University of Zurich

Socioeconomic Institute
Sozialökonomisches Institut

Working Paper No. 0308

Training Intensity and First Labor Market Outcomes of Apprenticeship Graduates

Rob Euwals and Rainer Winkelmann

September 2003

Socioeconomic Institute
University of Zurich

Working Paper No. 0308

Training Intensity and First Labor Market Outcomes of Apprenticeship Graduates

September 2003

Author's addresses

Rob Euwals
E-mail: R.W.Euwals@cpb.nl

Rainer Winkelmann
E-mail: winkelmann@sts.unizh.ch

Publisher

Sozialökonomisches Institut
Bibliothek (Working Paper)
Rämistrasse 71
CH-8006 Zürich
Phone: +41-1-634 21 37
Fax: +41-1-634 49 82
URL: www.soi.unizh.ch
E-mail: soilib@soi.unizh.ch

Training Intensity and First Labor Market Outcomes of Apprenticeship Graduates*

Rob Euwals
CBP Netherlands Bureau for Economic Policy Analysis, IZA, and CEPR

Rainer Winkelmann
University of Zurich, IZA, and CEPR

August 2003

Abstract

The apprenticeship system is the most important source of formal post-secondary training in Germany. Using German register data - the IAB Employment Sample - we find that apprentices staying with their training firm after graduation have longer first-job durations but not higher wages than apprentices leaving the training firm. Retention rates, first job durations, and post-apprenticeship wages are all increasing functions of training intensity. Some implications for the ongoing debate as to why firms are willing to invest in general training are discussed.

Keywords: Training, wages, job-duration
JEL Classification: C41, J24, J31, J44

1 Introduction

In international comparisons on education, vocational training, and labor market performance, Germany shows a high involvement of firms in the education and training of the young by means of a well-developed apprenticeship system with several hundreds of accredited training programs. This fact is often interpreted as a positive characteristic of the German education and labor market system, as it may contribute to low youth unemployment rates and high general skill levels of the workforce.

*An earlier version of the paper was written while we both were at the Institute for the Study of Labor (IZA), Bonn. It appeared as IZA Dicussion Paper No. 319 under the title "Why do firms train? Empirical evidence on the first labor market outcomes of graduated apprentices". We wish to thank Wiji Arulampalam, Jan van Ours, Stefan Wolter, and our former colleagues at IZA for valuable comments and helpful discussions, and the German Institute for Employment Research (IAB) for providing the data.

A particular feature of German apprenticeship training (AT) is that it mainly provides general training and thus portable skills. In addition, there is empirical evidence that firms make substantial net investments into AT (e.g., von Bardeleben et al., 1995). The existence of (partially) firm-financed apprenticeship training is not easily reconcilable with standard human capital theory under perfectly competitive labor markets according to which profit maximizing firms should not pay for general training but rather free-ride on the investment of others. Most explanations in the literature introduce some sort of market imperfection which allow training firms to recoup the costs of the training by extracting rents from the graduated apprentices who stay with the training firm. Such explanations are based on asymmetric information on the ability of apprentices (Acemoglu and Pischke, 1998), high mobility costs and low general labor turnover in the German labor market (Harhoff and Kane, 1997), or compressed wage structures (Acemoglu and Pischke, 1999a, 1999b). In these models firms are able to pay post- apprenticeship wages below marginal productivity and thus earn a return on their investment.

Concurrently, a number of empirical studies have investigated various aspects of AT. Two recent papers use firm level data to estimate the firms' demand functions for apprentices (Dietrich, 2000, Fougère and Schwerdt, 2001). Other papers have used individual level survey data to establish the wage returns to apprenticeship training (Winkelmann, 1996b), the effect of post-apprenticeship mobility on wages and wage growth (Werwatz, 1996, Winkelmann, 1996a, Dustmann et al., 1997, Clark, 2001, Bougheas and Georgellis, 2001), the wage effects of moves out of the training occupation (Werwatz, 1997, Clark, 2000), and the incidence of unemployment and non-employment spells during the transition from apprenticeship to regular employment (Winkelmann, 1996a, Franz et al. 2000, Franz and Zimmermann, 2000, Riphahn, 2002).

Despite the substantial size of the literature, few "stylized" facts have emerged. One of the recurrent findings is that about 70 percent of apprentices stay with their training firm initially after graduation, whereas the fraction of those staying for five years or longer is about one-third. Also, retention rates tend to be higher in larger firms than in smaller firms. More often than not, the empirical evidence is inconclusive or even contradictory.

For instance, there are no robust results available so far, whether jobs are more stable and wages higher for those AT graduates staying in the training firm than for those moving on to another firm. And yet, this type of evidence is crucial for understanding the nature of the AT, including the firms' motivation to contribute to it.

We argue in this paper that the lack of solid evidence is in part due to shortcomings of the data sources used in the previous literature, and we set out to address this deficit by using a more appropriate dataset that has become recently available. The problem with previous data sets, such as the German Socio Economic Panel (GSOEP) or the German Qualification and Career survey (Q&C survey), is that in order to understand the effect of AT on the trainees' skills and labor market choices, it is most instructive to study the period immediately following the apprenticeship, when the external labor market comes to its fullest force and before a host of other influences has obscured these initial relationships.

In the GSOEP, apprentices can be observed during their transition to employment in principle. However, the number of observed transitions is relatively small. The Q&C survey, by contrast, is large, with 20-30 thousand observations. However, people in the survey completed their apprenticeship training on average almost 20 years before the survey date. While some retrospective information is available, in particular on the transition pattern, its reliability is doubtful. Moreover, immediate post-apprenticeship wages are not collected at all. Finally, it is unclear whether the experiences of AT graduates of some twenty or thirty years ago have any close relation to the experiences of current and future cohorts of labor market entrants.

In our study, we overcome these difficulties by using official German social security register data. We have access to the IAB Employment Sample, a one percent sample of the complete employment histories of all workers subject to social security contributions at least once between 1975 and 1995. Even after restricting the sample to the part of the male cohort born between 1960 and 1965 who completed apprenticeship training, we retain 16,281 observations. The official register data provide accurate records on wages paid, as well as on job durations (with daily accuracy). The dataset offers a unique opportunity to provide more reliable empirical evidence on several of the key issues discussed in the

literature.

A second contribution of our paper is that we are able to analyze an additional aspect of apprenticeship training that has been neglected hitherto, namely the question of how variation in training intensity affects the first labor market outcomes of AT graduates. We are able to derive two new indicators of training intensity from the data, and we study how training intensity affects retention rates, post-apprenticeship wages, and first-job durations.

2 The Importance of Training Heterogeneity

Previous commentators have pointed out that the coexistence of "low-quality" and "high-quality" apprenticeship is a central feature of the German apprenticeship system (e.g., Steedman, 1993). This coexistence can explain why AT attracts school leavers with university entrance qualification headed for skilled jobs as well as less educated youths headed for semi-skilled employment. The breadth of AT may also increase adaptability of the system to structural changes in the labor market.

Failure to properly control for this quality component – the intensity of training and the amount of the firm's investment – affects the interpretation of previous empirical results in this area. The basic mechanisms we have in mind are as follows. First, firms providing more intensive training have a larger incentive to retain their apprentices. Conversely, firms which, for instance for technological reasons, find it optimal to heavily invest in their apprentices have a larger disincentive to train beyond their own expected needs for skilled worker recruitment than firms with a lower investment. They take on relatively fewer apprentices and offer skilled workers contracts to a higher proportion, if not to all, of them. Thus, immediate retention rates should be an increasing function of training intensity.

Second, post-apprenticeship wages should be an increasing function of training intensity as well. This is simply due to the fact that wages are linked to marginal productivity if workers can potentially move to other firms. Wages may not increase one-to-one with marginal productivity if part of the training is firm-specific, or if information problems or mobility barriers reduce the value of the outside option. Still, a positive association is to be expected. An immediate consequence is that mover-stayer post-AT wage differentials are

underestimated in models that ignore training intensity, because movers tend to be trained less intensively.

Third, the link between training intensity and the duration of the first job is likely positive as well. High-investment training firms have a larger incentive to offer a wage schedule that reduces turnover than low-investment firms. Although the initial investment cost are sunk at this stage, both training firm and outside firm are likely to engage in continuing on-the-job training, the amount of which typically is higher for more skilled workers, in this case apprentice graduates with high training intensity.

From the above, it is clear that training intensity is an important variable when analysing AT. Yet, training intensity is difficult to measure. The variables used in past research, most importantly occupation and firm size, do capture some aspects of training intensity. Unfortunately, their interpretation is difficult. First, while training intensity typically varies from occupation to occupation, so do a number of other things that affect labor market outcomes of graduates as well, such as skilled worker demand. This is even more so for firm size. Larger firms tend to have a larger training intensity, but they also have a larger internal labor market, which facilitates higher retention rates. The two effects complement each other.

We propose here two alternative indicators for training intensity, both of them constructed from the IAB data, that we find potentially superior. The first indicator is the wage paid during the apprenticeship training, twelve month after the beginning of training. It is known that the AT-programs with the most demanding training component also pay high wages. A reason may be that higher paying firms are able to attract more able apprentices, which is more relevant under high training intensity, assuming that ability and training are complements in the production of skills. In addition, there is a complementarity between training and capital intensity. Technologically more advanced firms have an incentive to pay higher training wages. Finally, training wages have a direct positive impact on the firms' training *investment* because of the cost and regardless of training intensity.

A second indicator is the duration of the apprenticeship. We classify apprenticeships as "short", "medium" or "long", depending on whether the duration up to December 31 of the

year prior to graduation was less than 2 years, 2-3 years, or longer than 3 years. The AT duration is an indicator of the firms' investment as national regulations of training content and intensity prevent firms from prolonging apprenticeships beyond what is necessary from a training point of view.

3 The IAB Employment Data

The data for this study are drawn from the IAB Employment Sample 1975-1995, which is a one percent sample of employment histories of all workers in former West Germany with at least one recorded spell during the 21-year period (see Bender, Haas and Klose, 2000). The selected subsample includes observations on 16,281 male apprentices born between 1960 and 1965.[1].

As the data come from official registers, accuracy is an important advantage. By law all employers are required to report information on their workforce to the German Social Security Administration, who administers the health insurance, statutory pension, and unemployment insurance. The population includes all workers and salaried employees, as long as they are not exempt from paying social security contributions. Exempt are civil servants, family workers (without pay), and those in marginal employment (this exemption was removed in 1999). Remarkably, and for the purpose of our study of great importance, the data also include apprentices who, regardless of their wages, are treated as regular employees by the social security system. All in all, the employee register covered nearly 80 percent of all employed persons in West Germany in 1995.

The data set has an event-history format that differs substantially from usual survey data such as the GSOEP or the Q&C survey. It is organized in terms of spells between "notifiable" events. Notifiable events include the start and termination of employment at a certain employer, and an obligatory end-of-year notification for continuing employment relations. Thus the maximum duration between two notifications is 365 days. Apart from

[1]Since the data start in 1975 and end in 1995, we observe the labor market histories for this cohort from the age of 15 to the age of 30- 35, which means that we observe all AT spells from the beginning. Moreover, the length of the observed post-apprenticeship labor market history is sufficient to measure the first employment spell without excessive right-censoring. See Bender and Dietrich, 1994, for details on identification of apprenticeship spells in the IAB data

notifications by employers, records from the social security administration are used to fill in information on unemployment spells, if any, and type of termination. Gaps in a worker's history indicate any period during which a worker either took up exempt employment or left the labor force.

Each spell-record provides information on the starting and end dates (in days), a personal identifier and an establishment number. Further variables include individual and family characteristics such as gender, year of birth, nationality, marital status, and highest qualification. In addition, each spell- record includes an occupational code, gross earnings[2], an industry code and the size of the establishment (number of workers).

The rules underlying the generation of this dataset imply that it is not possible in all cases to properly identify the day and month of apprenticeship completion.[3] Formally, apprentices can be in one of three states on the day following graduation: they continue working in the training firm, they immediately switch to a new employer, or they stop being employed. From the viewpoint of the IAB data, the second and third options imply termination of employment at the training firm. In these cases, a notification is given by the training firm, and the exact termination date of the apprenticeship is known. If the apprentice is retained as a regular employee, no notification is made. In this case, the mandatory end-of-year-notification by the training firm, now employer, reveals two status changes: the notified occupational status changes from "apprentice" to "skilled worker" and the educational level changes from "without vocational qualification" to "with vocational qualification". Only the year of graduation is known for these workers. We adopt the following two definitions:

Stayer A stayer is an apprentice whose first job after apprenticeship is in the training firm.

Mover A mover is an apprentice whose first job after apprenticeship is in a firm other than the training firm.

This definition allows for a non-employment gap between apprenticeship and first job for

[2]Earnings are top coded but the censoring threshold is not reached for any of the observed apprenticeship- or first post-apprenticeship wages

[3]The final certification process is frequently completed in early summer but many exceptions exist.

both movers and stayers. This is important, since many young men are drafted into the army (for 15 months) following their training. If such a person returns to the same firm after the military, he is classified as a stayer. Similarly, for men who are drafted while in their first job, the first job duration is calculated as the sum of the employment durations before and after the military related leave of absence.

In slight abuse of language, we often use the term "firm" when the data really identify only establishments. This is a certain weakness, as one cannot distinguish workers moving within a multi-establishment firm from those moving between such firms. Both types of move lead to a change of establishment number. This aspect may tend to overstate the post-apprenticeship mobility somewhat, in particular for sectors and occupations where multi-establishment firms are prevalent.

4 Empirical Results

Table 1 shows the mean values for selected variables, both overall and by birth cohort. Observations are almost equally distributed over the six cohorts. The average apprentice is 17 years old when he starts to train, with an upward trend. The proportion of trainees with Abitur increased steadily from 3 percent for the 1960 cohort to 6 percent for the 1965 cohort. Both the increased starting age and the increased proportion of apprentices with Abitur result from the secular trend towards higher schooling. In response to this trend, new apprenticeship programs were developed and existing ones modified in order to cater for the more educated as well.

——— TABLE 1 ———

The Table shows that very few apprentices are married. Only 2 to 3 percent are foreigners, substantially less than the overall proportion of non-Germans in the labor force (around 10 percent). This underrepresentation indicates that foreigners are less likely to start an apprenticeship than German nationals, even if they are second generation.

A substantial fraction of training occurs in very small or small firms. For example, 65 percent of all apprentices are trained in firms with less than 50 employees. 19 percent of all

apprentices are trained in very large firms with more than 500 employees. On average, the daily wage one year after starting training was 19.6 Deutsche Marks for the 1960 cohort, increasing by 9.7 percent (keeping prices constant in 1985 Deutsche Mark) to 21.5 Deutsche Mark for the 1965 cohort. 25 percent of all apprenticeships observed in the sample were of short duration, 53 percent of medium duration and 22 percent of long duration.

Finally, there is some information on the transition from AT to work. Unemployment spells and non-participation spells are quite common between training and the first job. Episodes of non-participation relate in many cases to mandatory military service (or its substitute for conscientious objectors). The risk of unemployment was highest for apprentices born in 1964 who entered the labor market around 1984. The correlation between the immediate post- apprenticeship unemployment incidence and the official German unemployment rate is about 0.9. Thus, the transition is affected by the business cycle, a relationship that Franz and Zimmermann (1999) study in more detail. The overall retention rate of trainees in their training firm was 71 percent.[4] Again, there is a strong cyclical component. The retention rate is lowest for the cohort that entered the labor market at peak unemployment.

4.1 The immediate retention rate

About 71 percent of the graduated apprentices start working in their training firm. They are, following our above definition, *stayers*. The remaining 29 percent are *movers*. Table 2 presents estimation results for linear probability models in which the dependent variable is 1 if the apprentice is a stayer and 0 if he is a mover. Three different models were estimated, including successively more explanatory variables. The most basic model in column (1) includes indicator variables for firm size only. The smallest size (1-9 employees) is the reference size. The coefficients confirm the findings in the previous literature based on different datasets. Retention rates are an increasing function of firm size. They increase from 66 percent for the small firms to 75 percent for very large firms with more than 500

[4]Note that both stayers and movers can have episodes of unemployment or non-employment between AT and first employment. For this reason, the percentages don't add up to 100.

employees.

——— TABLE 2 ———

The next model in column (2) adds a number of demographic variables as well as controls for training intensity. An F-test clearly confirms the joint significance of the additional variables. However, the demographic variables are individually rather unimportant. We find, as expected, that the effect of training intensity is large and highly significant. A 10 percent increase in the wage received during the apprenticeship increases the retention rate by 2 percentage points. Similarly, the retention rate increases with the duration of the apprenticeship. Incidentally, the magnitude of the duration effect is substantially larger than the firm size effect noted in model (1). We find a 16 percentage point difference in the retention rate between short and long apprenticeships. This effect is relatively robust to the inclusion of year and occupational dummies, which both are highly significant as a group. In other words, there is sufficient variation in training intensity *within* each of the 48 occupations to identify an intensity effect, and it is found to be large. The firm size effect, by contrast, is quite unstable once further variables are included in the model. In model (3), for instance, there is no longer any significant difference of having been trained in a firm with 10-49 employees or in a firm with 500+ employees.

4.2 Duration of the First Job

The second important determinant for a training firm's ability to recoup the costs of training is the duration of the first job. Training firms profit from long first job durations of their trained apprentices not only in the sense that they make use of their skills for a long time, but also in the sense that it keeps recruitment costs low.

There are two studies that have analyzed the duration of the first job after apprenticeship before: Winkelmann (1996a) and Franz and Zimmermann (1999). Winkelmann, using data from the GSOEP, reports five year retention rates of 30 percent, whereas Franz and Zimmermann, based on a sample from the 1991/2 German Qualification and Career Survey, including both men and women, find retention rates after five years of about 35

percent. In contrast, our estimated five-year retention rate is only 19 percent. A number of possible explanations for the discrepancy between our estimate based on register data and the previous estimates based on survey information come to mind. One is that looking at men only, or at men and women jointly, may make a difference. Second, register data are more accurate almost by definition. For instance, very short job spells at the start of ones career might be "overlooked" in surveys. Second, some of the discrepancy might be explained by definitional differences between establishments (in the register) and firms (in the surveys). In any case, it is clear that there is a large amount of job churning at the beginning of people's careers.

Based on a first job duration analysis for stayers, Franz and Zimmermann (1999) conclude that durations are longer for large firms and firms in the industrial sector. Winkelmann (1996a) compared the first job duration of the movers and stayers and found, somewhat counterintuitively, that movers tend to have longer durations than stayers. We have now the chance to settle this issue with much better quality data, and including controls for training intensity.

In principle, the job duration t is measured on a daily basis, which is precise enough to treat it as a continuous variable. However, for stayers we observe only an upper and a lower bound for the first job duration, and hence the likelihood contribution for stayers is discrete. A parametric duration model is ideal to handle such a mixed discrete-continuous duration problem, and we start from a mixed proportional hazard rate model.

The hazard rate (or instantaneous exit rate) $\lambda(t|x,\varepsilon)$ is defined as the limit of the probability that the spell is completed during the interval $t+h$ given that it has not been completed before time t, for $h \to 0$. Define a vector of individual characteristics x, a conformable parameter vector β, and an unobserved individual heterogeneity term ε and let

$$\lambda(t|x,\varepsilon) = \lambda_0(t)\exp(x'\beta+\varepsilon) \tag{1}$$

where $\lambda_0(t)$, the baseline hazard, is common to all individuals and allows for (non-monotonic) duration dependence. For any hazard function, it holds that the survivor function $S(t)$ is

given by

$$S(t) = \exp\left[-\int_0^t \lambda(s)ds\right] \qquad t \geq 0$$

Conditional on ε, the survivor function of the proportional hazard model has a particularly simple form:

$$S(t|x,\varepsilon) = \exp\left[-\int_0^t \lambda_0(s)ds \exp(x'\beta + \varepsilon)\right] \tag{2}$$

Moreover, one obtains the density $f(t|x,\varepsilon)$ by differentiation of $F(t|x,\varepsilon) = 1 - S(t|x,\varepsilon)$ with respect to t:

$$f(t|x,\varepsilon) = \lambda(t|x,\varepsilon)\exp\left[-\int_0^t \lambda_0(s)ds\right] \tag{3}$$

Thus, one has the basic inputs for maximum likelihood estimation, namely the density for non-censored observations and the survivor function for censored observations. However, to fully specify the model, we need to integrate the likelihood contributions over ε, and to specify parametrically the baseline hazard. First, we assume that $u = \exp(\varepsilon)$ is gamma distributed independently of x with expectation 1 and variance σ^2 (Lancaster, 1979, Abbring and Van den Berg, 1998). The closed form of the survivor function conditional on x can be shown to be

$$S(t|x) = \left[1 + \sigma^2 \int_0^t \lambda_0(s)ds \exp(x'\beta)\right]^{-(1/\sigma^2)} \tag{4}$$

with corresponding density

$$f(t|x) = \lambda_0(t)\exp(x'\beta)\left[1 + \sigma^2 \int_0^t \lambda_0(s)ds \exp(x'\beta)\right]^{-(\sigma^2+1)/\sigma^2} \tag{5}$$

Second, we specify a piecewise constant baseline hazard. In the most general case, we would estimate a different parameter for each day such that $\lambda_0(t) = \lambda_t,\ t = 1, 2, \ldots, T$, where T represents the maximum possible number of days. The integrated hazard is simply

$$\int_0^t \lambda(s)ds = \sum_{i=1}^{t} \lambda_i$$

As estimating T parameters would be too demanding, even with our large number of observations, we impose equality restrictions on these parameters on a monthly or yearly basis.

To tackle the problem of imprecisely measured first job durations for stayers, we proceed as follows. We know that the realized duration t is between a certain observed minimum duration $t_{\min}$ and a certain observed maximum duration $t_{\max}$. The probability of such an observation is:

$$P(t_{\max} \geq t > t_{\min}|x) = S(t_{\min}|x) - S(t_{\max}|x) \tag{6}$$

A third type of observation occurs due to the limited observation period: observations are right censored if graduates are still in their first job on December 31, 1995. Taking the three different kinds of observations together, the log-likelihood function can be written as

$$\log L(\beta, \sigma^2) = \sum_{\text{uncensored}} \log f(t|x) + \sum_{\text{partly censored}} \log[S(t_{\min}|x) - S(t_{\max}|x)] + \sum_{\text{right censored}} \log S(t|x)$$

In the spirit of the foregoing discussion, we are most interested in finding out, whether movers and stayers have different patterns with regard to their labor mobility and thus job durations, once we control for training intensity and other variables. We stratify the model by allowing for separate baseline hazards for movers and stayers. This approach is more general, and thus potentially more informative, than just allowing a proportional difference in the baseline for movers and stayers, as would be obtained if a dummy were included among the regressors.

——— TABLE 3 ———

Table 3 presents the maximum likelihood estimation results, again for three model versions. A first model includes the demographic variables, the training firm size dummies (which equal the first-job firm size for stayers but not necessarily so for movers) and a set of indicator variables for year and occupation. In a second model, the training intensity variables are added, while in a third, we add two further dummy variables capturing whether movers move to a smaller or to a larger firm. This last model specification accounts for the fact that movers frequently, although by no means exclusively, move to larger

establishments where job durations tend to be longer. An alternative, almost equivalent, specification would add separate controls for training firm size and first-job firm size.

We start with a remark on the variance of the unobserved heterogeneity term u. Given the substantial standard errors, the null hypothesis of the absence of unobserved heterogeneity is not rejected. This is not an uncommon finding, given that our baseline hazard is specified very flexibly and identification becomes a problem. We still decided to maintain the model with unobserved heterogeneity in order to protect estimates against spurious duration dependence.

Likelihood ratio tests show that model (3) is the preferred specification. Among the individual characteristics, marital status and pre-AT schooling have significant impacts: a graduated apprentice married at the time of starting the first job has a longer expected first job duration, as does an apprentice with Abitur. Both effects are quite large, as Abitur lowers the baseline hazard by approximately 42 percent, while being married lowers the baseline hazard by approximately 24 percent, a typical mobility cost effect. The age polynomial is insignificant, a result easily explained by the relative homogeneity of the sample since we look at men at the start of their careers only.

The training intensity variables are highly significant. For instance, a ten percent increase in the apprenticeship wage reduced the hazard of leaving ones first job by three percent. There is also strong evidence for substantial firm size effects, with lower hazards for employees in larger firms. In our specification, this expresses itself for movers in an increased hazard for those who move to a smaller firm and decreased hazard for those who move to a smaller firm (relative to the size of the training firm).

Figure 1 shows the stratified baseline hazard for movers and stayers for an otherwise identical representative individual.[5] According to theories based on asymmetric information and learning, the hazard rates should be different especially in the first months as for the stayers some information should be revealed already during the training period. Formal

[5]The reference person has German nationality, is 18 years old, unmarried, has no Abitur, ln(wage) = 3, medium training duration and the size of both training firm and employment firm is 50-499.

tests on the hazard rates reject the null hypothesis of no difference.

——— Figure 1 ———

With the exception of the first quarter, the probability of leaving the first job is indeed larger for movers than for stayers over the first three years. The largest differences occur between month 4 and month 12 after which differences become small. A likely explanation for the initial "excess separation rate" of stayers is an institutional one. Even if employers and apprentices agree to eventually separate, it is not uncommon that a temporary skilled-worker contract is offered in order to give apprentices time to search for another job, or to bridge a the temporary gap until the start of military service. In some industries, such temporary retention clauses are even part of collective bargaining agreements.

Figure 2 shows how these hazard rates translate into predicted survivor rates. After 1 year the survivor probability of stayers is larger than the survivor probability for movers. After 5 years the survivor probability of stayers is about 11 percent above the survivor rate of movers. Therefore a stayer trained and employed by a firm of a certain size has a longer first job durations than a mover trained and employed by a firm of the same size.

——— Figure 2 ———

However, the differences are not as large as one might expect. One can compare the mover-stayer differential to the effects of variations in other variables. For example, the five year retention rate of stayers with long AT duration exceeds the five year retention rate of stayers with medium AT duration by 10 percent.

Also, one can argue about the correct measurement of first job durations for stayers. One possibility is to set the clock back to zero once regular employment starts, as in Figures 1 and 2. The other possibility is to add the time of the apprenticeship to the overall job duration. Clearly, the job termination hazards of stayers would then exceed the hazards of movers at all comparable durations.

4.3 First Wages after Apprenticeship

A particular advantage of our data is that we observe the wages in the first job after the apprenticeship, whereas most other empirical studies on the AT observe the wages many years after the apprenticeship. Wages are measured on a daily basis (including weekends and holidays) in real 1985 DM. For the first wage in the first job we actually have to use the wage of the second notification in the first job, as for stayers the wage of the first notification is a weighted average of the last wage as an apprentice and the first wage in the first job. Using the wage of the second notification means that some selection is present in the data; some of the graduated apprentices have left the first job by then. This is the case for about 31 and 27 percent of the movers and stayers respectively. Secondly, for some apprentices the wage of the second notification is missing, as they did not receive a wage from their employer for that observational period (mostly due to mandatory military service). This is the case for about 2 percent of movers, and about 4 percent of stayers. Overall, the sample of valid observations is reduced from 16,281 to 11,199.

——— TABLE 4 ———

Table 4 presents the results for three different specifications. The first specification is a conventional model without the training intensity variables. Individual characteristics such as age, nationality, and Abitur each have a significant positive impact on the wage. Moreover, wages increase monotonously with the size of the training firm size. As the next model shows, much of the training firm size effect is due to the omission of training intensity as separate variables. The elasticity of post training wages with respect to training wages is 0.27. Similarly, there is a large premium for trainees with long training duration.

The inclusion of the training intensity variables also affects the stayer-mover differential which was insignificant in the first model. With training intensity variables included, there is a significant positive differential for movers, albeit of small magnitude (2.8 log points). The change confirms our initial conjecture that not controlling for training intensity will tend to lead to a downward bias in the mover-stayer differential (as the mover dummy is

negatively correlated with training intensity).

Why do movers command higher wages? 36.5 percent of all movers move to firms larger than their training firm; 29.0 percent of all movers move to firms smaller than their training firm. Thus, the direction of mobility is far from uniform. However, there is a slight dominance of "upward" mobility. Moving to a larger sized firm can be expected to be associated with a wage gain, as moving to a smaller firm leads to a wage loss. This is confirmed by the third column of Table 4 that shows that the premium of "moving up" is a 14.9 log point wage gain, while "moving down" is associated with a 10.7 log point wage loss. There is no mover stayer wage differential once we consider movers staying in the same size class.

5 Conclusion

The main goal of our analysis was to study the transition of apprenticeship trainees into their first job. Specifically, we studied the retention rates, the first job durations as well as initial post-apprenticeship wages. We argued that the IAB Employment sample is particularly well suited for this task, due to its accuracy, the long observation period (1975-95), and the large number of observations. Moreover, we emphasized the importance of controlling for training intensity, and showed that two relevant indicators of training intensity are available in the IAB data.

Using these data, we provided new evidence on two questions on which previous studies using alternative data sources have offered conflicting evidence. The first question was whether apprentices who are retained in the training firm receive higher or lower wages than apprentices who move to a new employer. Our finding is that there is no systematic wage differential between movers and stayers, once we control for training intensity and firm size. The second question was the relationship between apprenticeship training and the duration in the first job.

We find that retained graduated apprentices stay longer in their first job than apprenticeship graduates recruited by another firm. The difference, while statistically significant, may however be small in economic terms. Taken together, small differences in wages and

rather similar first job durations for movers and stayers tend to suggest that the ability of firms to extract rents from their trainees may be limited. What we cannot exclude, however, is that a generally compressed wage structure has some beneficial effects for training firms.

On the basis of the large mobility directly after graduation and the rather small differences in the first-job durations and wages we conclude that the knowledge obtained during apprenticeship training is quite transferable. This certainly is a benefit and strength of a system that provides post-secondary education for almost two thirds of youth in Germany.

References

Abbring and Van den Berg (1998) Limit Results for unobserved heterogeneity distributions in mixed proportional hazard models, Working Paper, Free University of Amsterdam.

Acemoglu, D., and J.S. Pischke (1998) Why Do Firms Train? Theory and Evidence, *Quarterly Journal of Economics*, 113(1), 79-119.

Acemoglu, D., and J.S. Pischke (1999a) Beyond Becker: Training in Imperfect Labor Markets, *Economic Journal*, 109, F112-F142.

Acemoglu, D., and J.S. Pischke (1999b) The Structure of Wages and Investment in General Training, *Journal of Political Economy*, 107, 539-572.

Bender, S. and H. Dietrich (1994) Occupational mobility in Germany - from apprenticeship to labor market integration, Nuremberg, working paper.

Bender, S., A. Haas, and C. Klose (2000) IAB Employment Subsample 1975- 1995. Opportunities for Analysis Provided by the Anonymised Subsample, IZA DP, No. 117.

Bougheas, S. and Y. Georgellis (2001) Apprenticeship training, earnings profiles and labor turnover: Theory and German evidence, paper presented at the 2001 EALE Conference.

Clark, D. (2000) How transferable is German apprenticeship training? Paper presented at the IZA Seminar on 14.11.2000.

Clark, D. (2001) Why do German firms subsidize apprenticeship training? Tests of the asymmetric information and mobility cost explanations, *Vierteljahreshefte zur Wirtschaftsforschung*, 1/2001.

Dietrich, H. (2000) Betriebliche Ausbildungsverhalten im Kontext der betrieblichen Altersstruktur, in: R. George and O. Struck (Eds.) *Generationenaustausch im Unternehmen*, Rainer Hampp Verlag: München, 159-176.

Dustmann, C., R. Euwals and A. van Soest (1997) Wages in the First Job after Apprenticeship: Movers versus Stayers, *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 30, 671-674.

Fougère, D. and W. Schwerdt (2001) What is the marginal productivity of apprentices? Paper presented at the 2001 CEPR/IZA Summer Symposium in Labor Economics, Ammersee.

Franz, W. J. Inkmann, W. Pohlmeier and V. Zimmermann (2000) Young and out in Germany: on the youth's chances of labor market entrance in Germany, in: D. Blanchflower and R. Freeman (eds.) *Youth employment and joblessness in advanced countries*, Chicago, University of Chicago Press, 381-426.

Franz, W., and V. Zimmermann (1999) Job mobility after apprenticeship training in West Germany, *Jahrbücher für Nationalökonomie und Statistik*, 219, (1+2).

Harhoff, D., and T.J. Kane (1997) Is the German Apprenticeship System a Panacea for the US Labor Market? *Journal of Population Economics*, 10, 171-196.

Lancaster. T. (1979) Econometric methods for the duration of unemployment, *Econometrica*, 47, 939-956.

Riphahn, R.T. (2002) Residential location and youth unemployment: the economic geography of school-to-work transitions, *Journal of Population Economics*, 15, 115-135

Steedman, H. (1993) The economics of youth training in Germany, *Economic Journal* 103, 1279-1291.

von Bardeleben, R., U. Beicht and K. Feher (1995) *Betriebliche Kosten und Nutzen der Ausbildung*, Berichte zur beruflichen Bildung No. 187: Bielefeld.

Werwatz, A. (1996) How firm-specific is German apprenticeship training? Discussion Paper No. 96-12, SFB 373, HU Berlin,

Werwatz, A. (1997) Mobility after Apprenticeship - How effective is the German apprenticeship system? Discussion Paper No. 97-55, SFB 373, HU Berlin

Winkelmann, R. (1996a) Employment Prospects and Skill Acquisition of Apprenticeship-Trained Workers in Germany, *Industrial and Labor Relations Review*, 49(4), 658-672.

Winkelmann, R. (1996b) Training, Earnings, and Mobility in Germany, *Konjunkturpolitik*, 42, 275-298.

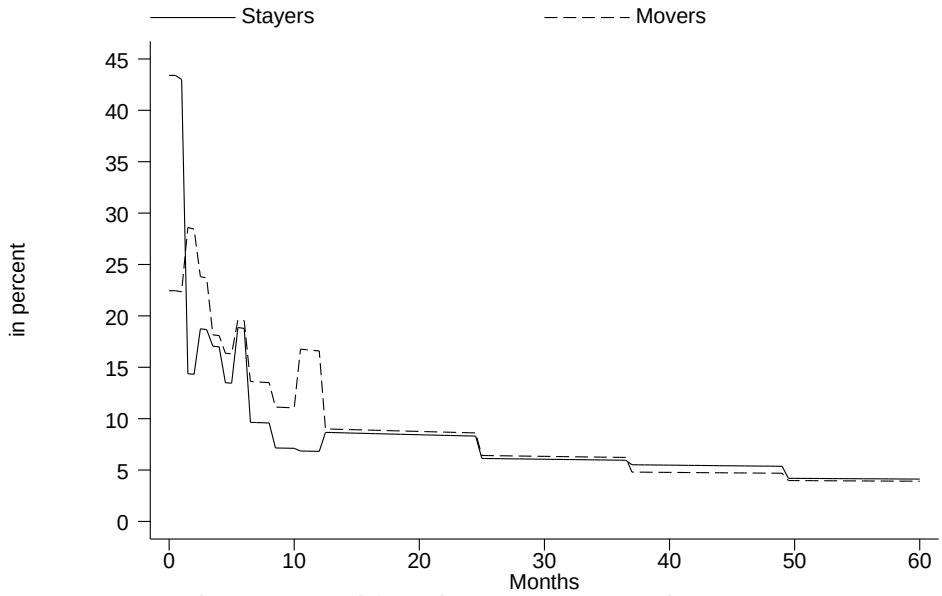

Fig 1. Hazard functions, movers and stayers

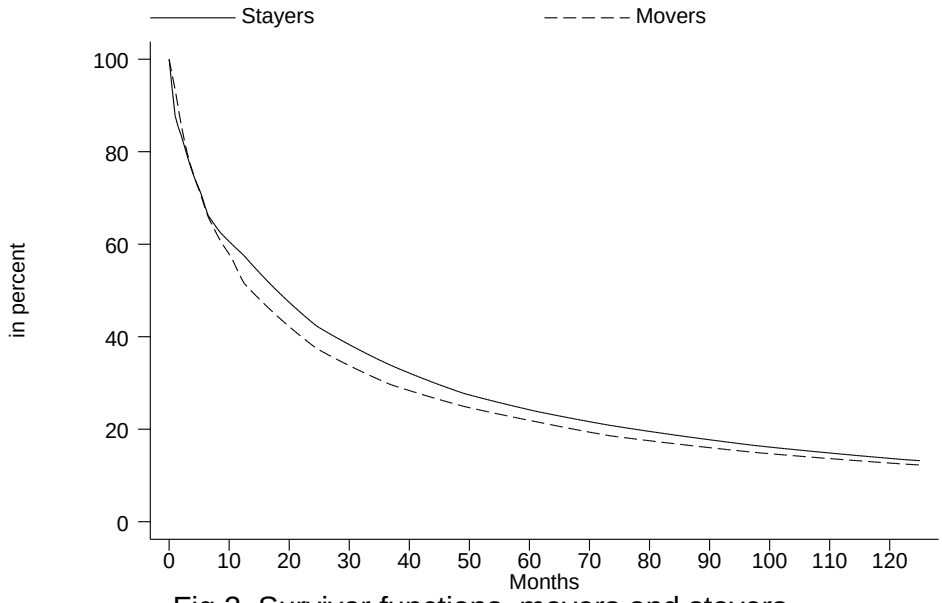

Fig 2. Survivor functions, movers and stayers

Birth year	1960	1961	1962	1963	1964	1965	Total
Demographics:							
German	0.984	0.987	0.979	0.977	0.969	0.970	0.978
Age at start of AT	16.5	16.9	17.0	17.0	17.1	17.2	17.0
Married	0.034	0.034	0.024	0.026	0.025	0.022	0.027
Abitur	0.030	0.032	0.042	0.050	0.049	0.061	0.044
Characterization of AT:							
Daily wage	19.6	20.2	20.8	21.3	21.6	21.5	20.8
Short duration	0.301	0.242	0.257	0.254	0.249	0.225	0.255
Medium duration	0.496	0.533	0.530	0.540	0.532	0.545	0.529
Long duration	0.203	0.266	0.213	0.206	0.219	0.230	0.216
1-9 employees	0.280	0.247	0.243	0.260	0.243	0.248	0.254
10-49 employees	0.283	0.304	0.305	0.292	0.299	0.280	0.294
50-499 employees	0.252	0.255	0.255	0.260	0.260	0.271	0.259
500+ employees	0.186	0.194	0.197	0.188	0.198	0.201	0.194
Employment transition:							
Stay with training firm	0.754	0.737	0.706	0.697	0.675	0.691	0.710
Unemployment	0.082	0.112	0.170	0.192	0.219	0.191	0.162
Further AT	0.024	0.022	0.025	0.027	0.020	0.023	0.023
Not in labor force	0.214	0.222	0.239	0.250	0.265	0.254	0.241
Number of observations	2790	2623	2511	2725	2857	2775	16281

Table 1: *Descriptive statistics*. Source: IAB sample, own calculations.

	(1)	(2)	(3)
Intercept	0.6577	-0.0023	0.0520
	(0.0074)	(0.3473)	(0.3468)
German national		-0.0846	-0.0681
		(0.1111)	(0.1096)
Age		-0.0313	-0.0229
		(0.0341)	(0.0339)
Age^2		0.0004	0.0002
		(0.0008)	(0.0008)
Married		-0.0016	-0.0097
		(0.0370)	(0.0366)
Abitur		0.0538***	0.0178
		(0.0195)	(0.0203)
Ln(training wage)		0.1952***	0.1440***
		(0.0122)	(0.0157)
Medium AT duration		0.1267***	0.1235***
		(0.0093)	(0.0093)
Long AT duration		0.1618***	0.1547***
		(0.0117)	(0.0125)
10-49 employees	0.0618***	0.0426***	0.0555***
	(0.0098)	(0.0098)	(0.0101)
50-499 employees	0.0630***	0.0189*	0.0395***
	(0.0101)	(0.0106)	(0.0117)
500 + employees	0.0903***	0.0185*	0.0644***
	(0.0107)	(0.0116)	(0.0136)
Year dummies (19)	no	no	yes***
Occupation dummies (47)	no	no	yes***
R-squared	0.0050	0.0404	0.0698

Table 2: *Linear probability estimates for stayer ($y = 1$) / mover ($y = 0$) model. Heteroskedasticity consistent standard errors in parentheses ($N = 16281$).*

	(1)	(2)	(3)
German national	-0.0311	-0.0465	-0.0145
	(0.1543)	(0.1566)	(0.1545)
Age	0.0630	0.0561	0.0676
	(0.1769)	(0.1779)	(0.1776)
Age^2	-0.0012	-0.0009	-0.0011
	(0.0041)	(0.0041)	(0.0041)
Married	-0.2389	-0.2414	-0.2396
	(0.1506)*	(0.1500)*	(0.1492)*
Abitur	-0.4198	-0.4249	-0.4105
	(0.2355)*	(0.2308)*	(0.2331)*
Ln(training wage)		-0.3177	-0.3338
		(0.0990)***	(0.0977)***
Medium AT duration		-0.1070	-0.1064
		(0.0574)*	(0.0566)*
Long AT duration		-0.2085	-0.2069
		(0.0759)***	(0.0748)***
10-49 employees	-0.0048	-0.0103	-0.0441
	(0.0605)	(0.0605)	(0.0613)
50-499 employees	-0.1312	-0.0876	-0.1459
	(0.0691)*	(0.0703)	(0.0728)**
500 + employees	-0.4416	-0.3696	-0.4365
	(0.0880)***	(0.0903)***	(0.0939)***
Move to smaller firm			0.2045
			(0.1039)**
Move to larger firm			-0.3027
			(0.1009)***
$\hat{\sigma}^2$	0.1401	0.1394	0.1298
	(0.1366)	(0.1301)	(0.1163)
Log-likelihood	-12773.5	-12763.6	-12742.8

Table 3: *Results for proportional hazard model for first job duration (for baseline hazard and survivor rates of movers and stayers, see figures 1 and 2). Other exogenous variables include dummy variables for year and occupation (N=16281).*

	(1)	(2)	(3)
Intercept	1.8189	1.5795	1.6364
	(0.2671)***	(0.2688)***	(0.2625)***
German national	0.1223	0.1075	0.0723
	(0.0631)*	(0.0595)*	(0.0505)
Age	0.1931	0.1408	0.1339
	(0.0246)***	(0.0244)***	(0.0238)***
Age^2	-0.0036	-0.0026	-0.0024
	(0.0006)***	(0.0006)***	(0.0005)***
Married	0.0243	0.0217	0.0192
	(0.0183)	(0.0181)	(0.0177)
Abitur	0.2287	0.2488	0.2334
	(0.0277)***	(0.0279)***	(0.0269)***
Ln(training wage)		0.2727	0.2660
		(0.0142)***	(0.0140)***
Medium AT duration		0.1234	0.1217
		(0.0089)***	(0.0088)***
Long AT duration		0.1869	0.1824
		(0.0111)***	(0.0109)***
10-49 employees	0.0263	0.0122	0.0401
	(0.0093)***	(0.0091)	(0.0090)
50-499 employees	0.1066	0.0627	0.1082
	(0.0099)***	(0.0100)***	(0.0100)***
500 + employees	0.1786	0.1106	0.1676
	(0.0115)***	(0.0117)***	(0.0117)***
Mover	0.0016	0.0276	0.0076
	(0.0078)	(0.0076)***	(0.0111)
Move to smaller firm			-0.1065
			(0.0110)***
Move to larger firm			0.1485
			(0.0143)***
R-squared	0.2410	0.2837	0.3049

Table 4: *Linear regression for the logarithm of wages in the first job. Robust standard errors in parentheses. Other exogenous variables include dummy variables for year and occupation (N=11199).*

Working Papers of the Socioeconomic Institute at the University of Zurich

The Working Papers of the Socioeconomic Institute can be downloaded from http://www.soi.unizh.ch/research/wp/index2.html

0308 Training Intensity and First Labor Market Outcomes of Apprenticeship Graduates
Rob Euwals and Rainer Winkelmann, September 2003, 25 p.

0307 Co-payments for prescription drugs and the demand for doctor visits – Evidence from a natural experiment
Rainer Winkelmann, September 2003, 22 p.

0306 Who Integrates?
Stefan Buehler and Armin Schmutzler, August 2003, 32 p.

0305 Strategic Outsourcing Revisited
Stefan Buehler and Justus Haucap, July 2003, 22 p.

0304 What does it take to sell Environmental Policy? An Empirical Analysis for Switzerland
Daniel Halbheer, Sarah Niggli and Armin Schmutzler, 2003, 30 p.

0303 Mobile Number Portability
Stefan Buehler and Justus Haucap, 2003, 12 p.

0302 Multiple Losses, Ex-Ante Moral Hazard, and the Non-Optimality of the Standard Insurance Contract
Michael Breuer, 2003, 18 p.

0301 Lobbying against Environmental Regulation vs. Lobbying for Loopholes
Andreas Polk and Armin Schmutzler, 2003, 37 p.

0214 A Product Market Theory of Worker Training
Hans Gersbach and Armin Schmutzler, 2002, 34 p.

0213 Weddings with Uncertain Prospects – Mergers under Asymmetric Information
Thomas Borek, Stefan Buehler and Armin Schmutzler, 2002, 37 p.

0212 Estimating Vertical Foreclosure in U.S. Gasoline Supply
Zava Aydemir and Stefan Buehler, 2002, 42 p.

0211 How much Internalization of Nuclear RiskThrough Liability Insurance?
Yves Schneider and Peter Zweifel, 2002, 18 p.

0210 Health Care Reform and the Number of Doctor Visits ? An Econometric Analysis
Rainer Winkelmann, 2002, 32p.

0209 Infrastructure Quality in Deregulated Industries: Is there an Underinvestment Problem?
Stefan Buehler, Armin Schmutzler and Men-Andri Benz, 2002, 24 p.

0208 Acquisitions versus Entry: The Evolution of Concentration
Zava Aydemir and Armin Schmutzler, 2002, 35 p.

0207 Subjektive Daten in der empirischen Wirtschaftsforschung: Probleme und Perspektiven.
Rainer Winkelmann, 2002, 25 p.

0206 How Special Interests Shape Policy - A Survey
Andreas Polk, 2002, 63 p.

0205 Lobbying Activities of Multinational Firms
Andreas Polk, 2002, 32 p.

0204 Subjective Well-being and the Family
Rainer Winkelmann, 2002, 18 p.

0203 Work and health in Switzerland: Immigrants and Natives
Rainer Winkelmann, 2002, 27 p.

0202 Why do firms recruit internationally? Results from the IZA International Employer Survey 2000
Rainer Winkelmann, 2002, 25 p.

0201 Multilateral Agreement On Investments (MAI) - A Critical Assessment From An Industrial Economics Point Of View
Andreas Polk, 2002, 25 p.