Working Paper No. 121

Information Advantages of Training Employers Despite Credible Training Certificates

Jens Mohrenweiser

Gabriele Wydra-Somaggio

Thomas Zwick

Universität Zürich
IBW – Institut für Betriebswirtschaftslehre

UNIVERSITÄT BERN

Working Paper No. 121

Information Advantages of Training Employers Despite Credible Training Certificates

Jens Mohrenweiser

Gabriele Wydra-Somaggio

Thomas Zwick

January 2017, this version: April 2017

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

www.economics-of-education.ch

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)

This version April 2017

ABSTRACT:

We show that employers that invest in general human capital can positively select the employees they retain although training contents are transferable and visible to all employers (they are certified by credible external institutions). To solve the puzzle, we distinguish between soft and hard skills. Hard skills are tested in graded final exams and can be signalled to outside employers while soft skills are typically not tested and cannot be signalled. Therefore, the information advantage about soft skills can explain why employers keep a positive selection of training participants and invest in certified and transferable skills.

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

Scholars frequently argue that 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 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). The information advantage allows training firms to get a return on investment because they can retain the trained employees they want to keep by paying them below productivity but above market earnings.[1]

The information advantage vanishes, however, if credible external institutions reveal the individual performance with (graded) training certificates and final exams (Arcidiacono et al., 2010). Certificates and final exams after training are widespread: examples are employer-sponsored Bachelor, Master and MBA courses (Cappelli, 2004; Benson et al., 2004; Pattie et al., 2006; Manchester, 2010, 2012; Benson, 2013), language 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; Muehlemann et al., 2013; Kriechel et al., 2014). In recent years, IT providers such as Microsoft, SAP or Oracle increasingly certify proficiency in their application software. All of these certificates overcome information asymmetries about training content and reduce firms' incentive to invest in training (Katz and Ziderman, 1990). This leads to a puzzle: how can training firms gain an information advantage if training courses are credibly certified by external institutions?

This paper provides a novel solution for this puzzle. It distinguishes between two types of skills: hard and soft skills. Hard skills are technical skills related to specific work tasks. They are quantifiable and the level of proficiency in hard skills is typically measured in standard exams after training or education (Nickson et al., 2012). Certificates and final exams allow

[1] A number of empirical papers have recently shown a wedge between wages and performance after company-sponsored training. Training has a larger effect on productivity than on wages in several European countries (Konings and Vanormelingen, 2015; 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)

training and outside firms to assess relative performance in hard skills in comparison to other training participants. Skill dimensions that are not hard skills but nevertheless valuable on the labour market are frequently called soft skills.[2] Soft skills comprise a broad range of intrapersonal and interpersonal skills such as for example communication, courtesy, integrity, responsibility, team-work, and work-ethics (Laker and Powell, 2011; Robles, 2012). Soft skills are typically hard to quantify and measure in traditional training certificates. Nevertheless, soft skills are generally considered to be more important than hard skills for employability, performance, and learning, particularly for entry-level jobs (Andrews and Higson, 2008; Nickson et al., 2012). Despite of their relevance for performance, soft skills are typically not part of training certificates. Hence, training firms might gain an information advantage about soft skills which remain private information even if training is certified by external institutions. In other words, information asymmetry on soft skills can generate the necessary information advantage that creates the wedge between earnings at training and outside firms.

To analyse information advantages of training firms after certified general training, we first introduce an analytical framework based on a model developed by Schönberg (2007). The model nests public and private (or asymmetric) information on skills. In the public information case, training and outside firms can equally assess hard and soft skills. In the asymmetric information case, the training firm knows more about soft skills than the outside firm. The model allows us to derive hypotheses that identify information asymmetries for soft skills between training and outside firms based on earnings offers and job mobility after training.

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 transferable and visible skills certified in final exams by external and independent institutions. The strongly regulated German apprenticeship training system provides training for around two thirds of each cohort at the start of their career. We use a unique database including grades of the final apprenticeship certificate among other

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

important information from administrative sources for the entire population of several cohorts of apprenticeship completers in one German federal state. The final grades serve as a public signal about hard skills. Since soft skills are typically not measured by standard exams and are not included for example as psychological tests in our data[3], we have to rely on an indirect measure. We follow the approach by Kahn (2013) and Kahn and Lange (2014) and assume that training employers reward hard as well as soft skills. Differences in earnings between observationally equal employees reflect subjective productivity assessments of an employer that also reveals latent skills. Hence, we use small additional payments at the end of an apprenticeship as a proxy for soft skills. More specifically, we take earnings differences in cells of apprentices in the same training firm, occupation, and cohort as indicator of performance. As we condition on grades in our earnings estimations, the additional payment identifies performance dimensions (latent soft skills) that are not captured in apprenticeship certificates (hard skills).

We show that soft skills indeed induce an information advantage for training firms. Hard and soft skills have distinct explanatory power for entry earnings after training completion. Entry earnings of stayers after training are correlated with final grades and with the earnings position during training but entry earnings of movers after training are only correlated with grades but not with the earnings position. This means that soft skills cannot be signalled to outside firms. As a consequence of the information advantage on soft skills, training firms are able to keep a positively selected group of apprentices after training.

The paper is structured as follows. The next section presents a model on the consequences of public and private information on hard and soft skills on the mobility after training and the determinants of earnings after training. The third section discusses our empirical strategy and the fourth section presents our data and variables. The fifth section shows the empirical results, this sixth discusses alternative explanations for our results and the last section concludes.

[3] Soft skills are usually measured in surveys with the help of psychological tests such as the Rotter Locus of Control Scale or the Rosenberg Self-Esteem Scale, for example employed in the National Longitudinal Survey of Youth, 1979 (NLSY79) (Heckman et al., 2006).

2 Theoretical framework

This section outlines a theoretical framework to identify whether training employers have the same information about hard and soft skills of training completers as outside firms (this is the public signal case) or whether training employers know more about trainees´ skills than outside firms (asymmetric information case).

We apply the model of Schönberg (2007) to the context that employers invest in transferable skills training but only hard-skills not soft skills are assessed in the final certificate.[4] The model consists of two periods. In the first period, training takes place and ends with a certificate.[5] 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. 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 employer offer plus an individual disutility shock incurred during training.[6] The disutility shock is not observable by the training employer but it has some expectations about the distribution of the disutility shock. Hence, the model generates some voluntary turnover for exogeneous reasons (similar to Acemoglu and Pischke, 1998; Autor, 2001).

The model distinguishes between two types of workers, high and low productivity workers, and between two skill dimensions, hard and soft skills. The training employer learns about the individual level of both dimensions during the first period. Hard skills are easy to observe for all market participants and public knowledge on the basis of the credible signal certificate. A high level of hard skills adds the same productivity to employees with low soft skills as to employees with high soft skills. The probability that the employee has high soft skills is higher when the employee has high hard skills – soft and hard skills are therefore positively correlated. The a-priori probabilities of high and low soft skills for the groups with high and low hard skills are known to outsider employers.

[4] In the appendix, we provide a more formal derivation that complements the intuitive account given in this section.

[5] In contrast to the model by Acemoglu and Pischke (1998), the amount of training provided is given, typically by the external provider.

[6] Examples for the disutility shock are problems with superiors or co-workers, the organizational climate in the training firm or changes in regional preferences (Acemoglu and Pischke, 1998). Cottini et al. (2011) have shown that an unsupportive line manager leads to a six percentage point increase in voluntary turnover probability.

The key question is whether the information about soft skills remains private or whether outside employers can reveal them. Outside employers need a reliable signal such as a job interview to infer soft skills (Lange, 2007). If this additional signal perfectly transmits the information on soft skills, outsider and training firms have the same (correct) information and soft skills become public information. If the additional signal is only white noise, the outside employer cannot infer anything about the soft skills of applicants in addition to the a-priori probability distribution of soft skills known for applicants with high and with low hard skill levels. In this case, soft skills are private information.

If outside employers can accurately assess soft skills, training and outside employers offer the entire additional productivity as extra payment for high hard and for high soft skills. This means that only the disutility shock allocates movers and stayers. As the disutility shock 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 soft skills, the outside employer is not able to fully pay for the additional productivity associated with high soft skills. In the asymmetric information equilibrium, the training firm can offer slightly higher earnings for completers with high soft 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 consider employer movers as having a higher risk of low soft skills or “lemons” (Greenwald, 1986; Gibbons and Katz, 1991).

Summing up, our model nests public and asymmetric information about soft skills on the labour market. We can derive the following hypotheses:

If soft and hard skills are public information, both skill types 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 about soft skills, hard skills have a stronger impact on the second-period earnings of movers than of stayers because the outside employers can only rely on the public signal. Soft skills have an impact on earnings of

stayers but not on earnings of movers. In particular, movers are negatively selected in comparison to stayers.

3 Empirical Strategy

Institutional Background

We use the German apprenticeship training system as an example for training in general human capital that ends with a generally accepted 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 in 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 therefore 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).

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 in order to get permission for apprenticeship training granted by the chambers of industry and commerce or the chambers of craft. Training curricula are published, tailor-made 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 which 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 graded by publicly funded and controlled vocational schools. It is important for our empirical strategy that apprentices receive graded certificates at the end of the training period because grades 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. Apprenticeship contracts are obviously no up-or-out contracts and in addition to that, a retention decision by the training firm is hard to observe by outside firms.[7] Outside firms therefore cannot easily verify whether a job applicant had an employment offer from the training firm or not.

Dependent variables: probability to move and entry earnings

We use two dependent variables: a binary variable of the status as staying or leaving apprenticeship completer and the entry earnings of apprenticeship completers. Entry earnings are measured in the first employment spells after graduation. These spells take clearly less than one year.[8] This means that a quick revelation of soft skills by a new employer cannot influence measured entry earnings.[9] The measure of entry earnings for leavers and stayers is also not biased by the influences of 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 those apprenticeship

[7] Some collective bargaining agreements and some voluntary pacts between employers and unions entail retention clauses for apprenticeship completers. It seems however that these agreements did not have an impact on the discretion of training firms to get rid of apprenticeship completers (Bispinck et al., 2002). The main reason is that there are no sanctions for employers that do not offer skilled labour contracts to all completers but sufficiently unattractive contracts with respect to salary, contract length, job or location that lead apprenticeship completers to look for another employer.

[8] The maximum spell length is less than one year. Employers have to report earnings at least once a year to the social security but status changes from apprenticeship to regular employment require an additional report (a spell in our data) during a calendar year. The average duration of the entry earnings spell is 178 days.

[9] 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, see Lange (2007).

completers who stay with them. Therefore, earnings of leavers and stayers are determined at exactly the same point in time.

Information on hard skills

We measure hard skills using the final grades apprenticeship completers receive in their final exams at the end of the training period. Their certificate provides easily accessible and unbiased publicly available information.[10] Outside employers can assess the hard skills of each apprentice on basis of the grades in several practical and theoretical subjects.[11] Since most employers who hire apprenticeship completers train themselves (Mohrenweiser, 2016), they have experience in assessing grades reported in apprenticeship certificates and supplementary information such as school grades, occupation, and selectivity of the training employer (Wagner and Zwick, 2012; Mohrenweiser and Zwick, 2015).[12]

Information on soft skills

In contrast to hard skills that can be learned by reading books and taking courses, soft skills such as work ethics, responsibility, courtesy, integrity, a positive and flexible attitude and communication skills are much harder to acquire. Soft skills are also, by nature, much harder to reliably quantify than hard skills and are typically not part of examinations and certificates. Soft skills are nevertheless relevant for productivity and transferable between employers. We rely on an indirect assessment of soft skills and assume that wages reflect productivity which depends on hard and soft skills in addition to other observable factors such as schooling, the characteristics of the training employer and individual characteristics such as age.

We use a cell of homogeneous peers to identify earnings differences in order to keep possible confounding factors such as individual and firm characteristics constant. A cell entails apprentices of one training employer, one occupation, and one cohort. A cell

[10] We also have information on the schooling level of trainees. We include this alternative signal on hard skills in our analysis in order to control for endogeneity induced by correlations between schooling level and grades in the apprenticeship final examinations (Schönberg, 2007).
[11] We include the average grades of the theoretical and practical final exams.
[12] Also compare Arcidiacono et al. (2010) for a discussion on the role of grades and selection into higher education institutions for the revelation of ability.

therefore comprises peers who are subject to the same training instructors, training content, selectivity into training, and collective bargaining rules. Within each cell, we define the difference between earnings of an apprentice and the minimum earnings in the cell measured at the last apprenticeship spell before training completion as additional payment. We argue that the additional payment is voluntary and reflects the opportunity of a training employer to differentiate earnings before the end of apprenticeship training. As we condition on grades in our earnings estimations, the additional payment identifies performance dimensions (latent soft skills) that are not captured in apprenticeship certificates (hard skills).[13]

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 collective bargained wage for their apprentices. Nevertheless, additional voluntary payments for selected apprentices on top of the bargained earnings is a wide-spread 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., 2013). [14]

The relative earnings position of the apprentice within the cell satisfies the hard-to-observe nature of the soft skills indicator because outside employers may observe the absolute individual earnings of an apprentice at the end of apprenticeship. Apprentices can however not credibly signal their relative position in the earnings hierarchy in order to improve their earnings opportunities. Moreover, employers never reveal their complete earnings structure (Waldman, 1990). For outside employers, higher earnings at the end of the training period might be a consequence of high additional payments or of a relatively high earnings level for all apprentices in the training firm.[15] As a consequence, the additional payment is private information available for the researcher. Training employers do not have an incentive to strategically use the additional payment in order to disguise the soft skills of apprentices.

[13] Unfortunately, we do not have information on subjective performance ratings.

[14] Appendix A provides empirical evidence and arguments for our interpretation of voluntary payments as indicators for skills.

[15] 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 in the regressions.

This characteristic of our employer generated information for soft skills is in contrast to publicly observable productivity indicators such as promotions that may be biased by strategic considerations (DeVaro and Waldman, 2012; Waldman, 2016; Waldman and Zax, 2016; Bognanno and Melero, 2016).

Both skill measures (additional payment and grades 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 the creation of the information and the earnings measure is not an issue. The training employer determines the additional payment before the grades in the final exams are known. The grades 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.

In order to be able to compare the impact of additional payments and grades 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 firm is the same as that of the outside employer. In other words: hard and soft skills have the same value in all entry jobs for apprenticeship completers. This assumption is less severe in our setting than in most other labour market settings 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. In addition, apprenticeship completers do not differ with respect to experience and they all recently obtained the same skills guaranteed by the training curricula and final exam certificates.

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). On the training-firm level, we pursue two strategies: we either control for the most important

employer characteristics, firm size and median earnings, or we include a cell fixed effect. The cell fixed effect additionally controls for training-firm unobservable variables such as training instructors, initial selection criteria into apprenticeship 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 public information for outside employers about the higher average productivity of these apprentices leading to higher entry wages for those who switch to another employer (Wagner and Zwick, 2012).

4 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 the basis of initial of name and surname, birthday, gender and the start of apprenticeship.

From exam records, we use the final exam grades, training occupation, and school-leaving qualification. Training occupation and school-leaving qualification are reported in more detail than in the Social Security Records. The occupation code corresponds to the training occupation named in the apprenticeship certificate and clearly distinguishes three and three and a half year apprenticeship occupations. From the Social Security Records, we use earnings, nationality, gender, training firm size, and the establishment identifier. Particularly the earnings information in the Social Security Records 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 chamber data.[16] We restrict our data to those apprentices whose apprenticeship spell in the Social Security Record ends close to the published chamber exam day (compare Table 1 for a detailed data selection procedure). We only consider completers in full-time employment in the first job after apprenticeship, with earnings above the marginal income threshold and with earnings less than the upper Social Security earnings threshold. We further drop apprentices who are older than 30 years and apprentices who earn less than the lowest or more than the highest percentile in their occupation during the apprenticeship training because these seems to be rare misreports.[17] Then, we identify peers of apprentices who learn in the same training establishment, occupation and cohort (our cell definition) and calculate the additional payment within cells. [18] Since calculating the additional payment is only possible in cells with two or more individuals, we drop individuals in cells with one apprenticeship completer only.

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 graduation.

5 Findings

Descriptive statistics

Table 2 summarises our variable definitions, means and standard deviations. The share of employer movers immediately after apprenticeship training is around 40 per cent. For employer movers and stayers, Table 3 separately displays the descriptive statistics. Movers have lower grades in their final apprenticeship exams than stayers, i.e. they have lower hard skills.[19] 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

[16] The identifier for successful graduation is a big advantage of the chamber data in comparison to the Social Security Records that entail only the status as apprentice. Hence, relying on the Social Security Data would require further assumptions to distinguish a successful apprenticeship graduate from drop-outs.

[17] Apprentices can legally earn less than the marginal income threshold. Since the Social Security Data entail a few earnings misreports, we apply the common rule to drop these employees in order to purge the data.

[18] Only 15% per cent of the cells with two or more apprentices, comprising 711 apprentices, have no earnings variation.

[19] In Germany, 1 is the best mark (with distinction) and 4 the lowest (just passed). For better interpretation, we multiply the grades by minus one. Hence, a higher number represents higher hard skills.

more likely to graduate from a lower secondary school track (*Hauptschule*), or without a school certificate.

Similarly, movers have lower soft skills because they receive a lower additional payment at the end of apprenticeship training. The raw earnings difference is substantial with around 0.14 log points accounting for somewhat less than 10 per cent of total earnings. The absolute amount of the average additional payments is small however with around 2 Euros a day (around 60 Euros a month). Movers receive on average a 28 cents a day (8.40 Euros a month) lower additional payment than stayers. This small absolute number supports our argument that outside firms can usually not observe the additional payment. Finally, stayers receive significantly higher earnings in their first job after apprenticeship than movers. The raw difference is again substantially with around 0.2 log points.

Soft and hard skills are correlated at a significance level below one per cent and Pearson's correlation coefficient is 0.17. The correlation is somewhat stronger for stayers with 0.18 than for movers with 0.14. This means that both skill indicators measure different skill dimensions and soft skills cannot be easily predicted from hard skills.

Finally, movers are significantly older and come from smaller training firms that pay lower earnings. We find no differences between movers and stayers 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 a large number of cell fixed effects and can interpret the coefficients as marginal effects. In addition, this regression produces usually similar marginal effects as a Probit or Logit regression if the dependent variable has an unconditioned probability around 50 per cent.

We find that the soft and hard skills 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 account for several individual and training establishment characteristics. An increase in the hard skills mark by one at the final apprenticeship exam reduces the probability to leave the training employer by 5.3 percentage points and a one Euro higher additional payment a day reduces the probability to leave by 8.4 percentage points.[20]

Model 3 includes cell fixed effects instead of establishment-level variables. This is our preferred estimation 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 grades in comparison to their peers. The point estimates of additional payment and grades increase in comparison to models 1 and 2. A one Euro higher additional payment reduces the probability to leave by 14.2 percentage points and a better mark by one increase the probability to leave by 9.8 percentage points. This reduction in the probability to leave is substantial given the unconditioned probability to leave of 40 per cent. Finally, model 4 in Table 4 includes individuals in all cells even if all completers leave or stay in the cell. The coefficients are qualitatively comparable for both productivity dimensions.

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 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 completers. Finally, 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 a negatively selected group with respect to the information on both skill dimensions.[21] Our model interprets the positive selection of stayers as a consequence of information advantages of training firms.

[20] Since the standard deviations of both variables are quite similar, we get qualitatively the same relevance for grades and bonus when we apply a reduction of one standard deviation.

[21] Acemoglu and Pischke (1998) also conclude that there is adverse selection for quitting apprenticeship graduates in Germany based on the observation that military quitters obtain comparably higher entry wages than apprenticeship graduates who voluntary quit and those who have been laid-off.

Asymmetric information on soft skills?

We test whether training firms have an information advantage on soft skills by comparing the influence of soft and hard skills on entry earnings of stayers and movers. Table 5 displays augmented Mincer entry earnings regressions for apprenticeship completers separately for stayers and movers. The first three models show regressions for stayers and the latter three models show the regressions for movers. 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. Including cell fixed effects improves the comparability between the peers in the sample at the price that we can only include cells with at least two stayers or movers. Hence, the sample is more biased against larger firms.

For stayers, soft skills have a significantly positive impact on entry earnings for apprenticeship completers. A one Euro higher additional payment a day during apprenticeship increases the entry earnings by around 1.1 per cent. For movers, soft skills have no significant impact on entry earnings. On the contrary, hard skills strongly influence entry earnings for both groups. A better mark by one unit increases the first full-time earnings of movers by 4.3 per cent and of stayers by 1.9 per cent. The regressions controlling for cell fixed effects confirm the results.

All control variables show the expected signs. Females have lower entry earnings than males. The school-leaving certificate is significantly positively associated with entry earnings for both groups but with stronger effects for movers than for stayers. Hence, schooling levels can also be used as a public signal on hard skills. Furthermore, stayers in high-paying training firms (measured as the mean earnings-level of apprenticeship completers in the establishment) also earn more in their first spell as skilled employee. More interestingly, apprentices leaving a high-paying training firm, also receive higher earnings in the first job after apprenticeship. This may reflect the fact that the new employer wage offer accounts for a superior initial selection by high-paying training employers. Staying German citizens receive lower entry earnings than staying foreign apprenticeship completers. Age has a positive impact on entry earnings for leavers only.

Robustness checks

It is important that we include both skill dimensions into the earnings equation in order to obtain the additional effect for each skill dimension on earnings. This interpretation however rests on the assumption that both skill dimensions are not too strongly correlated with each other and therefore measure different productivity dimensions. If both indicators are highly correlated, the inclusion of the second dimension would reduce the impact of the other in the entry earnings equation (Farber and Gibbons, 1996; Altonji and Pierret, 2001; DeVaro and Waldman, 2012). The reduction in the absolute size of the coefficient should be especially strong for stayers because the training employer can assess soft skills. We check the magnitude of the coefficients of each skill dimension when we exclude the other dimension. Table 6 shows the regressions reporting the key variables. The regressions control for the same variables as models 2 and 5 in Table 5. The magnitude of the coefficients of hard and soft skills remains remarkably stable when we eliminate the other skill dimension (Table 6 in comparison to Table 5). This is an additional indicator that both variables measure different dimensions of productivity (also see Heckman et al., 2006 for comparable evidence).

We then check, whether our results are robust when we reduce the sample of employer movers to those who are very likely to leave on exogenous reasons instead of being fired (compare the discussion in Gibbons and Katz, 1991). We follow the strategy by von Wachter and Bender (2006) who argue training employers that experience a severe short-term demand shock usually do not keep as many apprenticeship completers as usually. These firms reduce the retention rate of apprenticeship completers dramatically in one year but have comparatively high retention rates in all other years. We identify employer movers affected by an exogeneous labour demand shock using two retention rate reduction thresholds: first, a reduction by more than 50 per cent in one single year (model 1 in Table 7) and second, a 20 per cent 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 very similar point estimates for the private information and the public signal on entry earnings. Soft skills do not have a significant impact on entry earnings in both sub-samples. Hard skills are only significant in the larger sample (model 2). The coefficients in the smaller sample remain quite similar in size but 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.

As a final robustness check, we drop the cells with the highest variance (1% and 5%) and we also z-standardised the cells in order to reduce the influence of outliers on our results. The results (not shown here) are always qualitatively the same.

6 Discussion

The paper argues that apprenticeship training firms can positively select apprenticeship completers because they have an information advantage on soft skills. As a result, visible hard skills have a stronger influence on entry earnings for movers than for stayers. Grades therefore have in the words of DeVaro and Waldman (2012) the characteristics of a "publicly observable 'announcement' of the worker´s ability". On the contrary, soft skills are not visible and significantly increase the earnings of stayers but not of movers. This pattern fits perfectly to our interpretation of additional payments as private information. The apprenticeship certificate reveals hard skills but outside employers have no instrument to detect soft skills. Soft skills are therefore a possible channel for training firms to create monopsony power after apprenticeship training.

The relevant question is whether information advantage about soft skills is the only explanation or whether alternative arguments could explain our empirical findings. The literature of employer-sponsored training under imperfect labour market conditions discusses several additional 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 in principle explain our results, we argue below that these factors 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 from own experience (Mohrenweiser, 2016). Although our theoretical model and our estimation

strategy would allow for symmetric information on hard skills, this hypothesis is not supported.

Specificity of training contents can be created in the case of apprenticeship training, by voluntarily offering additional training beyond the contents required by the training curriculum. Apprenticeship training firms in Germany indeed train additional skills. Specificity in additional training content is in principle consistent with the pattern of mobility and entry wages but needs to fulfil two assumptions. First, training firms deliberately select only the best apprentices for additional training and second, the additional training is specific for example to machines or organisational routines in the training firm. We can only speculate about the plausibility of the first assumption, but we can use external data to investigate the second. Appendix Table C1 shows that 53 per cent of metal-working firms do 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 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 per cent of training firms state that they train additional skills not named before. This skills category could include firm-specific skills but also general skills not named in the previous categories. The category is also much smaller than the category “additional soft skills” (36 per cent) which we intend to measure by additional payments. Hence, there remains only a very small fraction of firms that 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. Moreover, Mueller and Schweri (2015) conclude that general and occupation-specific human capital are the most important components of apprenticeship training. Skills are therefore highly transferable within an occupational field based on their analysis of entry wages after apprenticeship training.

Regional monopsonies are also frequently used to explain 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 indeed stay in the same labour market regions when they move to another employer (Winkelmann, 1996). Therefore, regional employers are relevant alternatives for the majority of young skilled

employees who are willing to change the employer. The Saarland is typically classified as one labour market region (Kosfeld and Werner, 2012). We therefore cannot compare the relative importance of hard and soft skills in the selectivity of movers in regions with differences in labour market competition on the basis of our data set.

Hence, we cannot be certain that the positive selection of stayers in apprenticeship training firms is only a consequence of information advantages on soft skills of apprenticeship completers but additional evidence shows that the information advantage is the most plausible explanation for the findings.

7 Conclusions

This paper shows that German employers that invest in general skills of apprentices can keep a positive selection of apprenticeship completers with respect to hard and soft skills. Hard skills can easily be signalled to outside employers on the basis of graded final exams issued by chambers of trade or chambers of commerce. Soft skills however cannot credibly be signalled to outside firms and training firms can gain an information advantage about soft skills.

This paper provides a new explanation of information advantages of training firms. The information advantage about soft skills allows training firms to earn a return on their investment in transferable human capital. They can pay highly productive apprenticeship completers earnings below their productivity. Apprenticeship completers cannot signal their high soft skills to outside employers and therefore do not obtain better outside options. This paper is the first direct empirical evidence for a frequently stressed but never proven assumption: German firms are willing to invest in certified, visible and transferable human capital of their apprentices because they obtain an information advantage (Acemoglu and Pischke, 1998). This information advantage gives training firms the opportunity to retain a positively selected group of apprenticeship completers. Hence, the paper clarifies previous discussions about the effect of certification on firms' and individuals' training investments. As long as the certificate or the final exam covers not all relevant productivity 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 final exam issued by external and independent institutions. Employer-sponsored training leading to a widely recognised certificate is widespread, however. Examples are IT skill certificates, reimbursement programmes for MBA, Master or Bachelor studies in academic institutions, language certificates or mandatory health and safety certificates (Cappelli, 2004; Benson et al., 2004; Arcidiacono et al., 2010; Manchester, 2012; Benson, 2013, Sauermann, 2015). All these certificates have in common that they credibly reveal individual hard skills about IT, language or health and safety procedures shortly before they can be used as signals on the labour market. They however do not reveal soft skills. Hence, all these training programmes entail the possibility that employers can sustain an information advantage about soft skills.

Our results also demonstrate that training firms can positively select employer movers after training. To the best of our knowledge, our paper presents the first direct empirical identification of adverse selection processes in the labour market. Staying completers are positively selected on hard and soft skills, because soft skills are positively correlated with hard skills. Hence, our paper complements studies that indirectly identify adverse selection (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 soft skills 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 (Waldman, 1984; DeVaro and Waldman, 2012; Waldman and Zax, 2016; Bognanno and Melero, 2016). As a consequence, our information on soft skills 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), Kim and Usui (2014), and Kahn and Lange (2014) that also generate indicators for information asymmetries on skills that are valuable for the labour market exploiting earnings differences.

Finally, our paper adds to the literature on information asymmetries on the labour market. Previous studies distinguished between symmetric and asymmetric information on ability (Schönberg, 2007; Pinkston, 2009; Kahn, 2013). Waldman (2016) introduced a distinction between "academic" and "productive" abilities in a theoretical signalling model. Our paper provides a related empirical example that it is indeed important to distinguish between different skills dimensions when analysing information asymmetries between training and outside employer.

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 grades in apprenticeship exams induces apprentices to work harder for their final exam success. 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 or firms with works councils as additional reasons for the positive selection of apprenticeship completers who stay with their training firms and for training incentives. Finally, measuring soft-skills or particular dimensions of soft-skills more accurately and combine those measures with final exam results to assess signalled hard-skills, 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.

Akerlof G (1984). Gift Exchange and Efficiency-Wage Theory: Four Views, American Economic Review 74(2): 79-83.

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.

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.

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.

Bispinck R, Dorsch-Schweizer M and Kirsch J (2002). Tarifliche Ausbildungsförderung begrenzt erfolgreich – eine empirische Wirkungsanalyse, *WSI Mitteilungen* 04/2002, 213-219.

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.

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.

Cottini E, Kato T and Westergaard-Nielsen N (2011). Adverse workplace conditions, high-involvement work practices and labor turnover: Evidence from Danish linked employer–employee data, Labour Economics 18(6): 872-880.

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.

Dustmann C and Schönberg U (2009). Training and union wages, Review of Economics and Statistics 91(2): 363-376.

Farber H and Gibbons R (1996). Learning and Wage Dynamics. Quarterly Journal of Economics 111(4): 1007-1047.

Fitzenberger B, Licklederer 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.

Hu L and Taber C (2011). Displacement, Asymmetric Information, and Heterogeneous Human Capital, Journal of Labor Economics 29(1): 113-152.

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.

Kim S and Usui E (2014). Employer Learning, Job Mobility, and Wage Dynamics, working paper, University of Washington.

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.

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 (2012). Which firms train disadvantaged youth? Empirical Research in Vocational Education and Training 4(2): 115-130.

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, Zwick T and Backes-Gellner U (2013). Poaching and Firm Sponsored Training: First Clean Evidence, ZEW Discussion Paper 13-037, Mannheim.

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.

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, Hurrel, SA and Cullen AM (2012). Soft skills and employ-ability: 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.

Ryan P (2011). Apprenticeship: Between Theory and Practice, School and Workplace, Economics of Education Working Paper No. 64, Zurich.

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.

Stockinger B, Zwick T (2017). Apprentices Poaching in Regional Labor Markets, ZEW Discussion Paper 17-013, Mannheim.

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.

Wagner R and Zwick T (2012). How Acid are Lemons? Adverse Selection and Signalling for Skilled Labour Market Entrants, ZEW Discussion Paper 12-046, Mannheim.

Waldman M (1984). Job Assignments, Signalling, and Efficiency, RAND Journal of Economics 15(2): 255-267.

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)	5,792

Source: Ausbildungspanel Saarland 1998-2005.

Table 2: Definitions of variables and descriptive statistics

Variable	Definition (mean; sd)
Hard skills (public signal)	Grade in the final apprenticeship exam multiplied by minus 1 (-2.90; 0.74)
Soft skills (private signal)	Log daily earnings deviation from establishment/ occupation/ graduation year cell minimum plus 1 in € (0.58; 0.77)
Log earnings at first full-time employment	Log daily earnings at first full-time employment as skilled worker in € (4.03; 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 graduation (0.41; 0.49)
Female	Dummy variable equals 1, if apprentice is female (0.32; 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 (21.45; 2.01)
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 € (23.12; 6.85)
Number of employees (training employer)	Number of employees in the training employer (591; 1275)

N = 5,792; Source: Ausbildungspanel Saarland 1998-2005.

Table 3: Description of differences between stayers and movers

Variable	Movers	Stayers	t-test
Hard skills	-3.01	-2.82	5.92
Soft skills	0.500	0.638	4.25
Log earnings at first full-time employment	3.942	4.109	7.91
Female	0.331	0.324	0.25
Lower school track “Hauptschule” or drop-out	0.492	0.357	5.14
Medium school track “Realschule”	0.289	0.345	2.18
Upper school track “Abitur”	0.217	0.296	3.21
Age	21.66	21.30	6.75
German nationality	0.961	0.960	0.07
Establishment earnings level (training firm)	20.74	24.81	42.14
Number of employees (training firm)	292	803	424
Number of observations	2408	3384	

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

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

Variable	(1)	(2)	(3)	(4)
Soft skills (private information)	-0.093***	-0.084***	-0.142***	-0.051***
	(6.02)	(5.34)	(5.20)	(5.10)
Hard skills (public signal)	-0.043***	-0.053***	-0.098***	-0.064***
	(2.77)	(3.21)	(3.12)	(6.30)
Female		0.003	-0.023	
		(0.09)	(0.48)	
School certificate “Realschule”		-0.076***	-0.140***	
		(2.64)	(2.72)	
School certificate “Abitur”		-0.012	-0.102	
		(0.31)	(1.41)	
Age		0.024***	0.039***	
		(4.20)	(3.91)	
German		0.168***	0.283***	
		(3.20)	(3.09)	
Number of employees		-0.000***		
		(6.30)		
Establishment earnings level		-0.002		
		(0.96)		
Sector, Year, Occupation	Yes	Yes	No	Yes
Cell Fixed Effects	No	No	Yes	No
Pseudo R-sqr	0.08	0.11	0.26	0.22
Number of Observations	2320	2320	2320	5792

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; model (1)-(3) restricted to cells with movers and stayers, model (4) includes 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

	Stayers			Movers		
	(1)	(2)	(3)	(4)	(5)	(6)
Soft skills (private information)	0.027***	0.011**	0.012*	0.018*	0.008	0.016
	(4.84)	(2.12)	(1.72)	(1.91)	(0.82)	(0.98)
Hard skills (public signal)	0.025***	0.019***	0.007	0.059***	0.043***	0.062***
	(4.66)	(3.52)	(0.99)	(5.51)	(3.93)	(2.94)
Female		-0.017*	-0.019*		-0.030	-0.018
		(1.83)	(1.87)		(1.44)	(0.59)
School certificate “Realschule”		0.016*	-0.001		0.034**	0.013
		(1.95)	(0.17)		(2.23)	(0.52)
School certificate “Abitur”		-0.005	-0.008		0.042	0.049
		(0.40)	(0.53)		(1.60)	(1.08)
Age		0.003	0.004		0.008**	-0.003
		(1.66)	(1.43)		(2.24)	(0.49)
German		-0.036**	-0.008		0.012	-0.013
		(2.16)	(0.31)		(0.35)	(0.19)
Number of employees in training establishment		0.000			0.000	
		(0.68)			(0.91)	
Training establishment earnings level		0.014***			0.008***	
		(12.44)			(4.83)	
Sector, year, occupation dummies	Yes	Yes	No	Yes	Yes	No
Cell Fixed Effects	No	No	Yes	No	No	Yes
R^2	0.60	0.65	0.88	0.35	0.36	0.58
Number of observations	3384	3384	2846	2408	2408	1799

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; models (1), (2), (4) and (5) include all observations, model (3) includes only cells with at least two stayers, model (6) includes only cells with at least two movers *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$; Source: Ausbildungspanel Saarland 1998-2005.

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

	Stayers				Movers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Soft skills (private information)	0.011***	0.012*			0.008	0.018		
	(2.12)	(1.72)			(0.82)	(1.10)		
Hard skills (public signal)			0.019***	0.007			0.043***	0.062***
			(3.53)	(1.01)			(3.93)	(2.98)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Establishment controls	Yes	No	Yes	No	Yes	No	Yes	Non
Sector, year, occupation dummies	Yes	No	Yes	No	Yes	No	Yes	No
Cell Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.65	0.88	0.65	0.88	0.35	0.58	0.36	0.58
Number observations	3384	2846	3384	2846	2408	1799	2408	1799

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; models (1), (3), (5) and (7) include all observations, models (2) and (4) include only cells with at least two stayers, models (6) and (8) include only cells with at least two movers; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$; Source: Ausbildungspanel Saarland 1998-2005.

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

Variable	(1)	(2)
Soft skills (private information)	0.009	0.020
	(0.12)	(0.49)
Hard skills (public signal)	0.112	0.089*
	(1.18)	(1.74)
Individual controls	Yes	Yes
Cell Fixed Effects	No	No
R^2	0.56	0.68
Number of observations	156	511

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 percent below long-term average, model (3) restricted to cells with retention rate 20 percent 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 of relative earnings rank in an establishment/ occupation/ cohort cell is an indicator for apprentice skills.

The earnings rank of apprentices at the end of training has predictive power. The earnings rank predicts first full-time earnings of stayers (Mohrenweiser et al., 2013), i.e. additional payments during apprenticeship training are related to entry earnings training firms are willing to pay for apprenticeship completers they retain. 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., 2013). When skilled entry earnings reflect the market value (productivity) 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, 1986). 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., 2013).

We can interpret the earnings differences within a cell as relative individual productivity differences because apprentices learn the same skills and occupation-specific training curricula determine tasks that apprentices should perform and learn during each stage of apprenticeship. Therefore, two apprentices in the same occupation do not perform different tasks to obtain different earnings. Moreover, apprenticeship completers in the same training occupation in one firm are practically identical in terms of observable variables such as age,

education, start of apprenticeship and prior working experience. Nevertheless, many employers have explicit financial bonus rules for good grades at vocational school or good performance at work (Ryan, 2011; Ryan et al., 2013; Backes-Gellner and Oswald, 2014). Ryan et al. (2013) present evidence for individual and group-related performance pay for apprentices in 13 out of 18 analysed engineering and retailing firms in Germany.

Our measure filters out additional group payments frequently implemented on the establishment level and instead captures only the individual differences between apprentices in homogeneous cells. Group payments for all apprentices within a cell do not add variation that is required to calculate the difference between the individual earnings and the minimum earnings within a cell.

Unions, some works councils and the management of some training firms may regard explicit performance pay for apprentices as divisive or as a source of dysfunctional incentives to produce rather than to learn during training.[22] We therefore collected information about the internal validity of our interpretation of earnings differences in interviews with several personnel managers in training firms. Personnel managers confirm that all apprentices in a cohort receive the same base salary, according to collective agreements. However, many firms voluntarily pay for extracurricular activities by apprentices that are regulated in collective agreements. For example, many firms send their apprentices abroad for some weeks in the final training year. Firms state that apprentices who are sent abroad are a positive selection, that working abroad therefore is an implicit reward for good performance which is accompanied by additional payments. Other examples mentioned are payments for work during the weekend or overtime hours.[23] Again, the payment for these activities is 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 variance in the cells. Even these payments may be correlated with work-related performance if training firms are willing to grant these benefits for extraordinarily able apprentices only.

[22] Unfortunately, we do not have information on industrial relations in our data and therefore cannot check whether voluntary payments differ for example between establishments with and without collective agreements/works councils.

[23] The earnings definition in the LIAB data entails full-time earnings for apprentices. Therefore, bonus payments for overtime, weekend or shift work payment is included in the wage sum.

Appendix B: Mathematical derivation of model

We use the following notation: The individual soft skill measure "additional payment" *b* can take two expressions, high ($b = H$) and low ($b = L$). The individual hard skill measure "grade" *g* can also take two expressions, excellent ($g = E$) and average ($g = A$)[24]. 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 complements[25]. We therefore assume that an excellent grade adds a constant *c* to productivity in comparison to an average grade. Finally, we denote with *p* the probability that a training participant received an additional payment. We assume that soft and hard skills are positively related, or in other words training participants with excellent grades received a high additional payment with a higher probability than training participants with an average grade ($p^A < p^E$).[26] Table A1 gives an overview of our model.

Table B1: Specification of Productivity *y*

	Average grade *G*	Excellent grade *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 grades. Grades measure hard skills but do not reveal all dimensions of productivity. The missing dimension of productivity, soft skills, is known to training employers and to the training participant. Outside employers know the general probability of employees with average and excellent grades to receive an additional payment *p* and they observe a common noisy signal

[24] We do not use "high" and "low" for the differentiation of grades to ease reading.

[25] Schönberg (2007) proofs in her appendix A6 that the results hold if hard and soft skills are complements. As a consequence, the adverse selection should be stronger for apprenticeship graduates with high hard skills, a result also derived by Acemoglu and Pischke (1998). In addition, the difference between the impact of soft skills on entry wage offers of training and outside employers should be stronger for apprenticeship graduates with high soft skills than for apprenticeship graduates with low soft skills.

[26] The additional payment reflects the soft skill assessment of the training employer. We have to assume that the outside employer would assess soft skill 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.

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 grades.[27]

In our simple model, the potential signal *s* about soft skills takes only two values, good (*s = G*) and bad (*s = B*). The outside employer uses the signal in order to infer the soft skills of 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 grade is excellent or average:

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

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

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, $\Pr(L|k,B) = \Pr(L|k,G) = p^k$.

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

[27] 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 soft skills, the training firm therefore does not necessarily perform a job interview by itself in order to reveal the signal.

independent of the additional payment and grade (and the signal *s*). We assume that the disutility shock in the training firm is drawn from a distribution with the cumulative distribution function *G* with an associated probability function *g* that is bounded by $\left[\underline{\theta},\overline{\theta}\right]$ with $\overline{\theta} > 0$. The cumulative distribution function 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.

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 $w < v$[28]. 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 grades *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 probability that the training participant stays therefore can be written as Pr(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 grades *g* and signal *s*. The training employer therefore maximises profits by setting earnings according to:

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

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:

[28] Note that training participants with θ < 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.

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

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:

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

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 soft 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 grades, 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:

$v_L^k - w_L^k = v_H^k - w_H^k$. (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-bonus 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:

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

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

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

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

$$G\left(v^B - w_L^B\right) - G\left(v^G - w_H^G\right) > 0 \Leftrightarrow w_H^G - w_L^B > v^G - v^B, \quad (8)$$
$$G\left(v^G - w_L^G\right) - G\left(v^B - w_H^B\right) > 0 \Leftrightarrow w_L^G - w_H^B < v^G - v^B. \quad (9)$$

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-bonus 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:

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

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 to hold 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.[29]

We can turn now to the main feature of the model – the influence of additional payments and grades 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

[29] 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 payment as a measure of performance ratings.

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 *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 *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 grades 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 grades 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 grade not only reflect the productivity-enhancing effect of having a good mark, but also the fact that training participants with good grades have a higher probability to have a high additional payment p^E than training participants with low grades 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 grades equals:

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

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

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

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 grade has the same impact on productivity for low-productivity and high-productivity training participants, we find that:

$$\frac{\mathrm{d}v^{E,s}}{\mathrm{d}s} = \frac{\mathrm{d}w^{E,s}}{\mathrm{d}s} = 1.$$

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

$$\frac{\mathrm{d}w_i^{E,s}}{\mathrm{d}(1-p^E)} = \frac{\left[g_i^{E,s2} + \left(1-G_i^{E,s}\right) g_i^{E,s'}\right] / g_i^{E,s2}}{1+\left\{\left[g_i^{E,s2} + \left(1-G_i^{E,s}\right) g_i^{E,s'}\right] / g_i^{E,s2}\right\}} \frac{\mathrm{d}v^{E,s}}{\mathrm{d}(1-p^E)}.$$

Log-concavity of G gives $\frac{\mathrm{d}w_i^{E,s}}{\mathrm{d}(1-p^E)} < \frac{\mathrm{d}v^{E,s}}{\mathrm{d}(1-p^E)}$. Therefore an excellent grade 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.

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

This appendix shows the extent and content of additional training 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 his member firms about apprenticeship related questions ones 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 among the most expensive training investments for firms (Mohrenweiser and Zwick, 2009; Schönfeld et al, 2010) and these occupations serve usually 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 skill 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 Source: *Ausbildungspanel Südwestmetall*, 2011.