Discussion Paper No. 15-057

Work-Related Ability as Source of Information Advantages of Training Employers

Jens Mohrenweiser, Gaby Wydra-Sommaggio, and Thomas Zwick

ZEW
Zentrum für Europäische Wirtschaftsforschung GmbH
Centre for European Economic Research

Discussion Paper No. 15-057

Work-Related Ability as Source of Information Advantages of Training Employers

Jens Mohrenweiser, Gaby Wydra-Sommaggio, and Thomas Zwick

Work-Related Ability as Source of Information Advantages of Training Employers*

Jens Mohrenweiser (Bournemouth University)

Gaby Wydra-Sommaggio (IAB Regional Saarbrücken)

Thomas Zwick (University of Würzburg and ZEW Mannheim)

This version August 2015

ABSTRACT:

This paper addresses the puzzle how employers that invest in general human capital can gain an information advantage with respect to the ability of their employees when training is certified by credible external institutions. We apply an established model from the employer-learning literature and distinguish between two ability dimensions: cognitive and work-related ability. We apply this model to the German apprenticeship system and show that cognitive ability certified by external institutions at that the end of apprenticeship training can be signalled to outside employers. Apprenticeship graduates however cannot signal their work-related ability – measured by a small voluntary bonus paid by the training employer – to the outside market. We therefore show that the information advantage on work-related ability explains that training employers can positively select the apprentices they retain. As a consequence, this information advantage induces employers to invest in certified and transferable human capital.

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

Key words: training, employer learning, 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, 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 employee ability (Chang and Wang, 1995; Acemoglu and Pischke, 1998; Autor, 2001) or training content (Katz and Ziderman, 1990). The information advantage drives a wedge between workers' productivity and their `marketable' or outside wage. This wedge allows training firms to get a return on investment and nevertheless keep their trained employees by paying them a wage below productivity but above market wage[1].

Indeed, numerous empirical studies report employer-sponsored training in transferable skills such as college, 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), and apprenticeship training (Mohrenweiser and Zwick, 2009; Schönfeld et al., 2010; Ryan et al., 2013; Muehlemann et al., 2013; Kriechel et al., 2014). Those training courses usually end with a certificate issued by an external institution (Acemoglu and Pischke, 2000). It is however well known that particularly marked certificates issued by external institutions allow outside firms to infer individual ability (Arcidiacono et al., 2010). This kind of company-sponsored training therefore leads to a puzzle: how can training firms gain information advantages if training courses are credibly certified by external institutions?[2]

This paper provides a novel solution to solve this puzzle. Following recent developments in the employer learning literature[3], we distinguish between cognitive and work-related ability of employees (Waldman, 2014). Cognitive ability is measured in standard exams that allow training and outside firms to assess relative performance in comparison to other training participants. On the contrary, work-related ability is not measured and marked in traditional training certificates. Work-related ability comprises soft-skills, ability to work in teams or

[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).

[2] Katz and Ziderman (1990) explicitly analyse certification and conclude that certification overcomes information asymmetries about training content and reduces firms' incentive to invest in training.

[3] The employer learning literature assumes that employers are initially not able to fully assess their employees´ ability (and therefore productivity) but learn about it during the first years of employment.

behaviour towards co-workers, line managers and subordinates. Individuals learn about their cognitive ability already during school and are aware about it before they start an employment but they do not know their work-related ability before they start to work. This means that employer and employee learn simultaneously during employment about work-related ability (Waldman, 1984; deVaro and Waldman, 2012; Waldman, 2014; Waldman and Zax, 2014; Bognanno and Melero, 2015).

Hence, we argue that training firms gain an information advantage about work-related ability (private information) but not about cognitive ability (public signal) if training is certified by external institutions. In other words, information asymmetry on work-related ability can generate the necessary information advantage that creates the wedge between training firm and outside firm wages necessary to get a return on training investment.

To analyse the question of information advantages of training firms after certified general training, we first introduce an analytical framework based on an established employer-learning model of Schönberg (2007). The model nests symmetric and asymmetric learning. In the symmetric learning case, training and outside firms simultaneously learn about cognitive as well as work-related ability. In the asymmetric learning case, the training firm learns about individual work-related ability but not the outside firm[4]. This model allows us to derive hypotheses that identify information asymmetries for both ability types between training and outside firms based on wage offers and job mobility after training.

Second, we empirically test these hypotheses using particularities of the German apprenticeship training system. The strongly regulated apprenticeship system provides training for around two thirds of each labour market cohort. Apprenticeship training requires investments by employers in transferable human capital. It ends with certificates issued by independent institutions. We use a unique database that contains marks of the final apprenticeship certificate amongst other important information from administrative sources for the entire population of several cohorts of apprenticeship graduates in one

[4] Traditionally, studies on employer learning concentrate on symmetric learning about cognitive ability, for example Farber and Gibbons (1996), Altonji and Pierret (2001), Lange (2007), Mansur (2012), Kahn (2013), Kahn and Lange (2014), and Light and McGee (2015). However, Schönberg (2007), Pinkston (2009), Hu and Taber (2011), and Kim and Usui (2014) present models that nest symmetric and asymmetric learning about cognitive ability.

German federal state. The final marks serve as information about cognitive ability. We measure work-related ability with a bonus voluntarily paid by training employers on top of apprentices' base salary. Comparing bonus payments between apprenticeship graduates in the same training firm, in the same year, and in the same occupation allows us to detect differences of work-related ability between a very homogeneous group of peers.

We show that work-related ability indeed induces an information advantage of training firms. Bonus payments and exam marks are correlated but measure different dimensions of ability. They therefore both have distinct explanatory power in our entry earnings equations. Entry earnings of stayers after training are correlated with the bonus and final marks but entry earnings of movers after training are only correlated with marks but not with the bonus. This means that work-related ability identified by the bonus payments cannot be signalled to outside firms. As a consequence, training firms are able to keep a positively selected group of graduates.

This paper contributes also to a number of further open topics. First, it presents a direct assessment of adverse selection – we show that training employers are able to positively select apprenticeship graduates. Since work-related ability is positively related to cognitive ability, staying graduates are positively selected on both ability dimensions. 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) and studies using cognitive ability indicators as determinants of moving to another employer (Schönberg, 2007).

Second, our paper adds to recent developments in the employer learning literature. Previous studies analysed either employer learning on cognitive or work-related abilities. Waldman (2014) for the first time introduces a distinction between "academic" and "productive" abilities in one theoretical signalling model. Our paper provides an empirical example that it is indeed important to distinguish between different ability dimensions when analysing information asymmetries between incumbent and outside employer.

Third, our paper proposes a new measure for work-related ability in order to identify information advantages of incumbent employers: the relative wage position of an apprenticeship graduate in relation to peers in the same training firm, occupation, and cohort. A bonus is hard to observe for outside employers in comparison to the relatively easy to observe promotions frequently used in previous studies to identify work-related ability (Waldman, 1984; deVaro and Waldman, 2012; Waldman and Zax, 2014; Bognanno and Melero, 2015). As a consequence, our information on work-related ability 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) who also generate indicators for information asymmetries on the labour market derived from earnings trajectories between cohorts.

The paper is structured as follows. The next section presents a model on the consequences of symmetric and asymmetric employer learning on cognitive and work-related ability on the selection of employer movers and the determinants of earnings of apprenticeship graduates. The third section discusses our empirical strategy and the fourth section presents our data and variables. The fifth section shows and interprets the empirical results and the last section concludes.

2 Theoretical framework

We apply the framework of public and private information identified in the employer-learning literature to the literature of company-sponsored training in general human capital. The basic idea of employer learning models is to derive testable hypotheses about determinants of entry earnings for stayers and movers. As we restrict our sample to a homogeneous group of training graduates, we can infer whether training employers learn simultaneously with outside firms about the ability of training graduates (this is the symmetric learning case) or whether training employers learn more about their ability than outside firms (asymmetric learning case). Asymmetric learning enables training firms to use the information advantage to get a return on their training investments by paying graduates less than productivity but more than the market wage.

We apply the employer-learning model of Schönberg (2007) to the German apprenticeship training context. In the appendix, we provide a more formal derivation of the model that complements the intuitive account given in this section. The model consists of two periods. In the first period, apprenticeship training takes place and ends with a certificate.[5] The training employer can screen apprentices 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 a wage offer and the training employer can counter the offer. The apprenticeship graduate stays with the training employer if the wage offer is higher than the wage offer of the outside employer plus an individual disutility shock incurred during training.[6] The disutility shock is not observable by the training employer but it has some prior about the distribution of the disutility shock. Hence, the model generates some voluntary turnover for exogeneous reasons (Acemoglu and Pischke, 1998; Autor, 2001).

The model distinguishes between two productivity types of workers, high and low, in two productivity dimensions – cognitive and work-related ability. The training employer learns about the individual level of both productivity dimensions during the first period. Cognitive ability is defined as the knowledge necessary to perform a certain trade. This ability is certified and marked in extensive exams supervised and marked by external institutions at the end of the training period. Cognitive ability is easy to observe for all market participants (and the researcher) and therefore public on the basis of credible certificates. In the basic version of the model, a high level of cognitive ability adds the same productivity to employees with low work-related ability as to employees with high work-related ability. The probability that the employee has high work-related ability is however higher when the employee has high cognitive ability. The a-priori probabilities of high and low work-related ability for the groups with high and low cognitive ability are known to outsider employers.

Work-related ability is hard-to-observe information from the outside employers´ perspective. The key question is whether the information remains private or whether outside employers can reveal the individual work-related ability level. The revelation of work-related ability by the outside employer needs an additional signal such as a job

[5] In contrast to the model by Acemoglu and Pischke (1998), the amount of training provided is given in our model. In the empirical section, we only compare wages of apprenticeship graduates who have been trained in the same firm and learnt the same occupation. It seems plausible that all apprenticeship graduates in this homogeneous group received the same training input.

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

interview (Lange, 2007) because outside employers do not directly obtain reliable information on it. If this additional signal perfectly transmits the information on work-related ability, outsider and incumbent training firms have the same (correct) information about it and we have completely symmetric employer learning with respect to work-related ability.[7] If the additional signal is only white noise, the outside employer cannot infer anything about the work-related ability of the applicants in addition to the a-priori probability distributions of work-related ability for applicants with high and low cognitive ability levels. In this case, we have purely asymmetric learning.

If outside employers can accurately assess work-related ability, incumbent and outside employers offer the full additional productivity as bonus for high cognitive and for high work-related ability. This means that only the disutility shock allocates movers and stayers. As the disutility shock is equally distributed over ability levels, the training employer cannot retain a positive selection of employees.

If the training employer has superior information on work-related ability, the outside employer is not aware of and therefore not willing to pay for the full additional productivity associated with high work-related ability. For outside employers, the applicant has a low work-related ability level with a certain probability even if he or she has high cognitive ability. The training employer is able to pay fully the higher productivity and has the additional advantage that it can counter the wage offer of the outside employer because it has the last wage offer. In the asymmetric employer learning equilibrium, the training firm can offer a slightly higher wage for graduates with the high work-related ability and can keep them. The training firm can also select those graduates who have to leave after training – either, it does not give them a higher counter-offer or, it does not give them an employment offer after graduation. Hence, outside employers anticipate that incumbent training employers use their informational advantage and consider employer movers as having a higher risk of low work-related ability or are in other words “lemons” (Greenwald, 1986; Gibbons and Katz, 1991). As a consequence, the incumbent employer can pay graduates with high revealed work-related ability less than their productivity but more than the market. This difference between wages and productivity can be used to finance the training investment

[7] This is similar to the case of symmetric learning.

even if some training participants with high work-related ability leave because they experienced a high disutility shock (Acemoglu and Pischke, 1999; Leuven, 2005).

Summing up, our model nests symmetric and asymmetric employer learning and predicts the following: If information is symmetric about both ability dimensions for incumbent and outside employers, work-related and cognitive ability have the same impact on entry wages for those who stay with their employer and those who leave. If incumbent employers have an information advantage about work-related ability, cognitive ability has a stronger impact on the entry wages of movers than of stayers because the outside employer can only rely on this public signal. However, work-related ability has a stronger impact on entry wages of stayers than of movers because the additional signals on this private information are weaker for outside employers. If information on work-related ability is asymmetric, cognitive and work-related ability are higher for stayers than for movers (adverse selection). Both ability levels are equal for stayers and movers, if information is symmetric, however.

3 Empirical Strategy

Institutional Background

We use the German apprenticeship training system as an application for training in general human capital because it develops skills that are visible for outside employers and transferable between firms. Apprenticeship training in Germany traditionally provides the highest education degree for about two thirds of the German workforce and is therefore the backbone for medium skilled vocational training. Apprentices usually start the apprenticeship immediately after school at around age 16. The apprenticeship period for most of them is 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 ability dimensions of apprentices. Apprenticeships are generally considered as training investments for firms (Mohrenweiser and Zwick, 2009; Schönfeld et al., 2010).

Apprenticeships are strongly regulated by the state and subject to 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 and tailor-made for each occupation and describe the 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 in their region 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 graduates in one occupation enter the labour market at the same day. The theoretical part of the skill examination is administered and marked by publicly funded and controlled vocational schools. It is important for our empirical strategy that apprentices receive marked certificates at the end of the training period because marks allow all employers to assess cognitive ability of otherwise observationally equal apprenticeship graduates.

In addition, apprentices are free to choose an employer after the exam. Training costs reimbursement contracts for apprentices are forbidden by law. According to our model, training firms can always outbid an outside wage offer for those apprentices they would like to keep. 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. Outside firms therefore cannot easily verify whether a job applicant had an employment offer from the training firm or not. Promotions during the training period also cannot be used as alternative signal on work-related ability.

Dependent variables: probability to move and entry earnings

We use two dependent variables to assess the extent and existence of asymmetric information on ability: a binary variable of the status as staying or leaving apprenticeship graduate and the entry earnings of apprenticeship graduates. Entry earnings are measured in the first employment spells after graduation. These spells take clearly less than one year[8]. This means that quick employer learning cannot influence entry earnings.[9] Our measure of entry earnings for leavers and stayers also is not biased by the influences of the business cycle (Kahn, 2013). Most apprenticeship graduates 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 – clearly higher – earnings for those apprenticeship graduates who stay with them. Therefore, earnings of leavers and stayers are determined at exactly the same point in time. This is an important requirement for measuring employer learning (Farber and Gibbons, 1996; Pinkston, 2009).

Information on cognitive ability

We measure cognitive ability using the final marks apprenticeship graduates receive at the end of the training period. Their certificate provides easily accessible and unbiased publicly available information[10]. Independent institutions (the chambers of industry and commerce or crafts) and government-run vocational schools administer and conduct the final exams. Outside employers can assess the cognitive ability of each apprentice on basis of the marks in several practical and theoretical subjects. Most employers who hire quitting apprenticeship graduates train themselves (Mohrenweiser, 2015) and they therefore have experience in assessing marks reported in the apprenticeship certificate and other important information such as school marks, the occupation and the selectivity of the training employer (Wagner and Zwick, 2012; Mohrenweiser and Zwick, 2015).[11]

---

[8] The maximum spell length is less than one year. Employers have to report earnings etc. 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] When the entry spell would be longer, it may have been possible that earnings have increased for moving employees who are more productive than a-priori expected. This could confound the interpretation of entry earnings as expected productivity before the employer gets to know the employee, see Lange (2007).

[10] We also have information on the schooling level of trainees. We include this alternative signal on cognitive ability (Schönberg, 2007) in our analysis in order to control for endogeneity induced by correlations between schooling level and marks in the final apprenticeship examinations.

[11] Also compare Arcidiacono et al. (2010) for a discussion on the role of marks and selection into the schooling institution for the revelation of ability.

Information on work-related ability

We measure work-related ability using bonus payments of incumbent training employers on top of base salary. Paying a bonus 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. In collective bargaining contracts, the wages for apprentices in all occupations are specified and therefore employers could just pay the collective bargained wage for their apprentices during their training period. Nevertheless, paying a bonus for selected apprentices on top of the bargained wage is a wide-spread policy of training firms – almost all training firms differentiate their wages for observationally equal employees (Mohrenweiser et al., 2013). Our approach is comparable for example with the approach chosen by Kahn and Lange (2014) who argue that differences in wages between observationally equal employees partly reflect subjective productivity assessments of an employer.[12]

For calculating the bonus, we rely on the definition of a cell of homogeneous peers. A cell entails apprentices in one training employer, one occupation and one graduation cohort. The cell therefore comprises peers who are subject to the same training instructors, training content, selectivity into the training job, and collective bargaining rules. Within each cell, we define the bonus as the difference between earnings of an apprentice and the minimum earnings in the cell measured at the last apprenticeship spell before graduation.[13] The bonus reflects the opportunity of a training employer to differentiate earnings before the end of apprenticeship training. As we condition on marks at the end of apprenticeship training, the influence of the bonus on earnings captures work-related ability such as personality, capability to work in teams or social skills that are not captured in apprenticeship certificates.

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

[13] A number of previous studies exploit the earnings variation between apprentices close to graduation who learn the same occupation in the same establishment and graduation year. Some training firms pay a bonus for example for extra-ordinary performance to reward and reveal superior individual performance to the employees (Backes-Gellner and Oswald, 2014; Ryan, 2011; Ryan et al. 2013). Moreover, training employers frequently pay a bonus for those selected apprentices who are sent to subsidiaries abroad for some weeks, which can be considered as an incentive for good performance (Mohrenweiser et al., 2013). Finally, firms may use the bonus as a cheap possibility to increase the chance to retain positive reciprocal apprentices in the training firm after graduation.

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 relation to their peers in the training firm in order to improve their earnings opportunities on the labour market. 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 a high bonus or of a relatively high earnings level for all apprentices in the training firm.[14] As a consequence, the bonus is a private signal available only for the training employers, the apprentice and the researcher. Training employers do not have an incentive to strategically use the bonus payment in order to disguise the work-related ability of apprentices. This characteristic of our employer generated information for work-related ability is in contrast to publicly observable productivity indicators such as promotions that may be biased by strategic considerations (DeVaro and Waldman, 2012; Waldman, 2014; Waldman and Zax, 2014; Bognanno and Melero, 2015).

Both ability measures (bonus and marks) have the advantage that they have been created close to the actual decision about entry earnings after graduation and not before labour market entry. Therefore, unobservable heterogeneity such as more training for better apprentices between the creation of the information and the earnings measure is no problem here. Moreover, the training employer determines the bonus before the marks in the final exams are known. In addition, marks are not given by the training employer. This means that both measures are determined independently and that (as we will show below) they measure different ability dimensions.

In order to be able to compare the impact of the bonus and marks on entry earnings of movers and stayers, we have to assume according to Farber and Gibbons (1996) and Schönberg (2007) that the productivity assessment of the training firm is the same as that of the outside employer. In other words, cognitive and work-related ability have the same value in all jobs. This assumption seems not to be problematic in our setting because almost all employers that hire apprenticeship graduates also train apprentices themselves

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

(Mohrenweiser, 2015). Training and outside employers therefore do not seem to be structurally different with respect to their ability assessments and ability demands.

Control variables

Besides the measures for cognitive and work-related ability, a couple of additional individual and employer characteristics may have an impact on the probability to leave the training employer and on entry earnings after apprenticeship training. On the individual-level, we control for gender, age, nationality, and the schooling-level. These variables are 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 holds constant for example the training instructor, 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. Many papers stress that apprenticeships in larger and well-paying enterprises are more attractive for apprenticeship candidates (Soskice, 1994; Acemoglu and Pischke, 1998; Mohrenweiser and Zwick, 2015). These employers can apply stronger hiring standards and usually obtain better apprenticeship candidates. Graduating from well-paying and large training employers therefore may be public information for outside employers about the average productivity of apprentices (Wagner and Zwick, 2012).

4 Data and variables

We use the *Ausbildungspanel Saarland*, a data set linking Social Security Records with administrative exam files 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 between 1998 and 2005. We merge both data sets on the basis of the initial of name and surname, birthday, gender and the start date of apprenticeship.

From the chambers´ exam files, we use the final exam marks, the training occupation, and the school-leaving qualification. Training occupation and school-leaving qualification are more detailed in the data of the chambers 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 the earnings, nationality, gender, training firm size, and establishment identifier. Particularly the earnings information in the Social Security Records is very reliable because it is reported as exact daily earnings and is used to calculate social security and old age pension claims.

We detect apprenticeship graduates with a graduation identifier in the chamber data.[15] 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 graduates 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 of age 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.[16] Then, we identify peers of apprentices who learn in the same training establishment, occupation and graduation year (our cell definition) and calculate the bonus within cells. Since calculating the bonus is only possible in cells with two or more individuals, we drop individuals in cells with one apprenticeship graduate only.

Starting with 20,701 apprenticeship graduates, this procedure reduces our final sample to in 5792 individuals. The largest drop results when applying the two graduates per cell rule (observations drop by 58 per cent). The sample selection procedure generates an individual-

[15] 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.
[16] 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 clean the data.

level pooled cross-section dataset. Each remaining apprenticeship graduate enters the sample in the year of the first full-time employment after graduation.

5 Findings

Descriptive statistics

Table 2 summarises the variable definitions, means and standard deviations of the variables. The share of employer movers immediately after apprenticeship training is around 40 per cent in our sample. For employer movers and stayers, Table 3 separately displays the descriptive statistics. Movers have worse marks in their final apprenticeship exams than stayers.[17] 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 more likely to graduate from a lower secondary school track (*Hauptschule*), or without a school certificate.

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

However, marks and bonus payments are correlated at a significance level below one percent. The pairwise correlation is reasonable but low with a value of 0.17. The correlation is somewhat larger for stayers with 0.18 than for movers with 0.14. This means that both ability indicators measure different ability dimensions and work-related ability cannot be easily predicted from cognitive ability.

[17] Note that in Germany mark 1 is the best (with distinction) and mark 4 the worst assessment (just passed). For better interpretation, we multiply the marks by minus one. Hence, a higher number represents higher cognitive ability.

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.

Positive selection of stayers

We first test whether training firms are able to keep a positive selection of their apprenticeship graduates. We use a Linear Probability Model which has the advantage that we can include the 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 percent.

We find that the relative bonus payments and the final mark 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 graduate in order to analyse firms that select (or are able to attract) apprentices to stay. Models 1 and 2 take into account several individual and training establishment characteristics. An increase in the 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 bonus a day reduces the probability to leave by 8.4 percentage points.[18]

Model 3 includes cell fixed effects instead of establishment-level variables. This is our preferred estimation because it conditions on comparable peers learning and graduating under the same conditions[19]. The cell differences show us whether training firms are able to retain apprentices with the relatively best marks in comparison to their peers. The point estimates of bonus and marks increase in comparison to models 1 and 2. A one Euro higher bonus reduces the probability to leave by 14.2 percentage points and a better mark by one increases this probability by 9.8 percentage points. This reduction in the leaving probability is substantial given the unconditioned probability to leave of 40 percent. Finally, model 4 in

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

[19] This procedure renders marks to a private instead of a public signal since outside employers may have an idea about the general distribution of marks within an occupation but they do not know the relative differences within a cell.

Table 4 includes individuals in all cells, in other words even if all graduates leave or stay in the cell. The coefficients show comparable coefficients with regard to both productivity dimensions.

The control variables in these regressions have the expected signs. We also find a positive selection of stayers with respect to schooling background. Graduates with a medium-track secondary school certificate have a lower probability to leave than graduates holding a lower-track secondary school degree. German citizens have a higher probability to leave than graduates born abroad and older graduates have a higher probability to leave than younger graduates. Finally, graduates in larger training firms have a lower probability to leave than those graduating in smaller firms.

Hence, we can characterise employer movers after apprenticeship training as a negatively selected group with respect to the information on both ability dimensions.[20]

Asymmetric information on work-related ability?

Our test of asymmetric information in the labour market for apprenticeship graduates relies on the differentiation between the influence of work-related and cognitive ability on entry earnings of stayers and movers.

Table 5 displays augmented Mincer entry earnings regressions for apprenticeship graduates 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 respectively movers. Hence, the sample is more biased against larger firms.

We find that for stayers, the bonus has a significantly positive impact on entry earnings. A one Euro higher bonus a day increases the entry earnings by around 1.1 percent. For

[20] 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, Acemoglu and Pischke (1998) also conclude that there is adverse selection for quitting apprenticeship graduates in Germany.

movers, the bonus has no significant impact on entry earnings. On the contrary, marks in final exams strongly influence entry earnings. A one unit better mark increases the first full-time earnings of movers by 4.3 percent, a smaller impact than on the entry earnings of stayers (1.9 percent). The regressions controlling for cell fixed effects confirm the results and even have higher coefficients again.

All additional control variables show the expected signs. Females earn 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. This means that we find a pattern according to the assumption that schooling levels can also be used as a public signal on cognitive ability that mainly drives the entry earnings offer of outside employers. Furthermore, stayers in high-paying training firms (measured as the mean earnings-level of apprenticeship graduates in the establishment) not surprisingly also earn more as a skilled employee. More interestingly, apprentices leaving a high-paying training firm, also receive higher earnings in the first job after apprenticeship. This effect may result from the fact that the new employer takes into account a superior initial selection of apprentices by high-paying training employers that is known in local labour markets. Staying German nationals receive lower entry earnings than staying foreigners. Age has a positive impact on entry earnings for leavers only.

Our results show that cognitive ability has a stronger influence on entry earnings for movers than for stayers. Marks therefore have the characteristics of a public signal or in the words of DeVaro and Waldman (2012) a “publicly observable ‘announcement’ of the worker´s ability”. On the contrary, work-related ability significantly increases the earnings of stayers but not of movers. This pattern fits perfectly to our interpretation of bonus payments as private information. Hence, training employers are not able to generate an information advantage about cognitive ability but about work-related ability of their apprentices. The apprenticeship certificate renders the information about cognitive ability into a public signal but outside employers have no instrument to detect work-related ability. Work-related ability is therefore a possible channel for training firms to create monopsony power after apprenticeship training.

Robustness checks

It is important that we include both ability dimensions into the earnings equation in order to obtain the additional effect of each ability dimension on earnings. This interpretation however rests on the assumption that both ability dimensions are not too strongly correlated with each other and therefore measure different productivity dimensions. We therefore want to check, whether cognitive ability and work-related ability measures separate dimensions of productivity. If both indicators would be highly correlated, the addition of the second element would reduce the impact of the first element in the entry wage 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 as the employer can assess work-related ability. Therefore, we check the magnitude of the coefficients of each ability dimension when we exclude the other. Table 6 shows the regressions but only reports the key variables. The regressions control for the same variables as models 2 and 5 in Table 5. We find that the magnitude of the coefficients of marks and bonus remain remarkably stable when we eliminate the other (Table 6 in comparison to Table 5). This is an additional indicator that both indicators measure different dimensions of productivity.

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). Here, we follow the argument by von Wachter and Bender (2006) that training employers that experience a severe short-term demand shock usually do not keep as many apprenticeship graduates as usually. These firms reduce the retention rate of apprenticeship graduates dramatically in one year but have comparatively high retention rates in all other years. We implement this estimation strategy to identify those employer movers by using two retention rate reduction thresholds: first, training firms that reduce the average retention by more than 50 per cent in one single year (model 1 in Table 7) and second, training firms with a 20 per cent reduction (model 2 in Table 7). Table 7 reports the key variables (marks and bonus) 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. The bonus payments do not have a significant impact on entry earnings in both sub-samples. The marks 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.

6 Conclusions

This paper shows that training employers in the German apprenticeship training system can keep graduates who are better with respect to cognitive and work-related ability. Cognitive ability is identified by marks in the final exams and can easily be signalled to outside employers. Work-related ability however cannot credibly be signalled to outside firms and training firms can get an information advantage about this ability dimension. The information advantage leads to an adverse selection of training employer movers even if apprenticeship training is certified and marked by external institutions.

More specifically, training employers can base their entry earnings offer on hard to observe work-related ability (identified by bonus payments at the end of apprenticeship) but outside employers only obtain a noisy signal on this productivity dimension. Hence, work-related ability is only correlated with entry earnings of stayers but not of movers. On the contrary, training and outside employers easily observe reliable signals on cognitive ability (marks in the final exams). Cognitive ability has a stronger impact on the entry earnings offer of outside employers that can just rely on this signal, however.

This paper provides a new explanation of information advantages of training firms even if training is marked and certified by external institutions. The information advantage of training firms about work-related ability allows them to earn a return on their investment in transferable human capital. They can pay highly productive employees a wage below their productivity. Graduates cannot signal their high work-related ability to outside employers and therefore do not obtain better outside options. This paper therefore is the first direct empirical evidence for a frequently noted but never proven assumption (Chang and Wang, 1995; Acemoglu and Pischke, 1998; Leuven, 2005): German firms are willing to invest in certified, visible and transferable human capital of their apprentices because they obtain an information advantage on hard to observe ability dimensions. This gives training firms the opportunity to positively select the apprenticeship graduates they keep. However, this information asymmetry reduces at the same time the incentives of individuals to invest in their human capital.

Our results in addition give evidence that training firms indeed can adversely select employer movers after training. Apprenticeship graduates who stay with their training firm after graduation have better characteristics with respect to both ability dimensions in comparison to employer movers immediately after graduation. To the best of our knowledge, our paper represents the first direct empirical identification of adverse selection processes in labour markets and thereby complements studies on displacement losses such as the papers by Gibbons and Katz (1991) or more recently by Hu and Taber (2011), among others.

Moreover, this paper provides a novel indicator for work-related ability. Following DeVaro and Waldman (2012) and Waldman (2014), we introduce an indicator that can only be generated by the current employer. Our indicator is based on the earnings structure within a peer group of apprentices in the same training employer, occupation, and graduation year. The relative wage position of an apprenticeship graduate in relation to the peers is hard to observe for outside employers in comparison for example to relatively easy to observe promotions usually used in the literature. Our information on work-related ability therefore 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) who also generate indicators for information asymmetries on the labour market from earnings trajectories between cohorts.

Finally, the paper presents a relevant application of asymmetric employer learning that combines cognitive and work-related ability as recently proposed by Waldman (2014). It therefore combines two employer learning literatures that were traditionally separated in cognitive ability (Farber and Gibbons, 1996; Altonji and Pierret, 2001; Lange, 2007; Schönberg, 2007; Pinkston, 2009; Kahn, 2013; Kahn and Lange, 2014; Kim and Usui, 2014; Light and McGee, 2015) and work-related ability (Waldman, 1984, 1990; deVaro and Waldman, 2012; Waldman, 2014, Bognanno and Melero, 2015).

Finally, many empirical papers in the employer learning literature discuss the incentive effects (and their social welfare consequences) to obtain signals irrespective of their productivity enhancing effect (Lange, 2007; Waldman, 2014). It is beyond the scope of this paper but seems very interesting to assess whether the signalling value of good marks in apprenticeship exams induces apprentices to invest more than efficiently in their final exam success.

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.

Acemoglu D and Pischke J (2000). Certification of Training and Training Outcomes. European Economic Review 44: 917-927.

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.

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 (2012) Firm-sponsored general education and mobility frictions: Evidence from hospital sponsorship of nursing schools and faculty, Journal of Health Economics 32 (1): 149-159.

Benson G, Finegold D and Mohrmann S (2004). You Paid for the Skills now Keep Them: Tuition Reimbursement and Voluntary Turnover, Academy of Management Journal 47(3), 315-331.

Bognanno M and Melero E (2015). Promotion Signals, Age and Education, Journal of Economics and Management Strategy (forthcoming).

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.

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

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.

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.

Kriechel B, Mühlemann S, Pfeifer H, and S Schütte (2014). Works Councils, Collective Bargaining and Apprenticeship Training, Industrial Relations 53(2): 199-222.

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.

Light A and McGee A (2015) Employer Learning and the 'Importance' of skills, Journal of Human Resources (forthcoming).

Mansour H (2012). Does Employer Learning Vary by Occupation? Journal of Labor Economics 30: 415-444.

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 (2015). Recruitment and Apprenticeship Training, Industrielle Beziehungen (forthcoming).

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, Jahrbücher für Nationalökonomie und Statistik, forthcoming.

Mühlemann S, Ryan P and Wolter S (2013). Monopsony Power, Pay Structure, and Training, Industrial and Labor Relations Review 66(5): 1097-114.

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.

Pinkston J (2009). A Model of Asymmetric Employer Learning with Testable Implications, Review of Economic Studies 76: 367-394.

Ryan P, Wagner K, Teuber S and Backes-Gellner U (2013). Trainee Pay in Britain, Germany and Switzerland: Institutions, Market Forces, Market Power, European Journal of Industrial Relations 19(3): 201-220.

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

Sauermann J (2015). Reciprocity and Training Outcomes, paper presented at POEK 2015, Vienna.

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.

Wachter von 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 (2014). The Dual Avenues of Labor Market Signaling, working paper, Cornell University, Ithaca, NY.

Waldman M and Zax O (2014). An Exploration of the Promotion Signaling Distortion, MPRA working paper, Munich.

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	20,701
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)
Mark (public signal)	Mark in the final apprenticeship exam multiplied by minus 1 (-2.90; 0.74)
Log earnings bonus during training (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 apprentice graduates from nine year school track or does not pass the final school exam after nine school years (0.41; 0,49)
School certificate "Realschule"	Dummy variable equals 1, if apprentice graduates from a ten year school track (0.32; 0.47)
School certificate "Abitur"	Dummy variable equals 1, if the apprentice receives a university entrance allowance (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 the apprentice has German nationality (0.96; 0.19)
Establishment earnings level (training employer)	Average daily apprentice earnings at the end of apprenticeship 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
Mark in final apprenticeship exam	-3.01	-2.82	5.92
Log earnings bonus during training	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)
Earnings bonus (private information)	-0.093***	-0.084***	-0.142***	-0.051***
	(6.02)	(5.34)	(5.20)	(5.10)
Mark (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 graduates

	Stayers			Movers		
	(1)	(2)	(3)	(4)	(5)	(6)
Earnings bonus (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)
Mark(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)
Earnings bonus (private information)	0.011***	0.012*			0.008	0.018		
	(2.12)	(1.72)			(0.82)	(1.10)		
Mark (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)
Earnings bonus (private information)	0.009	0.020
	(0.12)	(0.49)
Mark deviation (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.

Mathematical appendix

We use the following notation: The individual work-related ability measure "bonus" *b* can take two expressions, high (*b = H*) and low (*b = L*). The individual cognitive ability measure "mark" *g* can also take two expressions, excellent (*g = E*) and average (*g = A*)[21]. In our basic model, we assume that the impact of information *g* on productivity *y* is the same for apprentices 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[22]. We therefore assume that an excellent mark adds a constant *c* to productivity in comparison to an average mark. Finally, we denote with *p* the probability that an apprenticeship graduate received a low bonus. We assume that apprentices with excellent marks received a high bonus with a higher probability than apprentices with an average mark ($p^A > p^E$). Table A1 gives an overview of our model.

Table A1: Specification of Productivity *y*

	Average mark *G*	Excellent mark *G*
Low bonus *b*	$y_L^A = b_L$	$y_L^E = b_L + c$
High bonus *b*	$y_H^A = b_H$	$y_H^E = b_H + c$
Proportion low bonus *p*	p^A	p^E

At the end of the training period, all employers and apprentices observe the marks of the apprenticeship graduates. Marks measure cognitive ability in the trade and, hence, do not reveal all dimensions of productivity. The missing dimension of productivity, work-related ability, is known to training employers and to their apprentices. Outside employers only observe a common noisy signal *s* about these privately known dimensions of individual productivity. An example for this signal may be the performance in a job interview (Lange,

[21] We do not use "high" and "low" for the differentiation of marks in order not to unnecessarily confuse notation.

[22] Schönberg (2007) proofs in her appendix A6 that all previous results also hold if cognitive and work-related ability are complements. As a consequence, the adverse selection should be stronger for apprenticeship graduates with high cognitive ability, a result also derived by Acemoglu and Pischke (1998). In addition, the difference between the impact of work-related ability on entry wage offers of incumbent and outside employers should be stronger for apprenticeship graduates with high work-related ability than for apprenticeship graduates with low work-related ability.

2007). This signal can be obtained by all (interested) labour market participants and is independent from marks.[23]

In our simple model, the potential signal *s* about work-related ability takes only two values, good (*s = G*) and bad (*s = B*). The outside employer uses the signal in order to infer the assessment of the incumbent employer on work-related ability of the apprenticeship graduate. This inference therefore is correct if an apprenticeship graduate with a good signal also received a high bonus and a graduate with a bad signal received a low bonus. 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 bonus apprentice when the information is good or bad, given the mark 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 symmetric information for incumbent and outside employers on the productivity of apprenticeship graduates) means that all low bonus apprentices generate a bad signal and that all high bonus apprentices 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 bonus apprentice is assessed purely by the a-priori probability depending on the distribution of high-bonus and low-bonus apprentices, $\Pr(L|k,B) = \Pr(L|k,G) = p^k$.

Apprentices may have had experienced disutility θ during training that is unknown to the training employer – it therefore cannot compensate the apprentice for individual disutility in order to retain the apprentice. The disutility shock is independent of the bonus and mark (and the signal *s*). We assume that the disutility shock in the training firm is drawn from a

[23] 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. Hereby the graduate can reveal his or her social ability, the training firm therefore does not necessarily perform a job interview by itself in order to reveal the signal.

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 in the outside employer is zero.

The graduate receives a wage offer from the incumbent employer (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 incumbent employer after training or not. In other words, the apprenticeship graduate stays with the training employer only if $w + \theta > v$.

The incumbent employer has the advantage that it can make counteroffers to those apprenticeship graduates it wants to retain when $w < v$[24]. The outside employer does not observe the wage offer of the training employer w and therefore a bidding war on apprenticeship graduates is not possible (Pinkston, 2009).

The training employer instead maximises the difference between productivity y_i^k (that depends on bonus and marks according to Table A1) and $w_i^{k,s}$ (that additionally depends on the signal), taking marks g, the productivity signal s, and the probability that the apprenticeship graduate stays with the training employer given v and the distribution of θ into account. The probability that the apprenticeship graduate 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 employers offer a wage taking into account marks g and signal s. The training employer therefore maximises profits by setting wages according to:

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

[24] Note that apprenticeship graduates with θ > 0 leave the training firm if $w = v$ and therefore the training firm has to offer a higher wage than the outside firm in order to reduce the risk of losing the trained apprenticeship graduate.

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

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

Outside employers pay the expected productivity of those apprenticeship graduates who switch firms after they received both wage offers. Outside employers can freely offer wages and even poach on the labour market for apprenticeship graduates, i.e. pay a wage that is higher than the wage the apprenticeship graduate would have earned when staying in the training firm. Apprenticeship graduates in addition do not have to refund training investments when they directly leave the training firm after graduation. In addition, there are usually more than one outside employer interested in the apprenticeship graduates. Outside employers cannot make a profit in the long run, accordingly. The outside wage offered can be written as:

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

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

Now, we can derive implications of symmetric and asymmetric information on the coefficients of the private and public productivity information for the entry wages of leaving and staying apprenticeship graduates. When information on apprenticeship graduates is symmetric ($q = 1$), signal s completely reveals the bonus payment and therefore the work-related ability 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 apprenticeship graduates sends a good signal, and $v_L^k = y_L^k$ if the apprenticeship graduate sends a bad signal. As the productivity difference between high- and low-bonus apprenticeship graduates is a constant (b_H - b_L = Δ) that is independent of marks, the training employer offers worker´s productivity minus this constant, $w_b^k = y_b^k - \Delta$. In other words, the wage offer w increases by the full amount of the productivity difference between low- and high-bonus apprenticeship graduates. As a consequence, also the difference between training and outside employers´ wage offers is equal and low- and high-bonus apprenticeship graduates have the same probability of leaving the incumbent 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 wage offer of outside employers between high-bonus and low-bonus apprenticeship graduates is smaller than Δ because a certain share of apprenticeship graduates with a good signal obtained a low bonus. Hence, high-bonus apprenticeship graduates have a smaller incentive to leave the training employer voluntarily than low-bonus apprenticeship graduates. 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-bonus apprenticeship graduate leaves the incumbent 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-bonus apprenticeship graduate leaves the incumbent 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 for the quit rate of low-bonus apprenticeship graduates to be higher than that of high-bonus apprenticeship graduates:

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

When q < 1, low-bonus apprenticeship graduates are more likely to leave the training employer than high-bonus apprenticeship graduates. This is unconditional on the signal observed by the outside employer. The probabilities that a low-bonus and a high-bonus apprenticeship graduate leave 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-bonus apprenticeship graduate than of a high-bonus apprenticeship graduate. When we assume q = 0.5, the outside employer offers the same wage for low-bonus and high-bonus apprenticeship graduates, the training employer however offers a higher wage for high-bonus than for low-bonus apprenticeship graduates. 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 wage mark-up for high-bonus apprenticeship graduates and low-bonus apprenticeship graduates increases stronger with better information quality of the signal q 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 wage offers for high-bonus apprenticeship graduates decreases with signal quality ($d(w_H^G - v^G)/dq < 0$) and at the same time, the difference between the wage offers for low-bonus apprenticeship graduates increases with the signal quality ($d(v^G - w_L^B)/dq < 0$). If we totally differentiate the first-order wage 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 $0.5 \leq q <$ 1are $d(v^G - w_L^G)/dq < 0$ 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-bonus apprenticeship graduates are more likely to leave the training employer unconditional on the signal observed under asymmetric information. In other words, with information symmetry, low- and high-bonus apprenticeship graduates have the same probability to quit. With information asymmetry, the chances of low-bonus apprenticeship graduates to leave the training employer are higher.[25]

[25] 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

We can turn now to the main feature of the model – the influence of bonus and marks on entry wages of apprenticeship graduates under symmetric and asymmetric information. From equation (5), it is clear that the bonus has the same impact on outside and training employer wage offers under information symmetry. With asymmetric information, wage offers of training employers vary more with bonus than wage offers of outside employers. The difference between expected wage offers *w* for high-bonus and low-bonus apprenticeship graduates 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 wage offers *v* for high- and low-bonus apprenticeship graduates 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-bonus apprenticeship graduates to be higher than the quit rate of high-bonus apprenticeship graduates in the asymmetric information case, compare equations (6) and (7).

The impact of marks on entry wages is again the same for training and outside employers under symmetric information. Conditional on the bonus, the difference between outside wage offers for apprenticeship graduates with average and excellent marks is the full productivity mark-up ($y_i^E – y_i^A = c$). In the case of asymmetric information, the wage offers of outside employers to apprenticeship graduates with a good mark not only reflect the productivity-enhancing effect of having a good mark, but also the fact that apprentices with good marks have a higher probability to have a high bonus p^E than apprentices with low marks p^A.

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

characteristics such as schooling, tenure, and job level. We use bonus payment as a measure of performance ratings.

$$E\left[w \mid E,L\right]-E[w \mid A,L]=q\left(w_L^{E,B}-w_L^{A,B}\right)+(1-q)\left(w_L^{E,G}-w_L^{A,G}\right). \quad \text{(A1)}$$

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

$$E\left[v \mid E,L\right]-E[v \mid A,L]=q\left(v^{E,B}-v^{A,B}\right)+(1-q)\left(v^{E,G}-v^{A,G}\right). \quad \text{(A2)}$$

We therefore need to show that the differences (A1) < (A2) or in other words $w_L^{E,s} - w_L^{A,s} < v^{E,s} - v^{A,s}$ for $s = G,B$. Since an excellent mark has the same impact on productivity for low-productivity and high-productivity workers, 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 wage maximisation problem of training firms (equation (3)) leads to:

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

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 mark has a higher impact on the wage offers of the outside employers than on the wage offers of training employers for low-productivity and high-productivity apprenticeship graduates.