The Scandinavian
Journal of Economics

Scand. J. of Economics 110(4), 733–753, 2008
DOI: 10.1111/j.1467-9442.2008.00559.x

Returns to Apprenticeship Training in Austria: Evidence from Failed Firms*

Josef Fersterer

Landesstatistischer Dienst, A-5010 Salzburg, Austria
josef.fersterer@salzburg.gv.at

Jörn-Steffen Pischke

London School of Economics, London WC2A 2AE, England
s.pischke@lse.ac.uk

Rudolf Winter-Ebmer

University of Linz, A-4040 Linz, Austria
rudolf.winterebmer@jku.at

Abstract

In the German-speaking countries, little is known about the payoffs to apprenticeship training for the participants. OLS estimates suggest that the returns are similar to those of other types of schooling. However, there is considerable heterogeneity in the quality of apprenticeships offered, and institutional descriptions suggest that there might be an important element of selection in who obtains an apprenticeship, and what type. In order to overcome the resulting ability bias, we estimate returns to apprenticeship training for apprentices in small Austrian firms which cease to operate. When a firm fails, current apprentices cannot complete their training in this firm. Since apprentices will be at different stages in their apprenticeship at that time, the failure of a firm will manipulate the length of the apprenticeship period completed for some apprentices. The time to firm failure can therefore serve as an instrument for the length of the apprenticeship completed both at the original firm and at other firms. We find instrumental variables returns which are similar or larger than the OLS returns in our sample, indicating relatively little selection.

Keywords: Human capital; returns to schooling; firm-based training; ability bias

JEL classification: $J24$; $J31$

I. Introduction

There is substantial interest in the functioning of the apprenticeship training system in German-speaking countries, which is often regarded as a potential model for other countries; see Heckman (1993) and Steedman (2001). The

* We thank two anonymous referees, Josh Angrist, Larry Katz, and participants at various seminars for helpful comments. We thank the Economic and Social Research Council (RES 000-22-01-0160) and the Austrian Science Fund (FWF-P-15422-G05) for financial support, and Pischke thanks the NBER for their hospitality.

importance of skills is increasing even at the bottom end of the skill distribution, and apprenticeships are seen as one way to create high skill levels for non-college-bound youths; see Nickell and Bell (1996) and Freeman and Schettkat (2001) for evidence of comparatively high skill levels at the bottom end of the German skill distribution. The apprenticeship systems in Germany, Austria and Switzerland are also believed to deliver relatively low youth unemployment rates because they facilitate an orderly school-to-work transition; see Büchtemann, Schupp and Soloff (1994), OECD (1998) and Ryan (2001).

The apprenticeship system is regarded as exemplary not only because of its role in the overall economy but also because it delivers desirable results for those who participate: a well-paying, interesting job, adequate job security, and possibilities for advancement. This suggests that the private returns to such training must be high. While the benefits of the apprenticeship system are often alleged, there are few existing studies of the returns to apprenticeship training. Krueger and Pischke (1995) and Winkelmann (1996) present OLS estimates of the returns to apprenticeship training in Germany in the order of 15–20 percent, and Fersterer and Winter-Ebmer (2003a) find returns of 15 percent for Austria—as return to a training period of three to four years.

Estimating the returns to apprenticeship training is not straightforward. Soskice (1994) argues that there is a lot of heterogeneity in the quality of apprenticeships, and firms select apprenticeship candidates extensively among secondary school leavers. This selection would bias any OLS estimates of the returns to training but there is virtually no empirical evidence on the degree of selection. Much of the heterogeneity identified by Soskice is related to firm size, and the wages for apprenticeship-trained workers indeed increases strongly in the size of the training firm; see, for example, Acemoglu and Pischke (1998), or Franz, Inkmann, Pohlmeier and Zimmermann (2000) for results for Germany.

In this paper, we try to obtain estimates of the return to apprenticeship training which are free from this selection bias. In order to overcome the selection, we focus on failed firms. When a firm fails, current apprentices cannot complete their training in this firm. Since apprentices will be at different stages in their apprenticeship at that time, the failure of a firm will manipulate the length of the apprenticeship period completed for some apprentices, as long as not all of them complete the apprenticeship elsewhere after their firm fails. The time from entering a failing firm to the actual exit of the firm therefore serves as an instrument for the length of the apprenticeship completed both at the original firm and at other firms.

The empirical work is based on the universe of social security records for Austria from 1975 to 1998. This allows us to identify a large number of apprentices in failing firms. The comparisons we make are limited to workers who started an apprenticeship in failing firms, because these firms will be different from firms that continue. Moreover, all the apprentices analyzed would have to find new employment during or soon after their apprenticeship. We analyze the wages of these workers once they have taken a post-apprenticeship job.

Our analysis of the data suggests that apprentices are typically hired every year until the eventual closure of the firm. The maintained assumption in the identification of the returns is therefore that apprentices hired about one to four years before a firm fails will be very similar. We believe that this assumption is most defensible in firms where the failure is rather sudden. We therefore limit the analysis to small firms. We also exclude firms where much of the decline in employment has already occurred before the eventual firm exit. Controlling for age rather than experience, we find instrumental variables estimates of the returns to training between 2.5 and 4 percent per year, which are similar or above our OLS estimates. It is important to keep in mind that our sample is a fairly selected sub-sample of training firms. Our main conclusion is that there seems to be little selection bias in the OLS estimates of the returns to training.

No previous instrumental variables (IV) studies of the returns to apprenticeship training exist so far. A specific instrument typically captures variation in a particular range of the endogenous variable. The only existing IV studies of the returns to schooling for Austria by Fersterer and Winter-Ebmer (2003b) and Ichino and Winter-Ebmer (2004) focus on secondary schooling. In Austria, differences in secondary schooling eventually tend to imply that an individual obtains either lower secondary schooling plus an apprenticeship or other vocational training, or a higher secondary degree plus academic education. This is not the relevant return for the marginal apprentice (i.e., someone at the borderline between undertaking an apprenticeship or no post-junior-high-school training at all). The same is true for IV studies of the returns to schooling in Germany by Ichino and Winter-Ebmer (1999, 2004) and Becker and Siebern-Thomas (2004).

The remainder of this paper is organized as follows. In the next section we describe the Austrian schooling and training system. Section III outlines our estimation strategy and Section IV our dataset. Section V reports the results. Section VI offers some concluding remarks.

II. The Austrian Schooling and Training System

The Austrian school system tends to be relatively complex, with a multitude of different educational routes at a particular age. Children start school at age six and attend a four-year primary school. After grade four, the school system tracks students into two types of secondary schools. Academic secondary schools lead to a university entrance exam (*Matura*) after grade 12. School leavers from this type of school typically enter universities, polytechnics (*Fachhochschule*) or other tertiary academies. General secondary school (*Hauptschule*) lasts for four years, and offers an academically less challenging curriculum than academic secondary schools. The school is supposed to prepare students for further vocational training at the end of the compulsory schooling age. Students who intend to pursue an apprenticeship supplement it with a one-year pre-vocational school before leaving full-time schooling at age 15. Alternatively, students can enter a full-time vocational school, either an intermediate vocational school, which intends to prepare students primarily for a profession, or an upper vocational school, which supplements training for a profession with preparation for the university entrance exam.

Selection into the different types of secondary school depends on a combination of primary school grades and recommendations by the primary school, a formal entrance exam and parental choice. There are also rules based on grades and exams, which allow students to transit from lower secondary schools to the academic secondary schools.

About 40 percent of a cohort complete an apprenticeship, mostly after leaving general secondary school. Twenty-three percent leave school at 15 without any formal qualifications (*Hauptschulabschluß*) and start work as an unqualified laborer. Apprentices receive training in a particular occupation in a firm. In addition, the apprentice attends a part-time vocational school (*Berufschule*), either one or two days a week or in block courses at certain times during the year. The firm provides both the practical component of the training as well as on-the-job experience. The vocational school supplements training by more theoretical components tailored to the chosen occupation, as well as providing further liberal arts education. In addition to learning at the firm and in vocational school, some training firms organize joint courses or send their apprentices to external training centers for certain aspects of the training. This happens more frequently in small firms, for example, because they lack certain machines, which are supposed to be part of the training curriculum.

The apprenticeship training is highly regulated. There are currently 275 apprenticeship occupations, with nationally legislated curricula. Each training firm has to satisfy certain requirements in order to be allowed to take on apprentices. In particular, there has to be a qualified trainer at the firm. The training firms are overseen by the Apprenticeship Offices of the regional chambers of commerce. They also administer the graduation exams for the apprentices. Apprentices obtain an allowance, negotiated in union contracts, which differs by occupation. The remuneration of apprentices starts out very low in the first year, and may rise up to about 80 percent of a skilled worker's wage in the last year.

Apprenticeships are not at all homogeneous, and the content and quality of training varies greatly between occupations and training firms. Apprenticeships last between two and four years. Durations of two and two-and-a-half years are rare (1 percent of all apprenticeships), while durations of three (62 percent) and three-and-a-half years (26 percent) are the most common; see Schneeberger and Kastenhuber (1996, Table 16). The duration of the apprenticeship may be reduced if the apprentice has prior relevant skills or a higher level of schooling, like a *Matura*. About 10 percent of apprentices pursue combination apprenticeships in two related occupations at the same time (*Doppellehren*). Where the content of the trades overlaps substantially, these combination apprenticeships can be completed in a much shorter time span than two individual apprenticeships. Examples are vehicle mechanic/vehicle electrician, plumber for water and gas/plumber for central heating systems, or cook/waiter. Eleven percent of apprenticeship contracts last four years, and almost all of these are combination apprenticeships.

The same occupation (say, vehicle mechanic) may be learnt either in a small crafts firm, or in a large industrial enterprise. In the small firm, training is largely on-the-job, and apprentices are well integrated into the business activity of the firm. In the large firm, the first year of training may be spent entirely in an internal training workshop with full-time trainers. Afterwards, apprentices may rotate to different types of jobs within the firm.

Apprentices can obtain additional credentials after accumulating some experience in their chosen trade. Additional training and exams may lead to the master craftsman credential (*Meister*) or an equivalent qualification, which allows the degree holder to train apprentices and is often the prerequisite for self-employment in a certain profession.

The vocational training system in Austria encompasses a wide variety of occupations. An apprenticeship may be in a traditional crafts profession (baker, cook, mechanic, hairdresser), in administrative and clerical occupations, sales or in technical professions (machinery electrician, chemical laboratory technician). Vocational training may lead to a qualification, which would be considered a fairly low-skilled occupation in most countries (waiter or sales clerk). But it may also lead to an occupation, which would require at least some college education in other countries, particularly when it combines an apprenticeship with more advanced school-based training (technician or middle manager).

Not all apprentices complete their apprenticeship. Some apprentices leave their apprenticeship early for a variety of reasons. Many of them continue their apprenticeship in another firm and/or occupation. About 18 percent of apprentices dropped out of the apprenticeship system completely in the 1980s. About one-fifth of apprentices had two apprenticeship contracts, and one-tenth had three or more; see Schneeberger *et al.* (no date).

Apprentices may also leave an apprenticeship without the resulting skilled-worker credential because they fail the final examination. About 85 percent of exam entrants pass. The exam can be repeated after three months if the candidate was unsuccessful. Roughly 70 percent of failed candidates attempt the exam again within a year and about 75 percent of repeaters pass; see Schneeberger *et al*. (no date). This suggests that more than 90 percent of apprentices pass the final exam eventually.

III. Empirical Framework

The following is a simple framework for thinking about the behavior of apprentices. Let L^* be the latent apprenticeship duration for a particular apprentice. An apprentice who does not train at a failing firm will realize that duration, so that the actual duration of the apprenticeship is $L = L^*$. This may be completed in a single firm, or the apprentice may quit the original training firm and complete the apprenticeship elsewhere. Variation in L^* combines the fact that different apprenticeship occupations require a different duration, as well as the fact that some apprentices quit early. L^* is likely related to the ability and motivation of an apprentice, apart from the fact that individuals with different levels of L^* will have accumulated different amounts of human capital.

We only consider failures in the first firm an apprentice joins. Let K be the time from the entry of the apprentice into this firm until the firm fails. This is not necessarily the same as the time an apprentice spends in the failing firm. For example, the apprentice may leave the firm early, either because $L^* < K$, or because the apprentice decides to complete the apprenticeship in another firm. The random variable K will be independent of L^* as long as the distribution of L^* does not vary among the apprentices who join the failing firm at different times before the failure. In other words, the last apprentices to join a failing firm have to have similar (observable and unobservable) characteristics as those who joined earlier. This is our key identifying assumption, and we discuss the consequences of its violation below. It is important to note that the actual time spent by an apprentice in the failing firm is not independent of L^*. In order to see this, note that apprentices who actually stay in the failing firm for time K will by definition have $L^* \geq K$, so that these apprentices now come from a truncated distribution of L^*.

Apprentices in failed firms will realize a total apprenticeship duration $L = L(K)$. The realized length of the apprenticeship for an apprentice in a failed firm may be equal to L^*, it may be shorter, or it may be longer. An apprentice in a failing firm may leave and complete the apprenticeship elsewhere in the original time. However, it is plausible that apprentices, who leave their original training firm, may take longer than initially anticipated to complete their training, maybe because the change of firms necessitates a change in occupation. Alternatively, an apprentice whose firm fails may decide not to complete his training or simply might not find a suitable training firm, although the apprentice would have completed it otherwise, i.e., $L(K) < L^*$.

An important assumption of our approach is that $E(L(K)) < E(L^*)$, i.e., being subject to the failure of the training firm on average reduces the completed training of the affected apprentices. This is similar to the existence of a first stage when estimating a system of equations by two-stage least squares, and it can easily be verified in our data. If all or most apprentices affected by firm failures took as long or longer to complete their apprenticeship as did those not affected, then there would be no useful variation in $L(K)$ for us to exploit. This assumption can be checked in our data since we observe $L = L^*$ among those apprentices whose firm fails after they completed their apprenticeship.

We therefore want to estimate the relationship

$$w_{it} = \alpha + \beta L_i + x_{it}'\gamma + \varepsilon_{it}, \tag{1}$$

where w_{it} is the log of the post-apprenticeship wage for individual i in year t, x_{it} are other covariates like age, and ε_{it} is an error term. The standard problem is that L_i is related to L_i^*, which may be correlated with the error term. Hence, our approach is to instrument L_i by K_i in the sample of apprentices for whom K_i is four years or less, i.e., who are potentially affected by the failure of their training firm. It is important to point out that the controls in x_{it} include age rather than experience. Conditional on age, a change in L_i affects both apprenticeship duration and experience. Hence, we only estimate their net effect. This is a standard problem in IV studies of the returns to schooling where it is common to control for age instead of experience; see e.g. Angrist and Krueger (1991).

If the returns to training differ by individual and/or by the stage of training, instrumental variables estimation will identify a local average treatment effect, as in Imbens and Angrist (1994), Angrist and Imbens (1995), and Angrist, Imbens and Rubin (1996), for the individuals whose behavior is affected by the firm failure (this group is often labeled "compliers"). For this interpretation to be correct, it is necessary that a monotonicity assumption holds. In our case, this means that a firm failure will not induce any individual to stay in the apprenticeship longer than they would have otherwise, or $L_i \leq L_i^*$. We cannot rule out the failure of monotonicity in our case *a priori*, but we discuss below whether there is evidence in our data that some individuals affected by firm failures have longer apprenticeship durations than the control group. We find that the fraction of individuals with particularly long durations is not higher among those affected by firm failures. This suggests that the monotonicity assumption seems to be satisfied in our data.

The group of compliers in our case, i.e., the apprentices for whom $L_i(K_i) < L_i^*$, are likely to be individuals who are closer to the margin of dropping out of the apprenticeship anyway. In addition, we are not able to measure the return to the entire apprenticeship period, since everybody in the sample will have completed some part of training in the failing firm. Hence, we measure the return to partial apprenticeships, and in particular the later years of the training period. There may be important non-linearities in the returns to apprenticeship training, for example due to credential effects of receiving the skilled-worker certificate. In estimating equation (1) by instrumental variables we may mistakenly attribute these returns to the part of the apprenticeship period not completed by compliers.[1]

The failure of a firm is not an event that will be completely unanticipated. It is well known from other studies that some workers tend to leave firms before the eventual failure; see Hamermesh and Pfann (2001). Some apprentices may leave their training firm because they expect the eventual failure of the firm, and the early leavers may not be selected from all apprentices at random.[2] This is not directly a problem for our approach. This behavior may affect who is in the group of compliers, but it does not affect the interpretation of the results for that group.

Note that the variation in K comes from the fact that some apprentices joined a failing firm just before the failure, while others joined the firm at an earlier time. As indicated above, our key identifying assumption is that apprentices who joined within the last four years before the failure of the firm are comparable. For example, we would overestimate the returns to training if the quality of the apprentices, who are hired before the firm fails, is declining over time because K is positively correlated with worker quality in this case. In order to avoid this, we focus our analysis on a sub-sample of small training firms, that fail suddenly, and on apprentices who joined these firms one year before the failure or earlier. We expect our identifying assumption to be satisfied most easily in this sub-sample.

Another issue arises in estimating equation (1). We have seen that L_i can be less than, equal to, or greater than the statutory length for an apprenticeship in the chosen occupation. It is common in the literature on returns to schooling to use the statutory length of an educational program. Typically, the endogenous regressor will be a variable like "highest grade completed".

[1] Unfortunately, we lack information on receipt of the actual apprenticeship credential, so we are not able to address this problem directly. In addition, our first stage is not powerful enough to use the instrument to separately identify a linear return to apprenticeship length and a credential effect.

[2] Such voluntary quits of apprentices may be very unlikely, unlike those of regular workers, because apprentices enjoy very strict dismissal protection in their training contract.

If a student repeats a grade, and hence spends additional time in school to reach the same grade level, this additional time is not counted. Therefore, it seems sensible for our purposes not to count time spent beyond the statutory length of the apprenticeship. Unfortunately, we do not know what the statutory length is for a particular apprentice. Hence, we truncate L_i at four years, the maximum length of an apprenticeship, and use the variable $\tilde{L}_i = \min(L_i, 4)$ instead.[3]

IV. The Data

We use employment records from the Austrian Social Security Administration. The dataset includes the universe of private sector workers in Austria, including apprentices, covered by the social security system. It excludes public sector workers and the self-employed. This is a drawback, since about 10 percent of apprenticeship-trained workers are self-employed, and self-employment is frequently a successful outcome for an apprentice. For each employment relationship, beginning and end dates are recorded and the establishment in which the worker is employed can be identified. The dataset covers the years 1972–1998 but we only use records from 1975 onwards, because the employer-level information is frequently inaccurate in the earlier years.

Each establishment has an employer social security number. Hence, an exit of an establishment in the data occurs when the employer identifier ceases to exist. While it is possible in principle to build a matched employer–employee panel at a daily frequency, this is neither computationally feasible nor all that valuable for our purpose. To define firm failures, we record the employment status and firm affiliation for each individual four times a year, on February 10, May 20, August 10 and November 10. We therefore attribute a firm exit to the last quarter before we no longer observe individuals with the firm identifier (and the firm does not reappear again later, for example, because it is in a seasonal industry). Hence, firm exits are defined at a quarterly frequency in our analysis. However, some of these cases may not be true firm exits, and (most of) the employees continue under a new identifier. We use a common strategy in this literature; as in e.g. Bender, Dustmann, Margolis and Meghir (2003) for France, to concentrate our analysis only on cases where we can be sure about firm exit status. If more than 70 percent of the employees continue under a new employer identification number we do not consider this a failure of the establishment. In some cases, the employer ID may refer to the employees in multiple sites; in other cases, multiple plants in a single firm may have

[3] Results do not change if we use the original durations.

separate numbers. This is not a particular problem for us, but it may weaken our first-stage relationship.

Furthermore, we limit our analysis to the sub-sample of failing firms where the loss in employment in the six months before the quarter of the failure is less than 25 percent. It is important for our identification strategy to analyze firms that exit suddenly, rather than those which have a protracted process of failing. In particular, firms which shed workers for a while before the eventual failure will likely have different hiring practices than those which fail rather suddenly. Hence, we are more confident that apprentices hired relatively close to the failure of the firm are similar to those hired earlier in this more limited sample.

We also limit the sample to small training firms with less than 10 employees. This is an interesting sub-sector, as these firms cover most of apprenticeship positions in the crafts sector. Our instrumentation strategy of using failed firms does not explain eventual apprenticeship durations very well in the case of apprentices from large firms; this may be due to concerted actions of the local chamber of commerce to place apprentices from big failed firms.

We identified firms that fail according to the definition above. We then began by recording all workers who received some apprenticeship training in the dataset, and who worked in one of these failing firms during some time in their career. For each worker, we recorded the exact number of days spent in apprenticeship training. Where this training takes place in multiple firms we summed the durations over all firms. Unfortunately, we only know whether a particular employment relationship is an apprenticeship but not whether training was successfully completed with a formal certificate. For each individual, we can then track at what time the failure occurred after the worker first entered the firm (this is our instrument K). We only retained those individuals whose firm fails within 16 quarters from the original entry date. This should include all apprentices who spent up to four years in the failing firm, and therefore up to the maximum length of an apprenticeship. Finally, we further restricted our estimation sample to those individuals hired within 4–15 quarters of failure. The reason is that the very last hires of the firm before it exits may be of different quality; we discuss this further below.

We only consider men in the analysis. In the social security administration records, monthly gross earnings are measured in May each year. Since some individuals are not employed for the full month, we standardize these earnings to give a daily wage for a typical day in employment. We neglect the problem of top-coding of earnings due to social security regulations, because it is largely irrelevant for apprenticeship-trained workers in their first years on the job.

V. Results

Our results exploit the fact that some apprentices do not complete their apprenticeship if it is cut short by the failure of their firm. Failing firms are not a random sample of all firms. Table 1 compares the population characteristics of apprentices and their training firms with various samples. Column (4) shows the full population of apprentices. Comparing columns (2) and (3) in the table to column (4) demonstrates that firms which fail are smaller than the average training firm. Firm sizes for apprentices who are potentially affected by the firm failure (those who joined the failing firms in the last 15 quarters before failure, column (2)) are not very different from those for apprentices who work in firms, which fail after their apprenticeship (column (3)). This is comforting, since it suggests that there is not much selection on the firm side on who hires apprentices until close to the time of failure. The biggest difference in the firm size distribution—as well as in other characteristics—is between apprentices who train in ever-failing firms, and all apprentices. Only 24 percent of all apprentices train in firms with less than 10 employees compared to 40 percent in column (3). Other characteristics of the apprentices and apprenticeship firms are not strongly affected by our sample selection.

Table 2 presents evidence on the relationship between the potential maximum duration of an apprentice in the failing firm, K, and the eventual length of the apprenticeship, L. Apprentices in our sample are grouped by the potential duration K, which is broken up into quarters. Each quarter of K is a column in the table. In addition, means for everybody unaffected by the failure of their firm are summarized in the last column, labeled 16+. The first five rows of the table tabulate the distribution of the length of the apprenticeship L, also in quarters, into five categories: $L < K$, $L = K$, $K < L < 11$ quarters, $L = 11$–13 quarters, L is 14 quarters or longer. The second category, $L = K$, also includes all apprentices who left the firm within one quarter before failure and did not continue elsewhere. Exits from a failing firm rarely occur all at once, even in relatively small firms, and there seemed to be some piling up of individuals also in the category $L = K - 1$. The third category, $K < L < 11$ quarters, should contain incomplete apprenticeships, but those that are longer than the duration in the failed firm. The fourth and fifth categories most likely contain completed apprenticeships. In fact, most apprenticeships in our data ended after 11–13 quarters, and this coincides with the most common contractual durations. Some apprenticeships last longer, or apprentices might have done combination apprenticeships. Hence, individuals in the fifth category (14 or 15 quarters) may have legitimately taken as long to complete an apprenticeship, or they could have taken longer to complete their apprenticeship, for example because they had to switch training firms.

Table 1. *Sample characteristics (in percent)*

	Sample			
	Estimation sample (1)	Joined failing firm within 15 quarters of failure (2)	Joined failing firm 16 or more quarters before failure (3)	All apprentices (4)
Age at beginning of apprenticeship				
15 years and younger	55.1	57.2	59.8	59.7
16 years and older	44.9	42.8	40.2	40.3
Residence				
Town	10.9	9.9	9.7	12.5
Rural	81.9	76.9	71.2	80.2
Information missing	7.2	13.2	19.1	7.3
Nationality				
Austrian	64.7	67.1	65.6	56.1
Other	3.4	3.4	1.6	2.3
Information missing	31.9	29.5	32.8	41.5
Firm size (beginning of apprenticeship)				
0–9 employees	100.0	42.4	40.2	24.4
10–19 employees		18.5	18.7	18.5
20–99 employees		26.6	26.1	29.8
100 and more		12.5	15.0	27.3
Industry classification				
Agriculture, mining, energy and water	4.2	1.3	0.9	2.2
Manufacturing	44.6	36.5	34.0	35.2
Construction	18.2	24.9	22.7	24.8
Wholesale and retail trade	14.8	13.0	14.3	20.1
Hotel and restaurant	7.9	9.4	7.3	6.6
Other private and public services	3.7	3.1	2.9	4.9
Information missing	6.5	11.8	17.9	6.2
Sample size	5,661	23,893	103,201	676,368

Notes: Estimation sample consists of apprentices who joined failing firms in quarters 4–15 before failure, and employment at the firm declined by less than 25 percent in the last two quarters before the firm exits. Sample in column (2) includes all apprentices who joined a failing firm within 15 quarters of failure. Sample in column (3) includes apprentices who joined a failing firm 16 or more quarters before failure. Sample in column (4) includes apprentices in all firms without restriction.

In considering Table 2, start by looking at the distribution for those apprentices who joined a failing firm 16 or more quarters before failure. These apprentices were not affected by the firm failure during their apprenticeship (and are not used in our analysis below). Fifty-nine percent of them completed their apprenticeship within the standard time (11–13 quarters), and another 27 percent took longer than that. About 14 percent dropped out early, and most likely did not complete their apprenticeship.[4]

[4] It has to be taken into account that we do not observe formal completion of the apprenticeship contract which requires successful completion of an exam.

Table 2. *Length of apprenticeships by potential duration in failed firms*

Eventual length of apprenticeship	Quarters to failure 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16+
Shorter than time to failure ($L < K$)	0.0	0.3	0.7	1.8	6.5	8.0	9.7	10.1	16.6	9.1	14.8	12.0	15.3	12.6	19.7	14.2
Within one quarter of failure ($L = K$)	10.1	10.4	11.5	14.2	11.6	8.2	12.5	14.1	10.9	16.3	5.6	0.0	0.0	0.0	0.0	0.0
Longer than failure but incomplete ($K < L < 11$)	11.7	9.1	9.5	12.4	8.8	7.5	8.6	3.7	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Complete (11–13 quarters)	58.9	57.4	58.1	48.7	53.8	56.1	48.4	51.0	50.7	48.8	56.1	74.4	70.2	61.6	50.7	58.9
Longer	19.3	22.8	20.3	23.0	19.3	20.2	20.8	21.1	17.8	25.8	23.5	13.6	14.5	25.8	29.7	27.0
Mean duration of apprenticeship (days)	979	999	994	973	971	995	978	1,003	976	1,054	1,028	1,063	1,050	1,076	1,061	1,078
No. of observations	195	298	148	226	398	415	279	298	477	559	392	594	731	834	458	32,682

Compare this distribution to those who joined the failing firm within 10 quarters or less to failure. Only around 48–55 percent of apprentices finish within 11–13 quarters in this group, while 18–25 percent take longer. There is no evidence that any apprentices affected by the firm failure during the apprenticeship took longer to complete their training than those not affected. This suggests that the monotonicity assumption is likely satisfied in our sample. On the flip side, 26 percent of affected apprentices left their apprenticeship early. This means that early termination in this group is about 12 percentage points larger than among those who trained in firms failing after 16 or more quarters. This difference is statistically highly significant and economically large.

This suggests that the failure of a firm induces some apprentices to leave the apprenticeship earlier than they would have otherwise. In fact, among apprentices who joined firms within 10 months of failure, about 14 percent of early terminations occur in the categories $L < K$ and $K < L < 11$, matching early terminations among those who joined 16 or more quarters before failure closely. About 12 percent are found in the category where $L = K$, indicating that the additional early terminations basically all occur right around the time of the firm failure.[5] Hence, it appears as if much of the dropping out induced by the failure of the firm seems to be coming from apprentices, who would have otherwise completed their apprenticeship.

Table 2 also displays the mean duration of the apprenticeship for each potential duration group. This is around 1,000 days, slightly less than three years or 12 quarters. Apprentices who joined the failing firms 16 quarters or more before failure have an average duration of 1,078 days which is 80 days more than those who joined closer to the failure. Since quarter to failure is our instrument, this gives a broad indication for the magnitude of the first-stage effects.[6]

Table 2 also displays the number of observations in each quarter to failure group. This number is relatively small for the apprentices who joined very close to failure. It becomes somewhat more stable for those apprentices who joined before five quarters to failure, although the numbers are highest for

[5] If one looks at the full cross-tab with quarterly information for every category of L instead of aggregating rows as in Table 2, the quarter where $K = L$ always has the highest cell percentage except for quarter 12, which is the regular apprenticeship duration.

[6] It is straightforward to decompose the total duration of apprenticeships into complete and incomplete durations according to $E(L) = E(L|\text{incomplete})P(\text{incomplete}) + E(L|\text{complete})(1 - P(\text{incomplete}))$, where we treat any duration up to 10 quarters as incomplete. For the group that joined firms within 10 quarters of failure, average incomplete durations are 562 days compared to 569 days among unaffected apprentices. For completed durations, the means are 1,147 and 1,162. These means are rather similar, also indicating that almost all the difference in durations is accounted for by the higher probability of not completing the apprenticeship among those who were affected by firm failure.

those who joined 12–14 quarters to failure. This indicates that within the last year before the failure of the firm it may be clear that the failure is imminent or possible. Therefore, fewer apprentices are hired at that stage. This is cause for concern that there may also be more selection among the group of apprentices who are still hired, or among the firms that still hire apprentices at that stage. Hence, we restrict the estimation sample to those apprentices who join the firm within four quarters or more to failure.

The relatively large number of apprentices who were in the failing firm for 12–14 quarters might indicate that firms delay their exit behavior when they have an apprentice whose completion of the apprenticeship is imminent. This behavior by firms would only impact our instrumental variables estimates if it were correlated with unobserved determinants of the later earnings of the apprentices affected.[7] However, it may also explain the relatively small number of apprentices who join 15 quarters to failure, and the fact that many of those who are observed tend to have relatively long apprenticeships. They may simply be apprentices in occupations with long apprenticeship duration.

Before proceeding, we merged the data on apprentices analyzed in the previous tables with the wages the apprentices earn later. There are up to 23 wage observations per apprentice, and the average number of observations per individual is 8.5. The resulting panel dataset has 47,881 observations. We account for the repeated observations on individuals in the following regressions by clustering standard errors at the level of individuals, using the "cluster" option in Stata 10.

We use two versions of the instrument. The first is a series of dummy variables for each quarter until the firm failure. Alternatively, we use a variable that is linear in the quarters until failure. The single linear instrument is more robust to weak instrument problems, and it produces somewhat different results. Table 3 displays the corresponding first-stage regressions.[8] The other regressors in the model are age and age-squared, a dummy for individuals without Austrian nationality, and dummies for 22 survey years and eight regions. The length of the apprenticeship is measured in days, but is scaled in terms of years, so that our estimates are easily comparable to standard estimates of returns to education. The coefficients on potential duration bounce around and are often insignificant for the first five quarters. Between nine and 12 quarters, the eventual length of the apprenticeship

[7] This would be the case, for example, if a firm decided to delay its exit when it has a particularly able apprentice but close early with a weak apprentice.

[8] We also regressed the linear version of the instrument on the few pre-apprenticeship covariates available in our data: age at the start of the apprenticeship, whether the individual is a foreigner, and education (which is only available for individuals who claim unemployment benefits at some stage during the observation period). These variables were neither individually nor jointly significant in our estimation sample.

Table 3. *First-stage regressions for length of apprenticeship in years*

	(1)	(2)
Potential duration in failing firm		
5 quarters	0.0457	
	(0.049)	
6 quarters	0.149*	
	(0.046)	
7 quarters	0.0988**	
	(0.051)	
8 quarters	0.0732	
	(0.048)	
9 quarters	0.104***	
	(0.045)	
10 quarters	0.234*	
	(0.044)	
11 quarters	0.237*	
	(0.045)	
12 quarters	0.303*	
	(0.042)	
13 quarters	0.288*	
	(0.042)	
14 quarters	0.355*	
	(0.042)	
15 quarters	0.266*	
	(0.045)	
Potential duration in failing firm		0.116
		(0.008)*
R-squared	0.08	0.07
F-statistic for joint significance of excluded instruments	79.58	749.56

Notes: Number of observations is 47,881. An observation is an individual-year. Regressions also contain a constant, dummies for the survey year and region, age, age-squared and a dummy for foreigners. Robust standard errors in parentheses are corrected for clustering at the level of the individual. Reference category for column (1) is four quarters. *Significant at 10 %; **significant at 5 %; ***significant at 1%.

rises, and finally reaches a plateau of more than a quarter of a year in the range from 12–15 quarters. Not surprisingly, this pattern mirrors the one we already found in Table 3 above.

The individual coefficients for quarters 10 and higher are highly significant. The F-statistic for the joint significance of the excluded dummy instruments is 80. This suggests that the correlation of our instruments with the endogenous regressor is high enough that "weak instrument" problems are unlikely. Stock, Wright and Yogo (2002) consider first-stage F-values below 10–20 to be problematic. Our F-statistic clearly exceeds these thresholds. F-statistics with the linear instrument are even higher. Not surprisingly, when we estimate our models with LIML we obtain identical results to standard 2SLS.

Table 4 reports our main results. Column (1) displays a relatively standard log wage regression on the length of the apprenticeship, controlling

Table 4. *Wage returns to length of apprenticeship training*

	OLS (1)	OLS (2)	IV (3)	IV (4)
Length of apprenticeship (in years)	0.052*	0.027*	0.023	0.038***
	(0.002)	(0.002)	(0.016)	(0.018)
Potential experience	0.048*			
	(0.001)			
Potential experience-squared	−0.001*			
	(0.000)			
Age		0.107*	0.108*	0.105*
		(0.004)	(0.005)	(0.005)
Age-squared		−0.002*	−0.002*	−0.002*
		(0.000)	(0.000)	(0.000)
Foreigner	−0.0212**	−0.0162	−0.0174	−0.0128
	(0.013)	(0.013)	(0.014)	(0.014)
Instruments	—	—	11 dummies for potential duration in failing firm	Linear variable for potential duration in failing firm
R^2	0.51	0.51	—	—

Notes: Number of observations is 47,881. Dependent variable is the log daily wage. Regressions also contain a constant, dummies for the survey year and region. Robust standard errors in parentheses are corrected for clustering at the level of the individual. *Significant at 10 %; **significant at 5 %; ***significant at 1 %.

for potential labor market experience. The return to an additional year of apprenticeship is about 5 percent. This is consistent with earlier work by Fersterer and Winter-Ebmer (2003a), who find returns to apprenticeship training of around 15 percent for the entire apprenticeship period. Given that an apprenticeship lasts on average about three years, this implies a return of about 5 percent a year.

Column (2) displays OLS estimates where we replace labor market experience with age. Labor market experience is correlated with the length of the apprenticeship, and hence also endogenous. Controlling for age rather than experience is therefore common in many IV studies of the returns to schooling. The return when controlling for age is now 2.7 percent, much lower than controlling for experience. It is common to find lower returns to education in regression models controlling for age rather than potential experience. In our sample of young workers, additional apprenticeship training replaces experience which is also quite valuable, so the difference in the estimates is large.

The IV estimates are shown in columns (3) and (4). The estimate of 2.3 percent in column (3) using the dummy instruments is practically the same as the OLS estimate. The return is now less precisely estimated, and is not significantly different from zero. This imprecision is a natural consequence of the fact that IV estimation removes much of the variance in the length of the apprenticeship variable. The magnitude of our standard errors is similar to those in other IV studies of the returns to schooling, for example, Angrist and Krueger (1991). Column (4) shows the results using the linear version of the instrument. The point estimate is higher than in column (3) but this estimate is still not significantly different from the OLS estimate. Our IV estimates therefore suggest returns which are similar or somewhat larger than the OLS returns.

It is important to keep in mind that our approach identifies the joint effect of apprenticeship training and any commensurate change in experience. Since the identifying firm closures take place during the apprenticeship period, any change in experience necessarily involves post-layoff experience. One worry regarding our study is therefore that some other variable, for example "time since layoff", rather than apprenticeship duration drives our results, as in von Wachter and Bender (2006). As a check on our approach, we therefore computed similar estimates on a sample of apprenticeship-trained workers, who experienced a layoff due to firm closure in their first job *after* rather than *during* the apprenticeship.

In this sample, everybody got to complete their apprenticeship but experienced a layoff within the first 16 quarters in the first post-apprenticeship job. This job may or may not be in the training firm, depending on whether and how long apprentices stayed at that firm. As before, workers in this sample also have to find a new job after the layoff, and we only consider wages after workers find jobs in new firms. The sample for this robustness check is slightly smaller (31,040 observations) since we have fewer post-layoff periods on the wage now.

In order to create results that are most closely comparable, we focus on the reduced-form regression, which corresponds to our IV model in column (4) of Table 4. The reduced form involves a regression of the wage on K_i, the potential duration in the firm where the layoff occurs. For simplicity, we focus on the specification with the linear potential duration. The reduced-form coefficient corresponding to column (4) of Table 4 is 0.0044 (with a standard error of 0.0021). In our sample for post-apprenticeship layoffs, the corresponding coefficient is -0.00014 (with a standard error of 0.0012). Hence, given our setup, the timing of layoffs in the first post-apprenticeship job does not affect later wages at all.[9] This is comforting since it suggests that it is unlikely that it is the dynamics of wages subsequent to a layoff which drives our results.

In summary, we find the IV estimates of the wage returns to apprenticeship training to be as high as or higher than the OLS estimates. We

[9] This is not necessarily inconsistent with the literature on wage losses and subsequent recoveries associated with layoffs; for example, von Wachter and Bender (2006). Our key regressor does not measure time *since* displacement but rather potential time *to* displacement in the layoff firm.

have exercised care in selecting a sample where biases in the IV estimates due to a potential correlation of the instrument with apprentice quality are less likely. As a result, our estimates are restricted to a specific set of apprentices who trained in a narrow set of firms. Therefore, we do not think that the actual magnitude of the results is as interesting as the relative similarity of the OLS and IV estimates, although the basic OLS results are not very different from earlier findings on a broader set of training establishments.

VI. Conclusion

Apprentices and non-apprentices differ in many dimensions, and there is anecdotal evidence that better training providers tend to select the best school leavers. This suggests that standard OLS returns to apprenticeship training may be biased upward because of the selection. This is the first study to use an exogenous event in an attempt to estimate returns which are free from selection bias.

Our focus on failed firms necessarily implies that we are not able to estimate returns for a very representative sample of apprentices. We focus on small firms. Both small firms and firms that fail in general may be relatively poor training providers, particularly in the final phase of their existence. Moreover, we estimate returns to partially completed apprenticeships, which is different from the return to the entire apprenticeship. In particular, we estimate primarily returns to the latter part of the apprenticeship, which may consist mainly of work experience rather than formal training, and therefore have lower returns. But it also includes any credential effects, which may raise returns. Our sample also excludes individuals who become self-employed, which may be associated with greater success and hence higher wages. Finally, we observe wages for workers when they are relatively young (26 years on average). For all these reasons our return estimates might be expected to differ from those in a more representative sample. Nevertheless, our basic OLS estimates are similar to earlier findings.

Our main result is that the estimated returns for apprentices affected by firm failure are not very different from the OLS returns in the same sample. This suggests that selection in the drop-out behavior of apprentices is not particularly important. OLS returns to apprenticeship training have previously been found to be of similar magnitudes to other forms of education. From this we conclude that apprenticeship training does not seem to be superior to other forms of school-based education, say, in colleges or vocational schools. This conclusion, of course, rests on the premise that we can extrapolate the degree of selection bias in our highly selected sample to the general population of apprentices. One drawback of our focus on small firms is that an important degree of selection may be due to the size of the training firm, with better apprentices recruited by larger firms.

There are some other important limitations to our study. We only estimate wage returns. Some of the benefits of apprenticeship training may be in the form of higher employment probabilities and a reduced incidence of unemployment. Our dataset is not well suited for analyzing this possibility since it only covers private sector employment. Hence, we cannot distinguish unemployment from employment in the public sector or self-employment. Apprenticeship training is often regarded as facilitating an easier school-to-work transition than is the case in purely school-based education systems. This is difficult for us to ascertain, because all of our apprentices had their ties to the original training firm severed because of the firm failure. This makes it difficult to draw conclusions about attachment to the firm and the labor force from this particular group. In addition, we only look at wage returns for apprentices while they are still relatively young. Some of the returns to apprenticeship training may accrue later on; for example, when some trained workers go on to become master craftsmen.

References

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

Angrist, J. D. and Imbens, G. W. (1995), Two-stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity, *Journal of the American Statistical Association 90*, 431–442.

Angrist, J. D. and Krueger, A. B. (1991), Does Compulsory School Attendance Affect Schooling and Earnings?, *Quarterly Journal of Economics 106*, 979–1014.

Angrist, J. D., Imbens, G. W. and Rubin, D. (1996), Identification of Causal Effects using Instrumental Variables, *Journal of the American Statistical Association 91*, 444–455.

Becker, S. and Siebern-Thomas, F. (2004), Supply of Schools, Educational Attainment and Earnings, manuscript, University of Munich.

Bender, S., Dustmann, C., Margolis, D. and Meghir, C. (2003), Worker Displacement in France and Germany, in P. J. Kuhn (ed.), *Losing Work, Moving On*, W. E. Upjohn Institute, Kalamazoo, MI.

Büchtemann, C. F., Schupp, J. and Soloff, D. (1994), From School to Work: Patterns in Germany and the United States, in J. Schwarze, F. Buttler and G. G. Wagner (eds.), *Labour Market Dynamics in Present Day Germany*, Campus, Frankfurt am Main, 112–141.

Fersterer, J. and Winter-Ebmer, R. (2003a), Are Austrian Returns to Education Falling Over Time?, *Labour Economics 10*, 73–89.

Fersterer, J. and Winter-Ebmer, R. (2003b), Smoking, Discount Rates and the Returns to Education, *Economics of Education Review 22*, 561–566.

Franz, W., Inkmann, J., Pohlmeier, W. and Zimmermann, V. (2000), Young and Out in Germany. On Youths' Chances of Labor Market Entrance in Germany, in D. Blanchflower and R. Freeman (eds.), *Youth Unemployment and Joblessness in Advanced Countries*, University of Chicago Press, Chicago, 381–425.

Freeman, R. and Schettkat, R. (2001), Skill Compression, Wage Differentials and Employment: Germany vs. the US, *Oxford Economic Papers 53*, 582–603.

Hamermesh, D. and Pfann, G. (2001), Two-sided Learning, Worker Turnover and Displacement, NBER Working Paper no. 8273.

Heckman, J. (1993), Assessing Clinton's Program on Job Training, Workfare, and Education in the Workplace, NBER Working Paper no. 4428.

Ichino, A. and Winter-Ebmer, R. (1999), Lower and Upper Bounds of Returns to Schooling: An Exercise in IV Estimation with Different Instruments, *European Economic Review 43*, 889–901.

Ichino, A. and Winter-Ebmer, R. (2004), The Long-run Educational Costs of World War II, *Journal of Labor Economics 22*, 57–86.

Imbens, G. W. and Angrist, J. D. (1994), Identification and Estimation of Local Average Treatment Effects, *Econometrica 62*, 467–475.

Krueger, A. and Pischke, J. S. (1995), A Comparison of East and West German Labor Markets Before and After Unification, in R. Freeman and L. Katz (eds.), *Differences and Changes in Wage Structures*, University of Chicago Press, Chicago, 405–445.

Nickell, S. and Bell, B. (1996), Changes in the Distribution of Wages and Unemployment in OECD Countries, *American Economic Review 86*, 302–308.

OECD (1998), Getting Started, Settling In: The Transition from Education to the Labour Market, *OECD Employment Outlook*, OECD, Paris, 81–122.

Ryan, P. (2001), The School-to-Work Transition: A Cross-national Perspective, *Journal of Economic Literature 39*, 34–92.

Schneeberger, A. and Kastenhuber, B. (1996), Anfänger- und Abschlußquoten in der Lehrlingsausbildung, Institut für Bildungsforschung der Wirtschaft, Schriftenreihe Nr. 102, Vienna.

Schneeberger, A. (Project Leader) with Freundlinger, A., Hinterstein, R., Kastenhuber, B., Klimmer, S., Matzner, G. and Nowak, S. (no date), *Informationen zur Lehrlingsausbildung 2000/2001*, Institut für Bildungsforschung der Wirtschaft, Vienna.

Soskice, D. (1994), Reconciling Markets and Institutions: The German Apprenticeship System, in L. M. Lynch (ed.), *Training and the Private Sector. International Comparisons*, NBER Series in Comparative Labor Markets, University of Chicago Press, Chicago, 25–60.

Steedman, H. (2001), Benchmarking Apprenticeship: UK and Continental Europe Compared, CEPR Discussion Paper no. 513.

Stock, J. H., Wright, J. H. and Yogo, M. (2002), A Survey of Weak Instruments and Weak Identification in Generalized Methods of Moments, *Journal of Business and Economics Statistics 20*, 518–529.

von Wachter, T. and Bender, S. (2006), In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers, *American Economic Review 96*, 1679–1705.

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