independent, external institutions. We use a unique database including marks of the final apprenticeship certificate among other important information from administrative sources for the entire population of several cohorts of apprenticeship completers in one German federal state. The marks serve as a public signal on skills obtained during apprenticeship training. To measure skill dimensions that might not be revealed in the certificate but nevertheless relevant for the training firm, we follow the approach proposed by Kahn (2013) and Kahn and Lange (2014) and assume that relative productivity differences between individuals are reflected in earnings, and that differences in earnings between observationally equal employees reveal private information on relevant skills. Hence, we use small earnings differences at the end of an apprenticeship as a proxy for private information of a training firm about relevant skills. More specifically, we take earnings differences between homogeneous peers in the same training firm, occupation, and cohort at the end of apprenticeship training as an indicator of individual skills not covered in training certificates. As we also condition on marks besides confounding factors in the entry job earnings estimations after apprenticeship, earnings differences at the end of apprenticeship identify performance dimensions that are not captured in apprenticeship certificates.

Our empirical analysis shows that training firms keep a positively selected group of apprentices after training and maintain an information advantage even if the programme is certified by credible external institutions. Additional earnings at the end of training influence entry earnings of stayers but not the entry earnings of movers. On the contrary, final marks in apprenticeship certificates influence entry earnings of movers significantly stronger than the entry earnings of stayers. This paper contributes to the literature on training motivations of employer-sponsored training by demonstrating that training firms can keep an information advantage on hard-to-observe skills such as soft skills even if the training course is certified. We also show that the information advantage enables training firms to keep a positively selected group of apprentices after training.

This paper is structured as follows. We derive hypotheses based on a theoretical framework in section two and describe the institutional background and the data set in section three. The fourth describes the variables and the fifth presents and discusses the findings. The final section concludes.

2 Theoretical framework

This section outlines a theoretical framework to identify whether outside employers have the same information about all relevant skill dimensions of training completers as training firms (the public signal case) or whether training employers know more about trainees' skills than outside firms (the asymmetric information case).

We apply the model proposed by Schönberg (2007) to the context that employers invest in transferable skills training but only one of two skill dimensions are assessed in the final certificate.⁵ The model consists of two periods. In the first period, apprenticeship training takes place and ends with an accredited certificate.⁶ The training employer can screen training participants during the first period and decide whom it would like to offer an employment contract in the second period (first job of apprenticeship completers). In the second period, an outside employer can make an offer and the training employer can counter the offer. The trainee stays if the offer is higher than the outside offer plus individual disutility from continued employment learned during the training period.⁷ Disutility is not observable by the training employer but it has expectations about the distribution of disutility. Hence, the model generates some voluntary turnover for exogenous reasons (similar to the models in Acemoglu and Pischke, 1998 or Autor, 2001).

The model distinguishes between two types of workers, high and low productivity workers, and between two skill dimensions. The training employer learns about the individual level of both skill dimensions during the first period. One skill dimension is easy to observe for all market participants on basis of the credible certificate. Another skill dimension is potentially

⁵ The electronic Appendix B provides a formal derivation of the model.

⁶ Differences in the quality of training firms that for example might have consequences for the selection of highly skilled apprenticeship candidates into training firms are not taken into account in the theoretical model. We capture these differences empirically, by introducing firm fixed effects.

⁷ Examples for disutility are problems with supervisors or co-workers, perceived underpayment or changes in regional preferences of the apprentice.

private information. A high level of the publicly signalled skill dimension adds the same productivity to employees with low and with high private information skills. The probability that the employee has high private information skills is higher when the employee has high public information skills. The a-priori probabilities of high and low skills in both dimensions and their positive correlation are known to all employers.

The key question in the model is whether an outside employer can infer the private information. Outside employers need a reliable signal such as a job interview or a reference letter to infer private information on skills (Lange, 2007). If this additional signal perfectly transmits the private information, outside and training firms have the same (correct) information and private information on skills becomes public. If the additional signal is only white noise, the outside employer cannot infer anything about the private information in addition to the a-priori probability distribution of skills. In this case, one skill dimension remains private information.

If outside employers can accurately assess both skill dimensions, training and outside employers offer the same wage for high skills. This means that only the disutility of training participants allocates movers and stayers. As the disutility distribution is equal for employees with low and high skill levels, the training employer cannot retain a positive selection of employees.

If the training employer has superior information on one skill dimension, the outside employer is not able to pay fully for the additional productivity associated with a high level in this skill dimension. In the asymmetric information equilibrium, the training firm can offer slightly higher earnings for training participants with high private information skills and can keep them. The training firm can also select those completers who have to leave after training – it either does not give them a higher counter-offer or it does not give them an employment offer. Hence, outside employers anticipate that training employers use their information advantage and they consider employer movers as having lower private information skills (Greenwald, 1986; Gibbons and Katz, 1991).

Summing up, we nest public and asymmetric information and derive the following hypotheses: If both skill dimensions are public information, they have a similar impact on entry earnings for those who stay after training and those who leave. In particular, movers

are not negatively selected in comparison to stayers. If training employers have an information advantage, public signal skills have a stronger impact on the entry earnings of movers than of stayers because the outside employers can only rely on the public signal. Private information skills have a stronger impact on earnings of stayers than on earnings of movers. Finally, movers are negatively selected in comparison to stayers.

3 Institutional Background and Data

Institutional Background

We use the German apprenticeship training system as an example for training in general human capital that ends with a generally recognised certificate. Apprenticeship training in Germany traditionally provides training for about two thirds of the German workforce and is the backbone for medium skilled vocational training. Apprenticeship training follows a dual track with training in firms and vocational schools. Firms decide whether they want to train apprentices and whom they train. School-leavers apply for apprenticeships in firms. Apprentices usually start the apprenticeship after school. The apprenticeship lasts three or three and a half years, depending on the occupation. This means that the first period in our model takes several years and offers enough time for the training employer to learn about all skill dimensions. Finally, apprenticeships are generally considered as training investments for firms (Mohrenweiser and Zwick, 2009; Schönfeld et al., 2010) because firms incur net costs until the end of the apprenticeship in most occupations.

Apprenticeships are strongly regulated by the Vocational Training Act and occupational specific training curricula. The Vocational Training Act describes the length of training, necessary equipment and requirements for training firms. Training firms have to fulfil these requirements to get permission for apprenticeship training granted by the chambers of industry and commerce or the chambers of craft. Training curricula are published, bespoke for each occupation and describe minimum skills, which have to be acquired in each training occupation for a successful graduation. The basic training contents are therefore identical for all apprentices in an occupation, irrespective of the training firm and region. The chambers observe the quality of apprenticeships in each enterprise and administer the final exam on the practical part of the skill examination. The chambers set the final exam day that typically takes place on the same day for all apprentices within one occupation in a region.

Each apprenticeship contract legally terminates at the day after the final exam and therefore all apprenticeship completers in one occupation enter the labour market at the same day. The theoretical part of the skill examination is administered and marked by public vocational schools (see Soskice, 1994 or Mohrenweiser et al., 2018). Apprentices receive marked certificates at the end of the training period. Marks allow all employers to assess hard skills of otherwise seemingly equal apprenticeship completers. Apprentices are free to choose an employer after the exam. Training costs reimbursement contracts for apprentices are forbidden by law. Outside firms cannot easily verify whether a job applicant had an employment offer from the training firm or not.

Data

We use the *Ausbildungspanel Saarland*, a data set linking Social Security records with administrative exam records from the chambers of industry and commerce and chambers of craft in the German federal state Saarland. The data comprise the entire apprenticeship graduation cohorts in this federal state between 1998 and 2005. We merge both data sets on basis of the initial of name and surname, birthday, gender and the start of apprenticeship.

From exam records, we use the final exam marks, training occupation, and school-leaving qualification, all of which are more precisely reported than in the Social Security records. The occupation code exactly corresponds to the training occupation named in the apprenticeship certificate and clearly distinguishes three and three and a half year apprenticeships. From the Social Security records, we use earnings, nationality, gender, training firm size, and the establishment identifier. In the Social Security records, the earnings information is very reliable because it is used to calculate unemployment benefits and old age pension claims.

We detect apprenticeship completers with a graduation identifier in the exam records. We restrict the sample to those apprentices whose apprenticeship spell in the Social Security record ends close to the exam record day (compare Table 1 for a detailed data selection procedure). We only consider completers starting full-time employment within 30 days after

apprenticeship⁸, with earnings above the marginal income threshold and with earnings below the upper Social Security earnings threshold. We further remove apprentices who are older than 30 years and apprentices with earnings below the lower Social Security threshold, because observations are very likely to be incorrect data entries.

The sample selection procedure generates an individual-level pooled cross-section dataset. Each apprenticeship completer enters the sample in the year of the first full-time employment after apprenticeship completion.

4 Empirical Strategy

Dependent variables: probability to move and entry earnings

Our first dependent variable is a binary variable of the status as staying or leaving apprenticeship completer. The second is entry earnings measured in the first employment spell after graduation. This spell takes clearly less than one year.⁹ This means that a quick revelation of skills by a new employer cannot influence measured entry earnings.¹⁰ The measure of entry earnings for leavers and stayers is also not biased by the business cycle because most apprenticeship completers started and finished their apprenticeship at the same point in time and they do not have prior experience on the labour market. As all apprenticeship contracts end at the day after the final exam, training employers also have to offer a new employment contract with new and clearly higher earnings for staying apprenticeship completers. Therefore, earnings of leavers and stayers are determined at exactly the same point in time.

Measurement of public signal

We measure the publicly signalled skill dimension with the marks of apprenticeship completers in their final exams at the end of the training period. The certificate provides

⁸ This restriction increases the comparability between movers and stayers. Without the restriction, the effects become larger.

⁹ Employers have to report earnings once a year but status changes from apprenticeship to regular employment require an additional report during a calendar year. The average duration of the entry earnings spell is 178 days.

¹⁰ For longer entry spells, it may have been possible that earnings have been increased after a while for movers who are more productive than a-priori expected. This could confound the interpretation of entry earnings as expected productivity before the employer knows the employee (Lange, 2007).

easily accessible and unbiased publicly available information on mainly hard skills.¹¹ Outside employers can assess the skills of each apprentice on basis of the marks in several practical and theoretical subjects.¹² Since most employers who hire apprenticeship completers train themselves (Mohrenweiser, 2016), they have experience in assessing marks reported in apprenticeship certificates and supplementary information such as school background, occupation, and selectivity of the training employer.

Measurement of private information

We follow Kahn (2013) and Kahn and Lange (2014) and measure the skill dimension that is potentially private information for the training firm by an indirect assessment. Assuming that wages reflect productivity, wage variations serve as a proxy for private information if we keep possible confounding factors constant. For identifying the private information part on productivity, it is crucial to find homogeneous peers. We therefore define cells of one training employer, training occupation, and apprenticeship completion cohort. In each cell, apprentices are subject to the same training instructors, training content, selectivity into training, and collective bargaining rules. As apprentices typically come directly from school and have no prior labour market experience, *earnings differences at the end of apprenticeship* between apprentices in these cells reflect the opportunity of a training employer to differentiate earnings before training ends. As we condition earnings differences on employer, occupation and cohort cells, we can interpret them as differences within a homogenous peer group. As we also condition on marks in our earnings estimations, the earnings differences identify performance dimensions that are not captured in apprenticeship certificates.

Within the peer group, a higher earnings position can be interpreted as additional payments. Paying more than the base salary for apprentices at the end of apprenticeship training is at first sight surprising because almost all apprentices in our sample are likely to work for employers with collective bargaining contracts. Employers covered by a collective agreement could just pay the collectively bargained wage for their apprentices. Nevertheless, additional

¹¹ We include schooling level as alternative public signal to control for endogeneity induced by correlations between schooling level and grades in the apprenticeship examinations (Schönberg, 2007).

¹² We include the average grade of the theoretical and practical final apprenticeship exams.

voluntary payments for selected apprentices on top of the bargained earnings is a widespread policy of training firms – almost all training firms differentiate their earnings for observationally equal apprentices when the end of apprenticeship approaches but not at the beginning (Mohrenweiser et al., 2018).¹³

The relative earnings position of the apprentice within a cell satisfies the hard-to-observe nature of the private information indicator. Although outside employers may observe the absolute individual earnings of an apprentice at the end of apprenticeship, apprentices cannot credibly signal their relative position in the earnings hierarchy of the training firm to improve their earnings opportunities. Employers never reveal their complete earnings structure (Waldman, 1990). Outside employers can interpret high earnings at the end of the training period as a consequence of high additional payments or of a relatively high earnings level for all apprentices in the training firm.¹⁴ As a consequence, training employers do not have an incentive to use the wage differentiation to disguise high performing apprentices. The employer-generated private information is not biased by strategic considerations in contrast to publicly observable productivity indicators such as promotions (DeVaro and Waldman, 2012; Waldman, 2016; Waldman and Zax, 2016; Bognanno and Melero, 2016).

Both skill measures (earnings differences during apprenticeship training and marks in final exams) have the advantage that they have been created closely before the actual decision about entry earnings after graduation (and not for example already before labour market entry). Therefore, unobservable heterogeneity such as additional training for higher skilled apprentices between information creation and the earnings measure is not an issue. The training employer determines the additional payment before the final exam is taken. Moreover, the marks are issued by external institutions. This means that both measures are determined independently and that (as we will show below) they measure different skills dimensions.

¹³ Appendix A provides empirical evidence and arguments that earnings differences at the end of the apprenticeship are indicators for apprentice productivity: They predict the entry earnings structure, they appear only in the final months of the apprenticeship and HR professionals report bonus payments for good marks or extraordinary citizenship.

¹⁴ In order not to confound relative and absolute earnings level, we also control for the average earnings level a training employer pays at the end of the apprenticeship period.

To compare the impact of earnings differences during the last spell of the training period and marks in final exams on entry earnings for movers and stayers, we follow Farber and Gibbons (1996) and Schönberg (2007) and assume that the productivity assessment of the training and outside firm is the same. In other words, all skill dimensions have the same value in all entry jobs for apprenticeship completers. This assumption is less severe in our setting because almost all employers that hire apprenticeship completers also train apprentices themselves (Mohrenweiser, 2016) and we only compare apprenticeship completers within the same occupation. Training and outside employers therefore should not be structurally different with respect to their skills assessments and skills demand.

Control variables

We include individual and establishment level control variables. On the individual-level, we control for gender, age, nationality, and the schooling-level – common determinants in earnings and mobility analyses (von Wachter and Bender, 2006; Göggel and Zwick, 2012; Mueller and Schweri, 2015; Dietrich et al., 2016). On the training-firm level, we pursue two strategies: either we control for the most important employer characteristics, firm size and median earnings, or we include a cell fixed effect. The cell fixed effect allows us to compare homogeneous peer groups (same training employer, occupation, and completion cohort) and additionally controls for training-firm unobservable variables such as training instructors, initial selection criteria into apprenticeship firm, occupation and collective bargaining rules. Training firm characteristics may inform the labour market about initial selection of high ability school-leavers into high-reputation training firms because apprenticeships at larger and well-paying enterprises are more attractive for apprenticeship candidates (Soskice, 1994; Acemoglu and Pischke, 1998; Mohrenweiser and Zwick, 2015). Completing the apprenticeship at well-paying and large training employers may indeed be a positive public signal.

5 Findings

Descriptive statistics

Table 2 summarises variable definitions, means and standard deviations. The share of employer movers immediately after apprenticeship training is around 40%. For employer

movers and stayers, Table 3 separately displays the descriptive statistics. Movers have on average lower marks in their final apprenticeship exams than stayers.¹⁵ They also have a significantly lower schooling background: they are less likely to graduate from a medium and upper secondary school track (*Realschule* and *Abitur*) and they are more likely to graduate from a lower secondary school track (*Hauptschule*), or without a school certificate.

Similarly, movers have significantly lower earnings at the end of apprenticeship training. They have an unconditional earning disadvantage of 20% (19.05€ daily earnings versus 23.78€). Finally, movers receive significantly lower earnings in their first job after apprenticeship than stayers. The raw difference is again substantial with about 15% (51.47€ daily earnings versus 60.76€).

The public and private information measures are significantly correlated with a Pearson's correlation coefficient of 0.25. The correlation between both indicators is somewhat stronger for movers than for stayers. The relatively low correlation demonstrates that both indicators measure different skill dimensions. Finally, movers are significantly older and come from smaller training firms that pay lower earnings. We find no differences with respect to nationality and gender.

Adverse selection

We first test whether training firms are able to keep a positive selection of their apprenticeship completers. We use a Linear Probability Model, which has the advantage that we can include cell fixed effects and interpret the coefficients of the interaction terms as marginal effects. In addition, this regression usually produces marginal effects similar to a Probit or Logit regression if the dependent variable has an unconditioned probability around 50% as in our case.

We find that the marks and additional payments are both negatively associated with the probability to leave the training firm (Table 4). The first three models in Table 4 restrict the sample to cells with at least one leaving and one staying apprenticeship completer in order to analyse firms that select (or are able to attract) apprentices to stay. Models 1 and 2

¹⁵ In Germany, 1 is the best mark (with distinction) and 4 the lowest (pass). To improve interpretation, we multiply the marks by minus one. Hence, a higher number represents higher skills.

account for several individual and training establishment characteristics. An improvement in the marks by one at the final apprenticeship exam reduces the probability to leave the training employer by 5.2 percentage points. A one Euro higher payment a day reduces the probability to leave by 1.9 percentage points (level-log model: $-0.419/21.75$) in Model 2.

Model 3 includes cell fixed effects instead of establishment-level variables. This is our preferred specification because it conditions on comparable peers learning under the same conditions. This specification identifies whether training firms are able to retain apprentices with the relatively higher skills in comparison to their peers. The point estimates of additional payment and marks strongly increase in comparison to models 1 and 2. A one Euro higher payment a day compared to the peers in the cell reduces the probability to leave by 4 percentage points and a better mark by one decrease the probability to leave by 10.7 percentage points. The reduction in the probability to leave induced by marks and additional payment is substantial given the unconditioned probability to leave of 40%.¹⁶ Finally, models 4 and 5 include individuals in all cells even if all completers leave or stay in the cell. Model 5 controls for cell fixed effects. The results remain qualitatively the same but the point estimates are smaller in model 5 compared to model 3. This finding supports the argument that variations in marks and additional payment are more likely to be used in firms that keep some but not all of their apprenticeship completers.

The control variables have the expected signs. We find a positive selection of stayers with respect to schooling background. Completers with a medium-track secondary school certificate have a lower probability to leave their training employer than completers holding a lower-track secondary school certificate. German citizens have a higher probability to leave than foreigners and older completers have a higher probability to leave than younger completers¹⁷. This finding could be the consequence of discrimination of training firms with respect to these minorities (Büsch et al., 2009; Kaas and Manger, 2012). Finally,

¹⁶ The negative effect of the private signal despite conditioning on the public signal might be surprising. The negative correlation between mobility and private information for asymmetric information however translates into a negative correlation between mobility and public signal given the positive correlation between public signals and private information.

¹⁷ The quadratic age term indicates that the maximum of the age effect is between 23 years (column 2, Table 4), and 40 years (column 5), well beyond the average age of completers (21.4 years).

apprenticeship completers in larger training firms have a lower probability to leave than those from smaller firms.

Hence, we can characterise employer movers after apprenticeship training as an adversely selected group with respect to both skill indicators. Our model interprets the positive selection of stayers as a consequence of information advantages of training firms.

Entry earnings

We test whether training firms have an information advantage represented in the additional payment during apprenticeship in addition to publicly visible marks. Table 5 displays augmented Mincer entry earnings regressions for apprenticeship completers. The first three models include additional payments during apprenticeship, marks and a mover dummy. The latter three models include an interaction effect between the mover dummy and the public/private information indicators. Models (1) and (2) as well as (4) and (5) control for individual and establishment characteristics and the models (3) and (6) use cell fixed effects that improve the comparability between peers and are our preferred specification.

The Models (1)-(3) show that apprenticeship completers with higher additional payments during apprenticeship and better marks get higher entry earnings in the first full-time skilled job. Moreover, movers get a wage penalty of about seven percent. Females and older completers receive a lower entry wage and a higher school degree additionally increases the entry earnings. In line with earlier results (Göggel and Zwick, 2012; Fitzenberger et al., 2015; Müller and Schweri, 2015; Dietrich et al., 2016), apprentices leaving a high-paying training firm receive higher earnings.

Models (4)-(6) include the interaction effects between the mover dummy and the public/private information indicators to test our theoretical predictions. The interaction effects confirm our hypothesis that training firms have private information on skills. Higher additional payments during apprenticeship have a significantly lower influence on entry wages for movers than for stayers. The point estimate for movers is insignificant (0.092 – 0.097). Movers cannot cash in their higher skills remunerated by the additional payment at the end of apprenticeship compared to the staying peers in the same firm, occupation, and cohort. In contrast, better marks have a significantly stronger influence on entry wages for

movers than for stayers¹⁸. Marks therefore have in the words of DeVaro and Waldman (2012) the characteristics of a “publicly observable ‘announcement’ of the worker’s ability”. Finally, apprentices can compensate for low private information with a high public signal. Nevertheless, given the public signal, they could get a higher wage when staying in the training firm (adverse selection effect).

Robustness checks

To identify the effect of public and private skill information on earnings, we include both skill indicators into the earnings equations. This approach rests on the assumption that both indicators are not strongly correlated with each other and therefore measure different skill dimensions. If both indicators are highly correlated, the inclusion of the second dimension would reduce the impact of the other in the earnings equation (Farber and Gibbons, 1996; Altonji and Pierret, 2001; DeVaro and Waldman, 2012). The reduction in the absolute size of the coefficients should be especially strong for stayers because the training employer can assess all relevant skill dimensions. We check the change in the magnitude of the coefficients of public and private information when excluding the other indicator (Table 6). The regressions control for the same variables as models (2), (3) and (5), (6) in Table 5. The magnitude of the coefficients of interest remains remarkably stable when we eliminate the other skill dimension (compare coefficients in Table 6 to those in Table 5). The stability of the skills coefficients is an additional indicator that both variables measure different dimensions of productivity (Heckman et al., 2006).

We also check whether our results are robust when we reduce the sample of employer movers to those apprenticeship completers who are very likely to leave on exogenous reasons (Gibbons and Katz, 1991). We follow the strategy proposed by von Wachter and Bender (2006) who argue that if training employers experience a severe short-term demand shock, they reduce the retention rate of apprenticeship completers dramatically as a cost efficient way to reduce total employment. The retention rates usually bounce back in the following year. A strongly reduced retention rate can serve as a valid test to assess potential

¹⁸ Movers have in model (6) on average 5.7 percent lower entry earnings than stayers $0.343 + (-0.097*2.947) + (0.038*-3.02)$. The marks are multiplied with -1 to facilitate the interpretation that higher numbers represent higher skills. The earnings deduction for movers in models (4) and (5) is comparable.

confounding factors such as individual career aspiration on mobility and entry wages of apprenticeship completers. We identify employer movers affected by an exogenous labour demand shock using two retention rate reduction thresholds: first, a reduction by more than 50% in one single year (model 1 in Table 7) and second, a 20% reduction (model 2 in Table 7). Table 7 reports the key variables in the entry earnings regression for employer movers and controls for the same variables as model 5 in Table 4. We find similar point estimates for the private information and the public signal on entry earnings. Private skills information does not have a significant impact on entry earnings in both sub-samples. Public skill signals are only significant in the larger sample (model 2). The coefficients in the smaller sample have low significance probably because the sample size is too small. We therefore conclude that we get similar results when we reduce our sample to mainly involuntary employer movers.

Moreover, we test if earnings in the first spell after apprenticeship completion are sensitive to reporting issues. We run the analysis using status and earnings in the first year after apprenticeship completion and obtain robust results. Finally, outside firms might infer hard-to-observe skill dimensions also from the fact that the apprenticeship completers got an employment offer (positive signal) by their training firm or by an employer move during apprenticeship training (negative signal). The binary retention and mobility during apprenticeship signals however cannot completely solve the information asymmetry on the continuous productivity rank of apprenticeship completers in a cell and the training firm retains a (reduced) information advantage.

Discussion

The paper argues that apprenticeship training firms can positively select apprenticeship completers because the credible training certificate does not capture all relevant skill dimensions. In this part, we discuss whether an information advantage is the only explanation for a positive selection of stayers or whether alternative arguments could explain our empirical findings. Besides information asymmetries on skills, the literature of employer-sponsored training under imperfect labour market conditions discusses several reasons why training firms can keep their best employees: information asymmetry on training contents, specificity of training contents, and regional monopsony power (Acemoglu

and Pischke, 1999; Booth and Zoega, 2004; Leuven, 2005). Even if these factors could explain our results, we argue that they are less plausible.

Information asymmetry on mandatory training contents is not plausible because the occupation-specific training curricula are standardised for all German regions and publicly accessible. In addition, most outside firms that hire apprenticeship completers immediately after completion also train apprentices themselves and know the training contents (Mohrenweiser, 2016).

Specificity of training contents can be created by learning-by-doing or if training firms voluntarily offer training beyond contents required by the training curriculum. Specificity of additional training could, in principle, be consistent with the pattern of mobility and entry wages. In this case, the skill dimension that is not measured in the training certificate measures firm-specific skills and the training certificate general skills. Here the firm-specific skill component needs to generate a rank order in skill proficiency within a cell that is weakly correlated with the (general) skill component. We can only speculate whether firm-specific learning-by-doing can generate the empirical pattern, but we can assess the relevance of firm-specific additional training using an external dataset of manufacturing firms. Appendix Table C1 shows that 53% of firms indeed train additional skills during apprenticeship. Manufacturing firms are usually considered to be more likely to use specific skills than service sector firms (Soskice, 1994). However, most firms state that the additional training covers contents such as language skills, IT skills, soft (or social) skills and skills usually learned in other metal-working training occupations. All of these categories comprise general skills. Only 13% of training firms state that they train additional skills not mentioned in other categories. This skills category could include firm-specific skills but also general skills not mentioned in the previous categories. The category is also much smaller than the category “additional soft skills” (36%). Hence, only a very small fraction of firms might train additional firm-specific skills. These additional skills might confound our findings but we think that their empirical relevance does not dominate our information asymmetry interpretation. Similar to this evidence, Göggel and Zwick (2012), Fitzenberger et al. (2015), and Müller and Schweri (2015) analysed entry wages after apprenticeship training and concluded that general and occupation-specific human capital are the dominant component of apprenticeship training.

Regional monopsonies are also an explanation why training firms can retain their best apprenticeship completers (Harhoff and Kane, 1997; Mühlemann and Wolter, 2011; Mühlemann et al., 2013). Most apprenticeship completers in Germany stay in the same labour market region when they move to another employer (Winkelmann, 1996). Our data come from the Saarland which is typically classified as one (of 114) labour market region (Kosfeld and Werner, 2012). We therefore cannot compare the relative importance of public and private information regarding the selectivity of movers across regions with different labour market competition. However, our results suggest a monopsony power mechanism within one labour market region in addition to possible regional monopsony effects.

Hence, we cannot be sure that the positive selection of stayers in apprenticeship training firms is solely a consequence of information advantages on skills not quantified in final apprenticeship examinations but additional evidence shows that the information advantage is the most plausible explanation.

6 Conclusions

This paper provides a novel answer to the old question why employers are willing to invest in externally certified training contents that are transferable and visible to all employers. Using the example of apprenticeship training in Germany, we show that apprenticeship completers who change their employer after training obtain higher returns for better marks in the final exams than apprenticeship completers who stay in the training firm. Marks can easily be signalled to outside employers on basis of marked certificates. We argue that the training certificate, however, does not signal all relevant skills of apprentices. Private information of training employers measured by additional payments during apprenticeship training does not yield higher entry earnings for apprenticeship completers who move to another firm after training. Higher payments however increase entry earnings for stayers. The information advantage about relevant skills not covered in the training certificate such as soft skills allows training firms to earn a return on their investment in transferable skills. They can pay highly productive apprenticeship completers earnings above the market but below their productivity. Hence, the paper clarifies previous discussions about the effect of certificates on firms' and individuals' incentives to invest in training. As long as the certificate covers not all relevant skill dimensions, the training firm can keep an information advantage.

Apprenticeships represent a specific form of employer-sponsored training in general skills that ends with a credible certificate issued by external and independent institutions. Employer-sponsored training in accredited training programmes is widespread, however. Such certificates usually do not reveal soft skills. Hence, training programmes with external certificates entail the possibility that employers can sustain an information advantage about relevant skill dimensions.

We demonstrate that training firms can positively select employer movers after apprenticeship training. Our paper presents a direct empirical identification of adverse selection in the labour market for apprenticeship completers in Germany. Staying apprenticeship completers are positively selected on basis of skills not covered in the training certificate. Hence, our paper complements studies that indirectly identify adverse selection of employer movers (Greenwald and Glasspiegel, 1983; Foster and Rosenzweig, 1993), studies using displacement losses as indicators of adverse selection (von Wachter and Bender, 2006; Göggel and Zwick, 2012; Fitzenberger et al., 2015), and studies using ability indicators as determinants of moving to another employer (Schönberg, 2007).

Our indicator for private information is based on the earnings structure within a homogeneous peer group of apprentices in the same training employer, occupation, and cohort. The relative earnings position of an apprenticeship completer in relation to peers is hard to observe for outside employers in comparison for example to promotions that are relatively easy to observe and frequently used in the literature (DeVaro and Waldman, 2012; Waldman and Zax, 2016; Bognanno and Melero, 2016). As a consequence, our private information indicator may not be strategically distorted by training employers in order to hide high-productivity employees. In this sense, our indicator complements recent analyses by Kahn (2013), and Kahn and Lange (2014) that also generate indicators for information asymmetries on skills exploiting earnings differences.

Finally, our paper adds to the literature on information asymmetries on the labour market analysing the consequences of symmetric and asymmetric information on skills and abilities (Schönberg, 2007; Pinkston, 2009; Kahn, 2013). Many empirical papers on information asymmetries on the labour market discuss the incentive effects (and their social welfare consequences) to obtain signals on skills irrespective of their productivity enhancing effect

(Lange, 2007; Waldman, 2016). It is beyond the scope of this paper but seems very interesting to assess whether the signalling value of better marks in apprenticeship exams induces apprentices to work harder for their final exam success. The model could also be extended to analyse if apprentices can overcome an initial stigma of a low signal in the long-run, particularly if they reduced their learning efforts because they have not been happy with their initial training firm or occupation choice. We also have to leave for future research to assess the relative importance of regional mobility costs and compressed wages for apprenticeship completers in unionised firms as additional reasons for the positive selection of apprenticeship completers who stay with their training firms and for training incentives. Finally, measuring potentially private information skill dimensions more accurately and combining those measures with final exam results to assess public skill signals, could further improve our understanding of information asymmetries about relevant skills.

References

- Acemoglu D and Pischke J (1998). Why Do Firms Train? Theory and Evidence, *Quarterly Journal of Economics* 113(1): 79-119.
- Acemoglu D and Pischke J (1999). The Structure of Wages and Investment in General Training, *Journal of Political Economy* 107(3): 539-572.
- Almeida R and Carneiro P (2009). The Return to Firm Investments in Human Capital, *Labour Economics* 16(1): 97-106.
- Altonji J and Pierret C (2001). Employer Learning and Statistical Discrimination, *Quarterly Journal of Economics* 116(1): 313-350.
- Andrews J and Higson H (2008). Graduate employability, 'soft skills' versus 'hard' business knowledge: A European study, *Higher Education in Europe* 33(4): 411-422.
- Arcidiacono P, Bayer P and Hizmo A (2010). Beyond signalling and Human Capital: Education and the Revelation of Ability, *American Economic Journal: Applied Economics* 2: 76-104.
- Autor D (2001). Why do Temporary Help Firms Provide Free General Skills Training?, *Quarterly Journal of Economics* 116(4): 1409-1448.
- Benson A (2013) Firm-sponsored general education and mobility frictions: Evidence from hospital sponsorship of nursing schools and faculty, *Journal of Health Economics* 32(1):

149-159.

- Benson G, Finegold D and Mohrmann S (2004). You Paid for the Skills now Keep Them: Tuition Reimbursement and Voluntary Turnover, *Academy of Management Journal* 47(3): 315-331.
- Bognanno M and Melero E (2016). Promotion Signals, Age and Education, *Journal of Economics and Management Strategy* 25(1): 111-132.
- Booth A and Zoega G (2004). Is Wage Compression a Necessary Condition for Firm-Financed General Training? *Oxford Economic Papers* 56(1): 88-97.
- Büsch V, Dahl S-A and Dittrich D (2009). An Empirical Study of Age Discrimination in Norway and Germany, *Applied Economics* 41(5): 633-651.
- Cappelli P (2004). Why do employers pay for college? *Journal of Econometrics* 121: 213-241.
- Chang C and Wang Y (1995). A Framework for Understanding Differences in Labor Turnover and Human Capital Investments, *Journal of Economic Behavior & Organization* 28(1): 91-105.
- Conti G (2005). Training, Productivity and Wages in Italy, *Labour Economics* 12: 557-576.
- Dearden L, Reef H and van Reenen J (2006). The Impact of Training on Productivity and Wages: Evidence from British Panel Data, *Oxford Bulletin of Economics and Statistics* 68(4): 397-421.
- De Grip A and Sauermann J (2012). The Effects of Training on Own and Co-Worker Productivity: Evidence from a Field Experiment, *Economic Journal* 122(560): 376-399.
- DeVaro J and Waldman M (2012). The Signalling Role of Promotions: Further Theory and Empirical Evidence. *Journal of Labor Economics* 30(1): 91-147.
- Dietrich H, Pfeifer H and Wenzelmann F (2016). The more they spend, the more I earn? Firms' training investments and post-training wages of apprentices, *Economics of Education Working Paper Series No. 0116*, University of Zurich.
- Farber H and Gibbons R (1996). Learning and Wage Dynamics. *Quarterly Journal of Economics* 111(4): 1007-1047.
- Fitzenberger B, Lickleder S and Zwiener H (2015). Mobility across Firms and Occupations among Graduates from Apprenticeship. *Labour Economics* 34: 138-151.
- Foster A and Rosenzweig M (1993). Information, Learning, and Wage Rates in Low-Income Rural Areas, *Journal of Human Resources* 28(4): 758-790.
- Gibbons R and Katz L (1991). Layoffs and Lemons. *Journal of Labor Economics* 9(4): 351-380.
- Göggel K and Zwick T (2012). Heterogeneous Wage Effects of Apprenticeship Training,

- Scandinavian Journal of Economics 114(3): 756-779.
- Greenwald B (1986). Adverse Selection in the Labour Market. *The Review of Economic Studies* 53(2): 325–347.
- Greenwald B and Glasspiegel R (1983). Adverse Selection in the Market for Slaves: New Orleans, 1830-1860, *Quarterly Journal of Economics* 98(3): 479-499.
- Harhoff D and Kane T (1997). Is the German apprenticeship system a Panacea for the U.S. labor market? *Journal of Population Economics* 2(1): 171–196.
- Heckman J, Stixrud J and Urzua S (2006): The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior, *Journal of Labor Economics* 24(3): 411-482.
- Heckman J and Kautz T (2012) Hard Evidence on Soft Skills, *Labour Economics* 19(4): 451-464.
- Hinz T and Mohrenweiser J (2017) The Effect of Regional Competition and Company-sponsored Training on the Productivity-Wage Wedge. Paper presented at the German Economic Society.
- Kaas L and Manger C (2012). Ethnic Discrimination in the Germany's Labour Market: A Field Experiment, *German Economic Review* 13(1): 1-20.
- Kahn L (2013). Asymmetric information between employers, *American Economic Journal: Applied Economics* 5(4): 165-205.
- Kahn L and Lange F (2014). Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures, *Review of Economic Studies* 81(4): 1575-1613.
- Katz E and Ziderman A (1990). Investment in General Training: The Role of Information and Labour Mobility. *The Economic Journal* 100(403): 1147-1158.
- Konings J and Vanormelingen S (2015). The Impact of Training on Productivity and Wages: Firm-Level Evidence, *Review of Economics and Statistics* 97(2): 485-497.
- Kosfeld R and Werner A (2012). Deutsche Arbeitsmarktregionen – Neuabgrenzung Nach Den Kreisgebietsreformen 2007–2011, *Raumforschung und Raumordnung* 70(1): 49-64.
- Kriechel B, Mühlemann S, Pfeifer H and Schütte S (2014). Works Councils, Collective Bargaining and Apprenticeship Training, *Industrial Relations* 53(2): 199-222.
- Jones DC, Kalmi P and Kauhanen A (2011). The effects of general and firm-specific training on wages and performance: evidence from banking. *Oxford Economic Papers* 64(1): 151-175.
- Laker DR and Powell JL (2011). The differences between hard and soft skills and their relative impact on training transfer, *Human Resource Development Quarterly* 22(1): 111-122.

- Lange F (2007). The Speed of Employer Learning. *Journal of Labor Economics* 25(1): 1-35.
- Leuven E (2005). The Economics of Private Sector Training: A Survey of the Literature, *Journal of Economic Surveys* 19(1): 91-111.
- Manchester CF (2010). Investment in General Human Capital and Turnover Intention, *American Economic Review: Papers and Proceedings* 100(3): 209-213.
- Manchester CF (2012). General Human Capital and Employee Mobility; How Tuition Reimbursement Increases Retention Through Sorting and Participation, *Industrial and Labor Relation Review* 65(4): 951-974.
- Mohrenweiser J (2016). Recruitment and Apprenticeship Training, *Industrielle Beziehungen* 23(1): 1-19.
- Mohrenweiser J and Zwick T (2009). Why Do Firms Train Apprentices? The Net Cost Puzzle Reconsidered, *Labour Economics* 16(6): 631-637.
- Mohrenweiser J and Zwick T (2015). Youth Unemployment after Apprenticeship Training and individual, Occupation and Training Employer Characteristics, *Journal of Economics and Statistics* 235(4+5): 418-432.
- Mohrenweiser J, Zwick T and Backes-Gellner U (2018). Poaching and Firm Sponsored Training: First Clean Evidence, *British Journal of Industrial Relations*, early view.
- Mueller B and Schweri J (2015). How Specific is Apprenticeship Training? Evidence from Inter-firm and Occupational Mobility after Graduation. *Oxford Economic Papers* 67(4): 1057-1077.
- Mühlemann S and Wolter S (2011). Firm-Sponsored Training and Poaching Externalities in Regional Labor Markets, *Regional Science and Urban Economics* 41(6): 560-70.
- Mühlemann S, Ryan P and Wolter S (2013). Monopsony Power, Pay Structure, and Training, *Industrial and Labor Relations Review* 66(5): 1097-114.
- Nickson D, Warhurst C, Commander J, Hurrell, SA and Cullen AM (2012). Soft skills and employability: Evidence from UK retail, *Economic and Industrial Democracy* 33(1): 65-84.
- Pattie M, Benson GS and Baruch Y (2006). Tuition Reimbursement, Perceived Organisational Support and Turnover Intention among Graduate Business School Students, *Human Resource Development Quarterly* 17(4): 423-442.
- Pfeifer C (2016). Intra-Firm Wage Compression and Cost Coverage of Training: Evidence from Linked Employer-Employee Data, *Industrial and Labor Relations Review* 69(2): 435-454.
- Pinkston J (2009). A Model of Asymmetric Employer Learning with Testable Implications, *Review of Economic Studies* 76: 367-394.

- Robles MM (2012). Executive perceptions of the top 10 soft skills needed in today's workplace, *Business Communication Quarterly* 75(4): 453-465.
- Ryan P, Wagner K, Teuber S and Backes-Gellner U (2013). Trainee Pay in Britain, Germany and Switzerland: Institutions, Market Forces, Market Power, *European Journal of Industrial Relations* 19(3): 201-220.
- Sauermann J (2015). Worker Reciprocity and the Returns to Training: Evidence from a Field Experiment, IZA working paper No. 9197, Bonn.
- Schönberg U (2007). Testing for asymmetric employer learning, *Journal of Labor Economics* 25(4): 651-692.
- Schönfeld G, Wenzelmann F, Dionisus R, Pfeifer H and Walden G (2010). *Kosten und Nutzen der dualen Ausbildung aus Sicht der Betriebe*. Bielefeld: Bertelsmann Verlag.
- Soskice D (1994). The German Training System: Reconciling Markets and Institutions, in L. Lynch (ed.) *International Comparisons of Private Sector Training*. University of Chicago Press.
- Von Wachter T and Bender S (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.
- Waldman M (1990). Up-or-Out Contracts: A Signaling Perspective, *Journal of Labor Economics* 8(2): 230-250.
- Waldman M (2016). The Dual Avenues of Labor Market Signaling, *Labour Economics* 41: 120-134.
- Waldman M and Zax O (2016). An Exploration of the Promotion Signaling Distortion, *Journal of Law, Economics, and Organization* 32(1): 119-149.
- Winkelmann R (1996). Employment Prospects and Skill Acquisition of Apprenticeship-Trained Workers in Germany, *Industrial and Labor Relations Review* 49: 658-672.

Tables

Table 1: Sample Selection

Step	Number of Observations
Successful merge between chamber and IEB data based on name initials, birthday, and start day of apprenticeship, last apprenticeship spell in IEB in the same year as graduation according to chambers	15,340
Age below 30 at graduation, apprenticeship ends in first or second quarter, no negative duration between end of apprenticeship and start of full-time employment	14,791
Earnings adjustment (earnings within Social Security contribution range and between 50% and 200% of occupational mean at the end of apprenticeship)	13,597
At least two apprentices per cell (establishment, occupation and graduation year)	5,813
Final sample (no missings in covariates)	5,733

Source: Ausbildungspanel Saarland 1998-2005.

Table 2: Definitions of variables and descriptive statistics

Variable	Definition (mean; sd)
Final marks (public signal)	Mark in the final apprenticeship exam multiplied by minus 1 (-2.91; 0.75)
Additional payments (private information)	Log daily earnings at last apprenticeship spell in € (3.08; 0.35)
Log earnings at first full-time employment	Log daily earnings at first full-time employment as skilled worker in € (4.04; 0.32)
Mover	Dummy variable equals 1, if the apprentice leaves the training firm and finds a skilled job in the training occupation within 30 days after apprenticeship completion (0.42; 0.49)
Female	Dummy variable equals 1, if apprentice is female (0.33; 0.47)
School certificate "Hauptschule" or school drop out	Dummy variable equals 1, if nine year school track graduate or does not pass the final school exam after nine school years (0.41; 0,49)
School certificate "Realschule"	Dummy variable equals 1, if ten year school track graduate (0.32; 0.47)
School certificate "Abitur"	Dummy variable equals 1, if twelve or thirteen year school track graduate (usually after 12 or 13 years in school) (0.26; 0.44)
Age	Age at apprenticeship graduation (21.44; 2.00)
German	Dummy variable equals 1, if German nationality (0.96; 0.19)
Establishment earnings level (training employer)	Average daily apprentice earnings at the end of apprenticeship in training firm in € (23.01; 6.78)
Number of employees (training employer)	Number of employees in the training employer (591; 1276)

N = 5,733; Source: Ausbildungspanel Saarland 1998-2005.

Table 3: Description of differences between stayers and movers

Variable	Movers	Stayers	t-test
Final marks (public signal)	-3.02	-2.83	5.93
Additional payments (private information)	2.947	3.169	10.17
Log earnings at first full-time employment	3.941	4.107	7.83
Female	0.330	0.324	0.25
Lower school track "Hauptschule" or drop-out	0.493	0.359	5.14
Medium school track "Realschule"	0.289	0.345	2.19
Upper school track "Abitur"	0.216	0.294	3.17
Age	21.65	21.29	6.75
German nationality	0.961	0.960	0.06
Establishment earnings level (training firm)	20.67	24.69	41.70
Number of employees (training firm)	293	804	42.14
Number of observations	2394	3339	

*** p < 0.01; ** p < 0.05; * p < 0.1; Source: Ausbildungspanel Saarland 1998-2005.

Table 4: Determinants of probability to move to another employer after training.

Variable	(1)	(2)	(3)	(4)	(5)
Additional payments (private information)	-0.318*** (-4.20)	-0.419*** (-4.54)	-0.861*** (-5.92)	-0.410*** (-7.34)	-0.468*** (-6.33)
Final marks (public signal)	-0.040** (-2.34)	-0.052*** (-2.92)	-0.107*** (-4.12)	-0.046*** (-4.40)	-0.044*** (-4.07)
Female		0.000 (0.01)	-0.025 (-0.67)		-0.014 (-0.83)
School certificate "Realschule"		-0.081*** (-2.64)	-0.147*** (-3.55)		-0.060*** (-3.40)
School certificate "Abitur"		-0.018 (-0.39)	-0.113* (-1.84)		-0.043 (-1.56)
Age		0.138* (1.92)	0.204* (1.93)		0.080* (1.73)
Age squared		-0.003 (-1.60)	-0.004 (-1.54)		-0.001 (-1.33)
German		0.168*** (3.34)	0.277*** (3.87)		0.120*** (4.08)
Number of employees		-0.006*** (-2.58)			
Establishment earnings level		0.012 (1.60)			
Sector, Year, Occupation	Yes	Yes	No	Yes	No
Cell Fixed Effects	No	No	Yes	No	Yes
R ²	0.08	0.10	0.26	0.26	0.68
Number of Observations	2291	2291	2291	5733	5733

Dependent variable: Dummy equals one if the apprentice leaves the training firm after graduation, zero otherwise; OLS regression; reference level for school certificate: "Hauptschule"; standard errors clustered at the cell-level; models (1)-(3) restricted to cells with movers and stayers, models (4) and (5) include all observations *** p < 0.01; ** p < 0.05; * p < 0.1. Source: Ausbildungspanel Saarland 1998-2005.

Table 5: Determinants of first full-time skilled employment log earnings for apprenticeship completers

	(1)	(2)	(3)	(4)	(5)	(6)
Additional payments (private information)	0.233*** (10.72)	0.084*** (2.85)	0.051 (1.28)	0.288*** (11.09)	0.137*** (4.66)	0.092** (2.33)
Final marks (public signal)	0.034*** (5.90)	0.030*** (5.06)	0.031*** (3.84)	0.021*** (3.65)	0.017*** (2.90)	0.016** (1.97)
Mover	-0.070*** (5.99)	-0.067*** (5.89)	-0.069*** (5.00)	0.355*** (3.29)	0.356*** (3.34)	0.343** (2.30)
Mover * additional payments				-0.106*** (3.50)	-0.107*** (3.59)	-0.097** (2.25)
Mover * final marks				0.033*** (2.81)	0.032*** (2.77)	0.038** (2.53)
Female		-0.016 (1.53)	-0.022* (1.81)		-0.017 (1.58)	-0.023* (1.87)
School certificate "Realschule"		0.027*** (3.49)	0.018* (1.81)		0.028*** (3.58)	0.017* (1.69)
School certificate "Abitur"		0.017 (1.29)	0.038** (2.26)		0.018 (1.41)	0.036** (2.12)
Age		-0.089*** (3.19)	-0.087** (2.32)		-0.090*** (3.30)	-0.085** (2.33)
Age squared		0.002*** (3.36)	0.002** (2.32)		0.002*** (3.47)	0.002** (2.32)
German		-0.017 (0.97)	-0.013 (0.51)		-0.017 (0.96)	-0.012 (0.49)
Number of employees in training establishment		0.001 (0.98)			0.001 (0.85)	
Training establishment earnings level		0.010*** (5.58)			0.010*** (5.61)	
Sector, year, occupation dummies	Yes	Yes	No	Yes	Yes	No
Cell Fixed Effects	No	No	Yes	No	No	Yes
R ²	0.48	0.49	0.72	0.48	0.49	0.72
Number of observations	5733	5733	5733	5733	5733	5733

Dependent variable: log daily earnings in first full-time employment after apprenticeship graduation; OLS regression; reference level for school certificate: "Hauptschule"; standard errors clustered at cell-level; *** p < 0.01; ** p < 0.05; * p < 0.1; Source: Ausbildungspanel Saarland 1998-2005.

Table 6: Determinants of first full-time skilled employment log earnings, separate control for earnings differences and marks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Additional payments (private information)	0.091*** (3.05)	0.054 (1.35)			0.136*** (4.60)	0.085** (2.15)		
Final marks (public signal)			0.031*** (5.19)	0.032*** (3.85)			0.021*** (3.59)	0.018** (2.26)
Mover	-0.070*** (6.07)	-0.072*** (5.18)	-0.069*** (6.02)	-0.071*** (5.16)	0.206** (2.25)	0.160 (1.22)	0.002 (0.07)	0.024 (0.53)
Mover * additional earnings					-0.089*** (3.05)	-0.075* (1.76)		
Mover * final marks							0.025** (2.18)	0.033** (2.23)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Establishment controls	Yes	No	Yes	No	Yes	No	Yes	Non
Sector, year, occupation	Yes	No	Yes	No	Yes	No	Yes	No
Cell Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
R ²	0.49	0.72	0.49	0.72	0.49	0.72	0.49	0.72
Number observations	5733	5733	5733	5733	5733	5733	5733	5733

Dependent variable: log daily earnings in first full-time employment after apprenticeship graduation; OLS regressions; standard errors clustered at cell-level; individual controls: gender, school certificate, age, nationality; establishment controls: median earnings and number of employees; *** p < 0.01; ** p < 0.05; * p < 0.1; Source: Ausbildungspanel Saarland 1998-2005.

Table 7: Skilled entry earnings determinants of employer movers: cells with demand shock only

Variable	(1)	(2)
Additional payments (private information)	0.039 (0.14)	0.033 (0.25)
Final marks (public signal)	0.024 (0.28)	0.074* (1.67)
Individual controls	Yes	Yes
Cell Fixed Effects	No	No
R ²	0.66	0.69
Number of observations	152	507

Dependent variable: log daily earnings in first full-time employment after apprenticeship graduation; OLS regression; standard errors clustered at the cell-level; individual controls: gender, school certificate, age, nationality, year, occupation and sector; model (1) restricted to cells with retention rate 50% below long-term average, model (2) restricted to cells with retention rate 20% below long-term average; *** p < 0.01; ** p < 0.05; * p < 0.1; Source: Ausbildungspanel Saarland 1998-2005.

Appendix A: Validation for the interpretation of relative earnings differences as indicator for skills within an establishment/ occupation/ cohort cell

In this section, we present a couple of additional arguments for our interpretation that relative earnings in comparison to peers within an establishment/occupation/cohort cell during the last spell of apprenticeship are an indicator for apprentice skills.

Relative earnings of apprentices at the end of training have predictive power. They predict first full-time earnings of stayers (Mohrenweiser et al., 2018). A Spearman Rank Correlation test shows that the earnings rank remains stable between the end of the apprenticeship and the first full-time employment of stayers within a cell (Mohrenweiser et al., 2018). When skilled entry earnings reflect the market value of an employee, we can conclude that our measure reflects individual skill differences of apprentices at the end of the apprenticeship period when training firms know the relative productivity of their apprentices.

We have to discuss why firms differentiate earnings between apprentices, despite earnings being determined by collective bargaining or other rules that prohibit undercutting certain earnings levels. Our argument is that training firms use the information on skills by voluntarily sharing a part of the additional rent created by high performing apprentices (Farber and Gibbons, 1996). This could give training enterprises a head start after the end of the apprenticeship period because high performing apprentices feel more obliged to stay or they are more loyal according to gift exchange considerations (Akerlof, 1984). Supporting evidence for this hypothesis is that additional payments strongly increase absolutely and relatively just before apprenticeship termination and the variance within homogeneous cells of apprentices is very low during the first years of apprenticeship (Mohrenweiser et al., 2018).

In interviews with personnel managers, we collected further evidence for the hypothesis that extra payments can explain the wage dispersion within homogeneous groups of apprentices in the same occupation, establishment, and cohort cell. We found evidence that extra payments are positively related to skills – also to hard-to-observe skills. HR managers confirm that all apprentices in the same training year receive the same base salary, according to collective agreements. However, many firms voluntarily pay supplementary allowances for good performance by their apprentices. For example, many firms send some apprentices abroad for some weeks in the final training year. Apprentices who are sent abroad are a positive selection, according to our interview partners. Working abroad is a

reward for good performance and it is accompanied by additional payments according to collective agreements. Additional evidence for the existence of additional payments for apprentices is given by the latest Federal Institute of Vocational Education and Training costs/benefits study, a survey of the costs and benefits structure of German apprenticeship training firms conducted every six years (Jansen et al., 2015). This survey indicates that 45% of training firms use additional payments. These payments can include for example accommodation and travel allowances for apprentices working abroad for some time. In addition, there is evidence for explicit financial bonus rules for good marks at the vocational school or for good performance at work: Ryan et al. (2013) find individual and group-related performance pay for apprentices in 13 out of the 18 engineering and retailing firms. Backes-Gellner and Oswald (2014) find that a number of firms pay bonuses for good marks in vocational schools. Other examples mentioned are payments for work during the weekend or overtime hours.¹⁹ Again, the payments for these activities are typically considered as reward for good performance. Finally, fringe benefits for example for apprentices commuting larger distances or accommodation allowances may be a reason for the observed earnings differences. Even these payments may be correlated with work-related performance if training firms are willing to grant these benefits for extraordinarily able apprentices only. We therefore can conclude that unions and works councils accept differences in earnings of apprentices within one employer and occupation when these differences are according to transparent rules. These rules relate earnings differences to individual productivity differences. Objective rules reduce the negative effect of additional payments on the work motivation of apprentices who do not get them. Objective rules also reduce possible expectations for higher future compensation by the recipients.

Our measure filters out additional group payments frequently implemented on the establishment level and instead captures only the individual differences between apprentices controlling for occupation, training firm and apprenticeship completion cohort. Group payments for all apprentices within a group of apprentices in one firm, occupation and apprenticeship year do not add variation.

Appendix B: Mathematical derivation of model

We use the following notation: The individual private information measure “additional payment” b can take two expressions, high ($b = H$) and low ($b = L$). The individual public signal²⁰ measure “mark” g can also take two expressions, excellent ($g = E$) and average ($g = A$)²¹. In our basic model, we assume that the impact of information g on productivity y is the same for training participants with a low and a high level of b – this assumption is relaxed later and replaced by the possibility that g and b are complements²². We therefore assume that an excellent mark adds a constant c to productivity in comparison to an average mark. Finally, we denote with p the probability that a training participant received an additional payment. We assume that public and private information on skills are positively related, or in other words training participants with excellent marks received a high additional payment with a higher probability than training participants with an average mark ($p^A < p^E$).²³ Table A1 gives an overview of our model.

Table B1: Specification of Productivity y

	Average marks G	Excellent marks G
Low additional payment b	$y_L^A = b_L$	$y_L^E = b_L + c$
High additional payment b	$y_H^A = b_H$	$y_H^E = b_H + c$
Proportion of additional payment p	p^A	p^E

At the end of the training period, all employers and training participants observe the marks. Marks measure not all dimensions of productivity. The missing dimension of productivity is known to training employers and to the training participant. Outside employers know the general probability of employees with average and excellent marks to receive an additional

²⁰ In the original Schönberg model, the easily to observe dimension is called “education” and the hard to observe dimension is called “cognitive ability and communication skills”.

²¹ We do not use “high” and “low” for the differentiation of marks to ease reading.

²² Schönberg (2007) proves in her appendix A6 that the results also hold if public and private information is a complement. As a consequence, the adverse selection should be stronger for apprenticeship graduates with high public signals, a result also derived by Acemoglu and Pischke (1998). In addition, the difference between the impact of private skill information on entry wage offers of training and outside employers should be stronger for apprenticeship graduates with high private skill information.

²³ The additional payment reflects the private information of the training employer. We have to assume that the outside employer would assess them similarly given it had the same information. This is a common assumption in the literature, see the discussion in Kahn (2013) on the general applicability of AFQT as indicator of productivity.

payment p and they observe a common noisy signal s about this privately known dimension of individual productivity. An example for this signal may be the performance in a job interview (Lange, 2007). This signal can be obtained by all (interested) labour market participants and is independent from marks.²⁴

In our simple model, the potential signal s about private skill information takes only two values, good ($s = G$) and bad ($s = B$). The outside employer uses the signal in order to infer the additional payments received by the training participant. This inference is correct if a training participant with a good signal also receives a high additional payment and a training participant with a bad signal receives a low additional payment. In our model, $q \geq 0.5$ denotes the probability that outside employers can observe the correct signal. According to Bayes' Law, the outside employers compute the probability of having a low additional payment training participant when the information is good or bad, given the mark is excellent or average:

$$p^{k,B} = \Pr(L | k, B) = \frac{p^k q}{p^k q + (1 - p^k)(1 - q)}, \quad k = A, E, \quad (1)$$

$$p^{k,G} = \Pr(L | k, G) = \frac{p^k (1 - q)}{p^k (1 - q) + (1 - p^k)q}, \quad k = A, E. \quad (2)$$

Here, $q = 1$ (or public information for training and outside employers on the productivity of training participants after training) means that all low additional payment training participants generate a bad signal and that all high additional payment training participants generate a good signal, in other words: $\Pr(L | k, B) = 1$ and $\Pr(L | k, G) = 0$. When information is completely asymmetric (or $q = 0.5$), the probability to get a high additional payment training participant is assessed purely by the a-priori probability depending on the distribution of high additional payment and low additional payment training participants, $\Pr(L | k, B) = \Pr(L | k, G) = p^k$.

²⁴ It is important to note that the training and the outside employers receive the same signal. In our case it would be for example sufficient that the training employer knows that the apprenticeship graduate is able to perform well in a job interview at an outside employer and hereby is able to reveal his or her additional payment, the training firm therefore does not necessarily perform a job interview by itself in order to reveal the signal.

Training participants may have had experienced disutility θ during training that is unknown to the training employer – it therefore cannot compensate the training participant for individual disutility in order to retain the training participant. The disutility shock is independent of the additional payment and mark (and the signal s). We assume that the disutility in the training firm is drawn from a distribution with the cumulative distribution function G with an associated probability function g based on $[\underline{\theta}, \bar{\theta}]$ and $\bar{\theta} > 0$. The function

$1-G(\theta)$ is log-concave, which means that $\frac{g(\theta)}{1-G(\theta)}$ is non-decreasing in θ , i.e. it continuously

increases with θ but with a decreasing rate. The expected value of disutility is zero.²⁵ Please note that even if the apprentice knows before the end of the apprenticeship training he or she wants to move to another employer after apprenticeship completion, this should not lead to a reduction in training efforts. The main reason is that apprentices need good public signals and if possible a retention offer from the training firm in order to obtain an adequate job at another quality employer.

The training participant receives an earnings offer from the training employer after completing training (w) and from the outside employer (v) and takes his or her disutility at the current employer (θ) into account for the decision whether to stay with the training employer or not. In other words, the training participant stays with the training employer only if $w + \theta > v$.

The training employer has the advantage that it can make counteroffers to those training participants it wants to retain after training when $w < v$ ²⁶. The outside employer does not observe the offer of the training employer w and therefore a bidding war on training participants after training completion is not possible (Pinkston, 2009).

The training employer instead maximises the difference between productivity y_i^k and $w_i^{k,s}$, taking marks g , the productivity signal s , and the probability that the training participant stays with the training employer given v and the distribution of θ into account. The

²⁵ The literature has analysed additional reasons for individuals to leave a firm, such as social ties, peer effects (Wegener, 1991), life events (Lee and Mitchell, 1994) and several sources of job dissatisfaction (March and Simons, 1958; Mobley, 1977; Griffith et al., 2000).

²⁶ Note that training participants with $\theta < 0$ leave the training firm if $w = v$ and therefore the training firm has to offer higher earnings than the outside firm in order to reduce the risk of losing the training participant.

probability that the training participant stays can be written as $\Pr(\text{stay} | g, s) = 1 - G(v^{g,s} - w^{g,s})$ with $g = A, E$, and $s = B, G$. The outside employer offers earnings taking into account marks g and signal s . The training employer therefore maximises profits by setting earnings according to:

$$\max w_b^{g,s} \left[1 - G(v^{g,s} - w_b^{g,s}) \right] (y_b^g - w_b^{g,s}) \quad \text{with } b = L, H; g = A, E; s = B, G.$$

The first-order condition shows that the training employer pays according to the productivity of the training participant minus an expression that decreases with disutility θ :

$$w_b^{g,s} = y_b^g - \frac{1 - G(v^{g,s} - w_b^{g,s})}{g(v^{g,s} - w_b^{g,s})}. \quad (3)$$

Outside employers pay the expected productivity of those training participants who switch firms after they received both offers. Outside employers can freely offer earnings and even poach training participants after training, i.e. pay a wage that is higher than what the training participant would have earned when staying in the training firm. If training participants do not have to refund training investments when they directly leave the training firm after training and several outside employers are interested in the participant, outside employers cannot make a profit in the long run. The outside earnings offered can be written as:

$$v^{g,s} = \frac{\Pr(L | \text{move}, s)y_L^g + \Pr(H | \text{move}, s)y_H^g}{\Pr(L | \text{move}, s) + \Pr(H | \text{move}, s)} = \frac{p^{g,s}G(v^{g,s} - w_L^{g,s})y_L^g + (1 - p^{g,s})G(v^{g,s} - w_H^{g,s})y_H^g}{p^{g,s}G(v^{g,s} - w_L^{g,s}) + (1 - p^{g,s})G(v^{g,s} - w_H^{g,s})} \quad (4)$$

with $p^{g,s}$ according to equations (1) and (2).

Now, we can derive implications of public and asymmetric information on the coefficients of the private and public productivity information for the earnings of leaving and staying training participants after training. When information is public ($q = 1$), signal s completely reveals the additional payment and therefore the private skills assessment of the training firm, i.e. $p^{k,B} = 1$ and $p^{k,G} = 0$. Outside employers therefore offer $v_H^k = y_H^k$ if the training participant sends a good signal, and $v_L^k = y_L^k$ if the training participant sends a bad signal. As the productivity difference between high and low additional payment training participants is

a constant ($b_H - b_L = \Delta$) that is independent of marks, the training employer offers worker's productivity minus this constant, $w_b^k = y_b^k - \Delta$. In other words, the offer w increases by the full amount of the productivity difference between low and high additional payment training participants. As a consequence, also the difference between training and outside employers' offers is equal and low and high additional payment training participants have the same probability of leaving the training employer:

$$v_L^k - w_L^k = v_H^k - w_H^k. \quad (5)$$

When the information is asymmetric ($q < 1$), the difference in the offer of outside employers between high and low additional payment training participants is smaller than Δ because a certain share of training participants with a good signal obtained a low additional payment. Hence, high additional payment training participants have a smaller incentive to leave the training employer voluntarily than low additional payment training participants. This leads to adverse selection of movers and outside employers anticipate that training firms use their informational advantage and consider movers as having a higher risk to be "lemons". The formal argument is as follows: The probability that a low additional payment training participant leaves the training employer equals:

$$qG(v^B - w_L^B) + (1-q)G(v^G - w_L^G), \quad (6)$$

and the probability that a high additional payment training participant leaves the training employer equals:

$$qG(v^G - w_H^G) + (1-q)G(v^B - w_H^B). \quad (7)$$

Therefore, the following two conditions are sufficient that the quit rate of low additional payment training participants is higher than the rate of high additional payment training participants:

$$G(v^B - w_L^B) - G(v^G - w_H^G) > 0 \Leftrightarrow w_H^G - w_L^B > v^G - v^B, \quad (8)$$

$$G(v^G - w_L^G) - G(v^B - w_H^B) > 0 \Leftrightarrow w_L^G - w_H^B < v^G - v^B. \quad (9)$$

When $q < 1$, low additional payment training participants are more likely to leave the training employer than high additional payment training participants. This is unconditional on the signal observed by the outside employer. The probabilities that a low respectively a high additional payment training participant leaves the training employer are according to equations (6) and (7). Therefore, inequalities (8) and (9) have to hold in order to get a higher quit rate of a low additional payment training participant than of a high additional payment training participant. When we assume $q = 0.5$, the outside employer offers the same for low and high additional payment training participants, the training employer however offers higher earnings for high than for low additional payment training participants. Therefore, inequality (8) always holds because $w_H^G - w_L^G > 0$ and $v^G - v^B = 0$. If $q = 1$, $w_H^G - w_L^G = v^G - v^B = a_H - a_L$. If we can show that for $0.5 \leq q < 1$, the difference between the earnings mark-up for high additional payment training participants and low additional payment training participants increases stronger with better information quality q of the signal for outside employers than for training employers ($d[(w_H^G - w_L^G) - (v^G - v^B)]/dq < 0$), then inequality (8) holds irrespectively of the signal quality. This is the case, if the difference between the earnings offers for high additional payment training participants decreases with signal quality ($d(w_H^G - v^G)/dq < 0$) and at the same time, the difference between the offers for low-additional payment training participants increases with the signal quality ($d(v^G - w_L^B)/dq < 0$). If we totally differentiate the first-order earnings setting condition (3) for the training firms, we obtain:

$$\frac{dw_i^s}{dq} = \frac{\left[g_i^{s2} + (1 - G_i^s) g_i^{s'} \right] / g_i^{s2}}{1 + \left\{ \left[g_i^{s2} + (1 - G_i^s) g_i^{s'} \right] / g_i^{s2} \right\}} \frac{dv^s}{dq}.$$

By log-concavity of G , the expressions in the square brackets are larger than zero and dv^s/dq is multiplied by a number smaller than one.

Next, consider inequality (9). For $q = 0.5$, the condition is satisfied, because the left-hand side is zero and the right-hand side is negative. The analogous requirements for $0.5 \leq q < 1$ are $d(v^G - w_L^G)/dq < 0$ and $d(w_H^B - v^B)/dq < 0$. That these conditions hold under the assumption of log-concavity of G has been shown for inequality (8). Therefore, low additional payment training participants are more likely to leave the training employer unconditional on the signal observed under asymmetric information. In other words, with

public information, low and high additional payment training participants have the same probability to quit. With information asymmetry, the chances of low additional payment training participants to leave the training employer are higher.²⁷

We can turn now to the main feature of the model – the influence of additional payments and marks on earnings after training under public and asymmetric information. From equation (5), it is clear that the additional payment during training has the same impact on outside and training employer earnings offers after training under public information. With asymmetric information, earnings offers after training of training employers vary more with additional payments than earnings offers of outside employers. The difference between expected earnings offers of training firms w for high and low additional payment training participants equals:

$$E[w|H] - E[w|L] = [qw_H^G + (1-q)w_H^B] - [qw_L^B + (1-q)w_L^G].$$

The difference between the earnings offers of outside firms v for high and low additional payment training participants however equals:

$$E[v|H] - E[v|L] = [qv^G + (1-q)v^B] - [qv^B + (1-q)v^G].$$

Therefore, $w_H^G - w_L^B + v^B - v^G > 0$ and $v^G - w_L^G + w_H^B - v^B > 0$. These inequalities correspond to the conditions for the quit rate of low additional payment training participants to be higher than the quit rate of high additional payment training participants in the asymmetric information case, compare equations (6) and (7).

The impact of marks on earnings after training is again the same for training and outside employers under public information. Conditional on the additional payment, the difference between outside earnings offers for training participants with average and excellent marks is the full productivity mark-up ($y_i^E - y_i^A = c$). In the case of asymmetric information, the earnings offers of outside employers to training participants with a good mark not only reflect the productivity-enhancing effect of having a good mark, but also the fact that

²⁷ DeVaro and Waldman (2012) argue analogously that adverse selection of employer movers can be detected by a negative influence of performance ratings on the probability to quit given all other observable employee characteristics such as schooling, tenure, and job level. We use voluntary additional payment as a measure of performance ratings.

training participants with good marks have a higher probability to have a high additional payment p^E than training participants with low marks p^A .

Analogously to Appendix A5 in Schönberg (2007), the proof is presented for low-productivity training participants. The average difference between earnings offers of training employers for training participants with excellent and average marks equals:

$$E[w | E, L] - E[w | A, L] = q(w_L^{E,B} - w_L^{A,B}) + (1-q)(w_L^{E,G} - w_L^{A,G}). \quad (10)$$

The analogous wage difference between earnings offers of the outside employers equals:

$$E[v | E, L] - E[v | A, L] = q(v^{E,B} - v^{A,B}) + (1-q)(v^{E,G} - v^{A,G}). \quad (11)$$

We therefore need to show that the differences (10) < (11) or in other words $w_L^{E,s} - w_L^{A,s} < v^{E,s} - v^{A,s}$ for $s = G, B$. Since an excellent mark has the same impact on productivity for low-productivity and high-productivity training participants, we find that:

$$\frac{dv^{E,s}}{ds} = \frac{dw^{E,s}}{ds} = 1.$$

The total differentiation of the earnings maximisation problem of training firms (equation (3)) leads to:

$$\frac{dw_i^{E,s}}{d(1-p^E)} = \frac{[g_i^{E,s^2} + (1-G_i^{E,s})g_i^{E,s'}] / g_i^{E,s^2}}{1 + \left\{ [g_i^{E,s^2} + (1-G_i^{E,s})g_i^{E,s'}] / g_i^{E,s^2} \right\}} \frac{dv^{E,s}}{d(1-p^E)}.$$

Log-concavity of G gives $\frac{dw_i^{E,s}}{d(1-p^E)} < \frac{dv^{E,s}}{d(1-p^E)}$. Therefore an excellent mark has a higher impact on the earnings offers after training of the outside employers than on the earnings offers of training employers for low-productivity and high-productivity training participants.

An apprentice might also engage in activities that deliberately sabotage his or her public signal or private information because they want to work in another occupation or for another employer after completing their apprenticeship. A high public signal is necessary however in order to obtain a take-over and/or a high earnings offers if an apprenticeship

completer wants to change the occupation. The public and the private signals are the basis for the job and earnings offers because the employers induce from them expected ability and productivity in the new occupation. Even if they want to change the employer and there is asymmetric private information, apprentices should not neglect investments into private information. The main reason is that public signals and private information are positively correlated. It therefore negatively affects your public signal if you sabotage your private information dimension that helps you to get a positive outside earnings offer. As a consequence, only few apprentices sabotage their private information and the private information reveals an ability dimension. We observe this because high private information is negatively correlated with mobility and positively with entry earnings of stayers.

Please note that our simple theoretical model does not include firm-specific human capital and firm heterogeneity. Both factors potentially influence the value of information asymmetry for the training firm and the decision of the apprenticeship completer to change the employer. We leave it for future research to analyse the impact of these elements on the apprenticeship training system.

Appendix C: Additional evidence about specificity of additional training during apprenticeship

Table C1 in this appendix shows the extent and content of additional training beyond the requirements during apprenticeship using the 2011 wave of the *Ausbildungspanel Südwestmetall* (apprenticeship survey of the employer association for the metal-working industry in Southwest Germany). The employer association asks its member firms about apprenticeship related questions once a year. The response rate is very high with around 70% per year (for more information about the survey, see Mohrenweiser, 2012). The 2011 survey entails a question whether apprenticeship training firms in metal working blue-collar manufacturing occupations train additional skills that go beyond the respective training curriculum.

Metal-working occupations are usually considered to be among the most expensive training investments for firms (Mohrenweiser and Zwick, 2009; Schönfeld et al., 2010) and these occupations usually serve as examples for firm-specific skills linked to specific machines and production processes (Soskice, 1994).

Table C1: Training of additional skills during apprenticeship

	Number of firms	Per cent
stick to training curriculum	237	47.4
additional skills training	263	52.6
if additional skills, ... (multiple answers possible)		
... skills from other blue-collar metal working occupations	80	16.0
... skills from white-collar training occupations	58	11.6
... foreign language	70	14.0
... soft skills	182	36.4
... something else	65	13.0
... additional skills are externally certified	33	6.6

N = 500 apprenticeship training firms, blue-collar manufacturing occupations only; multiple answers possible if additional training offered. Source: *Ausbildungspanel Südwestmetall*, 2011.

References for Appendices

- Acemoglu D and Pischke J (1998). Why Do Firms Train? Theory and Evidence, *Quarterly Journal of Economics* 113(1): 79-119.
- Akerlof G (1984). Gift Exchange and Efficiency-Wage Theory: Four Views, *American Economic Review* 74(2): 79-83.
- Backes-Gellner U and Oswald Y (2014). Learning for a Bonus; How Financial Incentives Interact with Preferences, *Journal of Public Economics* 118(C): 52-61.
- DeVaro J and Waldman M (2012). The Signalling Role of Promotions: Further Theory and Empirical Evidence. *Journal of Labor Economics* 30(1): 91-147.
- Farber H and Gibbons R (1996). Learning and Wage Dynamics. *Quarterly Journal of Economics* 111(4): 1007-1047.
- Griffeth RW, Hom PW and Gaertner S (2000). A Meta-Analysis of Antecedents and Correlates of Employee Turnover: Update, Moderator Tests, and Research Implications for the Next Millennium. *Journal of Management* 26(3): 463-488.
- Jansen A, Pfeifer H, Schönfeld G and Wenzelmann F (2015). Apprenticeship Training in Germany Remains Investment-Focused – Results of the BIBB Cost-Benefit Survey 2012/13, BIBB Report 1/2015, Bonn.
- Kahn L (2013). Asymmetric information between employers, *American Economic Journal: Applied Economics* 5(4): 165-205.
- Lange F (2007). The Speed of Employer Learning, *Journal of Labor Economics* 25(1): 1-35.
- Lee TW and Mitchell RR (1994). An Alternative Approach: The Unfolding Model of Employee Turnover, *Academy of Management Review*, 19(1): 51–89.
- March JG and Simon HA (1958). *Organizations*. New York: Wiley.
- Mobley WH (1977). Intermediate Linkages in the Relationship between Job Satisfaction and Employee Turnover. *Journal of Applied Psychology*, 62(2): 237-240.
- Mohrenweiser J (2012). Which firms train disadvantaged youth? *Empirical Research in Vocational Education and Training* 4(2): 115-130.
- Mohrenweiser J and Zwick T (2009). Why Do Firms Train Apprentices? The Net Cost Puzzle Reconsidered, *Labour Economics* 16(6): 631-637.
- Mohrenweiser J, Zwick T and Backes-Gellner U (2018). Poaching and Firm Sponsored Training: First Clean Evidence, *British Journal of Industrial Relations*, early view.
- Pinkston J (2009). A Model of Asymmetric Employer Learning with Testable Implications, *Review of Economic Studies* 76: 367-394.

- Ryan P, Wagner K, Teuber S and Backes-Gellner U (2013). Trainee Pay in Britain, Germany and Switzerland: Institutions, Market Forces, Market Power, *European Journal of Industrial Relations* 19(3): 201-220.
- Schönberg U (2007). Testing for asymmetric employer learning, *Journal of Labor Economics* 25(4): 651-692.
- Schönfeld G, Wenzelmann F, Dionisus R, Pfeifer H and Walden G (2010). *Kosten und Nutzen der dualen Ausbildung aus Sicht der Betriebe*. Bielefeld: Bertelsmann Verlag.
- Soskice D (1994). The German Training System: Reconciling Markets and Institutions, in L. Lynch (ed.) *International Comparisons of Private Sector Training*. University of Chicago Press.
- Wegener B (1991). Job Mobility and Social Ties: Social Resources, Prior Job, and Status Attainment, *American Sociological Review* 56(1): 60-71.