

IZA DP No. 378

The Promise of Workplace Training for Non-College-Bound Youth: Theory and Evidence from German Apprenticeship

Damon Clark
René Fahr

October 2001

The Promise of Workplace Training for Non-College-Bound Youth: Theory and Evidence from German Apprenticeship

Damon Clark

Nuffield College, Oxford, CEP and IZA, Bonn

René Fahr

IZA, Bonn and University of Bonn

Discussion Paper No. 378
October 2001

IZA

P.O. Box 7240
D-53072 Bonn
Germany

Tel.: +49-228-3894-0
Fax: +49-228-3894-210
Email: iza@iza.org

This Discussion Paper is issued within the framework of IZA's research area *Mobility and Flexibility of Labor Markets*. Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent, nonprofit limited liability company (Gesellschaft mit beschränkter Haftung) supported by the Deutsche Post AG. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public. The current research program deals with (1) mobility and flexibility of labor markets, (2) internationalization of labor markets and European integration, (3) the welfare state and labor markets, (4) labor markets in transition, (5) the future of work, (6) project evaluation and (7) general labor economics.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character.

ABSTRACT

The Promise of Workplace Training for Non-College-Bound Youth: Theory and Evidence from German Apprenticeship*

This paper assesses the potential of 'workplace training' with reference to German Apprenticeship. When occupational matching is important, we derive conditions under which firms provide 'optimal' training packages. Since the German system broadly meets these conditions, we evaluate the effectiveness of apprenticeship using a large administrative dataset. We find returns to apprenticeship for even the lowest ability school-leavers comparable to standard estimates of the return to school, and show that training is transferable across a wide range of occupations, such as a one-digit occupation group. We conclude that the positive experience with German Apprenticeship Training may guide the design of similar policies in other countries.

JEL Classification: C29, J24, J31, J62

Keywords: German apprenticeship training, human capital, occupational mobility, wages

Damon Clark
Centre for Economic Performance
London School of Economics
11 Furnival St
London, WC2A 2AE
United Kingdom
Tel.: +44 (0)20 7955 6955.
E-mail: d.clark@lse.ac.uk

* This paper, also available as CEP Working Paper No. 1153, has greatly benefited from discussions with Steve Nickell, Hilary Steedman and Margaret Stevens. Brian Bell, Sami Berlinski, Jordi Blanes, Alan Manning, Tom Nicholas, Steve Pischke, Imran Rasul, Wendelin Schnedler, Andrew Sweeting, Jon Temple and Rainer Winkelmann also provided helpful comments. Financial support from ESRC and DFG is gratefully acknowledged.

1 Introduction

In countries such as the UK and the US, where there exists no well-established path from school to work, policies designed to help non-college-bound school-leavers into the labour market are high up the policy agenda.¹ Although school-to-work policies comprise a diverse mix of ideas and proposals, the common denominator is an emphasis on workplace training as a means of smoothing the transition from school to work.² Proponents of these policies subscribe to the widely-held view that education and training (particularly of non-college bound youth) is central to economic performance, but also draw on recent findings casting doubt on the efficacy of publicly provided training and education programs (see, for example, Heckman (1999)). In both of these respects, the basic thrust of the workplace training idea is very much in step with the conventional policy wisdom.

In other respects however, the idea of workplace training for school-leavers goes against the grain of modern thinking about the labour market. In particular, whilst it was thought in the 1980s that high rates of job mobility amongst school-leavers were in some sense problematic, influential research in the 1990s has emphasised the importance of the role played by the labour market in sorting young workers into the firms and the occupations in which they are most productive (Topel and Ward (1992), Neal (1999)). In this context, workplace training may restrict school-leaver's job shopping opportunities by tying them to firms and occupations that they may not be well suited to. As Neal (1999) puts it, "institutions that limit returns from search may lead to...an inefficient assignment of workers to tasks in the economy" (p. 257).

Is workplace training for school-leavers a good thing or a bad thing? It is instructive to start from the social planner's perspective. Workplace training is good to the extent that it improves productivity within a particular firm or occupation, but bad in that it is costly to provide, and because it introduces costs to productive job-shopping. That is, the prospect of losing firm- or occupation-specific capital prevents young workers from shopping for their most productive match. Clearly, this depends on the training being 'specific'. The more transferable is the training provided, the less costly it is to shop for new jobs and occupations.

To analyse the problem more formally, the first part of this paper breaks with the assumption that training consists of the sum of general and firm-specific components and instead presents a formulation of workplace train-

¹See Ryan (2001) for a cross-country survey of the school-to-work transition.

²For the US, see Hughes, Bailey, and Mechur (2001) for a positive assessment of existing policies, Bassi and Ludwig (2000) for a detailed account of several school-to-work schemes and Krueger and Rouse (1998) for an analysis of a workplace training scheme aimed at a wider age group. For the UK, Steedman, Green, and Ryan (1998) propose a radical extension of the UK's relatively small 'Modern Apprenticeship' scheme.

ing in which firms train in *intensive* and *extensive* dimensions.³ Of course Becker himself argued that training can not always be represented as the combination of firm-specific and general components: “(s)ome training may be useful not in most firms nor in a single firm but in a set of firms defined by a product, type of work or geographical location” (Becker (1975), p.35). We define ‘intensive training’ as training that improves a worker’s productivity in a particular task or occupation, whilst ‘extensive training’ increases the transferability of the intensive training to other occupations. The advantages of this training technology are twofold: first, they allow for an asymmetry in the transferability of training between occupations. Unlike with the general-specific dichotomy, we can, given some assumptions about occupational distance, say that intensive training in occupation A is more transferable to occupation B than occupation C. Secondly, this training technology is also attractive from an empirical point of view, since we can measure the returns to intensive training within an occupation and infer the transferability of training between occupations.

Returning to the social planner’s problem, we show that given this training technology, the planner will maximise social welfare by providing training that contains an ‘extensive’ component designed to facilitate post-training occupational mobility, even though occupational mobility would still be lower than it would be without training (since training will not be ‘general’). Under the assumption that there are a large number of firms competing to provide training and recruit trained workers within every occupation, we then show that the privately optimal training programme exactly mirrors the socially optimal one. In principle therefore, we would expect the ‘optimal’ workplace training programme to be offered to school-leavers.

Since our model is concerned with occupational matching, it is related to the literature that considers this issue in more depth. This includes Miller (1984), McCall (1990) and Neal (1999). The reason we keep the occupational mobility decision so simple - workers can inspect occupational matches on the job and so move when match improvements offset human capital losses - is in order that we can investigate the human capital investment decision.

The effect of future turnover on the human capital investment decision has previously been analysed in the context of firm matching but not occupation matching. In particular, Stevens (1994), Chang and Wang (1996) and Scoones (2000) explore the firm’s human capital investment decision when there is matching of firms to workers, or at least some turnover. The essential differences between these models and ours are first, that we deal with occupational turnover and secondly, that since we assume ‘competitive’ occupations, firms earn no rents, there are no externalities and the privately optimal training decision mirrors the socially optimal one.

To bridge the gap between theory and practise, the second part of the

³This distinction was first introduced by Kim (1989).

paper considers an existing workplace training programme that broadly corresponds to the situation outlined. German Apprenticeship Training (GAT) is a mass workplace training program regularly completed by more than two-thirds of German school-leavers. Training typically lasts for three years, is organised along occupation lines, and competition for potential apprentices and newly apprenticed workers is stiff.

Interestingly, many advocates of workplace training in the UK and US regard German Apprenticeship Training (GAT) as a blueprint or model for school-to-work policies. For example, with reference to GAT, Baily, Burtless, and Litan (1992) advocate the subsidy and regulation (via skill standards and certification) of workplace training for non-college-bound US youth. Others however, are more cautious, advocating only that “work-based learning, as instituted in the Germany style...is an idea worth trying out on a small scale” (Heckman, Roselius, and Smith (1994), p.141).

An exchange of views about GAT between these authors goes to the heart of the workplace training debate.⁴ Burtless (1994) argues that GAT is to be admired because, *inter alia*, the returns to training seem high. Heckman, Roselius, and Smith (1994) counter first, that the returns are up for debate and secondly, that even if the returns to training *within the training occupation* are high, the occupation ‘specificity’ of the training may prevent future occupation shopping: “the very narrow technical training and rigid curriculum of the apprenticeship program may contribute to diminished options in later life” (p.99). On the second point, Burtless counters that GAT must be relatively transferable to facilitate certain types of occupational mobility.⁵

In our empirical work, we attempt to resolve these issues by estimating the returns to training within the training occupation and by estimating the transferability of training to other occupations. The first of these exercises entails a comparison of the earnings of apprentices trained inside their occupation with non-apprentices. The second involves an analysis of the loss of earnings experienced upon moving out of the training occupation. Our approach therefore combines an analysis of the returns to apprenticeship training based on an earnings equation written in levels (see Card (1999) for a review) and an analysis of the effects of displacement on earnings (see for example Neal (1995) and Dustmann and Meghir (1999)). The difference between our approach and the one typically taken to estimate the returns

⁴See the comments of Burtless (1994) on Heckman, Roselius, and Smith (1994) and the reply of the latter.

⁵For example, it is an oft-stated fact within the GAT literature that apprentices trained as bakers often work for motor companies such as Ford upon completing apprenticeship. Heckman, Roselius, and Smith (1994) would probably argue that this was a product of the fact that the returns to apprenticeship are low, so that there is very little cost to this type of mobility. Burtless (1994) on the other hand would argue that the costs of this kind of mobility are low because returns to apprenticeship are high and apprenticeship training is transferable across occupations.

to schooling is that we focus only on apprentices still working within their training occupation. The difference between our approach and the displacement literature is that we focus on displacement from an occupation rather than a firm or industry.

To preview our results, we find returns to GAT within the occupation trained in comparable to the returns to a year of schooling typically reported in the literature. Moreover, we find this training to be transferable within a broad occupational group, such as a 1-digit occupation. Consistent with this finding, we show that there exists a high degree of mobility out of the apprenticeship occupation, although as we would expect, the mobility of German workers across 1-digit occupational groups is lower than that for untrained workers in the US.

The paper is organised as follows: section 2 sets up a model of workplace training and in section 3 we describe how we plan to estimate the parameters of interest. Section 4 provides an outline of GAT and section 5 describes the data used in later sections. Section 6 describes the occupational mobility of German apprentices and we estimate the returns to GAT, and the transferability of GAT in section 7. Section 8 concludes with some remarks on the policy implications of our findings.

2 A Model of Workplace Training

In this section, we present a model of workplace training. The model will enable us to shed light on the issues raised in the Introduction, and will pave the way for our empirical analysis. The basic structure of the model is simple, but flexible enough to allow us to consider both training and matching as sources of productivity and wage growth. Before we outline the basic assumptions of the model, we consider the related literature.

2.1 Related Literature

The existing training literature - both in its theoretical and empirical form - starts from the hugely influential distinction made by Becker (1975) between general training (of equal use to every firm in a labour market) and specific training (of use only to the training firm). Whilst this serves as a neat way of distinguishing who is most likely to bear the costs of training⁶, it does not necessarily provide an accurate representation of what we typically think of as training. For example, it is not clear how the ability to repair TVs, to program computers, or to write economics papers can be said to be the combination of general and specific training. As noted in the Introduction,

⁶As is well known, Becker claimed that firms would be unwilling to finance the costs of general training since they could capture none of the returns in competitive labour markets; it was claimed that the costs and returns to specific training would be shared between trainee and firm.

Becker himself appreciated that the general-specific dichotomy was not a useful classification of all types of skills.

In this spirit, we adopt an alternative, and we believe, a more natural characterisation of workplace training. In particular, we suppose that the choices are made over *intensive* and *extensive* training. Intensive training is training in a particular skill or task. Extensive training involves learning about related tasks and acquiring a deeper understanding of the task being performed.⁷ The key characteristic of this kind of training is that it does not improve the worker's performance within her current firm, but does improve her performance in other firms.

Our model considers occupational matching in a very simple way. In particular, we assume a two-period structure in which all occupational match information is revealed between the two periods. Papers focussing on occupational mobility making more realistic assumptions about match revelation include McCall (1990) and Neal (1999). Neal (1999) presents a model with occupation and firm matching in which workers have to switch firms in order to realise different occupation matches. The implication is that workers will have an optimal two-stage rule by which they will match first to an occupation and then a firm. Using NLSY data, Neal (1999) finds some evidence in favour of this pattern. This can be thought of as a simpler version of McCall (1990), in which information can arrive at different times, so that the optimal search policy is more complicated. In fact McCall (1990) focusses on the prediction that if a worker observed in her second job did not change occupations, the hazard on this second job should be a decreasing function of previous tenure.

The advantage of our simpler characterisation of match revelation is that we can consider the effects of occupational turnover on the human capital investment decision. The effect of firm turnover on human capital investment has been analysed by Chang and Wang (1996), although the key assumption in their model is asymmetric information regarding training levels. Without this, firms would never pay for general training and would only under-invest in specific training given the assumption that surpluses are shared between worker and firm. In contrast, Stevens (1994) shows that firms and workers may over-invest in specific training. In her model, the force generating turnover is post-training productivity shocks rather than worker-firm matching, and this has the effect of generating expected rents

⁷This distinction was first introduced by Kim (1989) to analyse the connection between labour specialisation and market size. To construct an example close to hand, suppose that a research assistant is hired to estimate a series of equations using a particular dataset. Intensive training would involve showing the research assistant how to load the data into a software package, execute the estimation commands and obtain the output. Extensive training would involve learning some of the the econometric theory underlying the estimation methods and knowing what to do if for example, the dependent variable was limited in some way or the estimated errors displayed serial correlation.

for third parties. Firms and workers *over-invest* in specific training to reduce these rents.

The focus of Scoones (2000) is more directly related to our paper, since it analyses the worker’s decision to invest in both general and specific training in the context of firm matching. In this model, the social planner’s training decision is identical to ours: that is, the social planner takes account of the probability of turnover in determining the optimal specific training investment (the optimal general training investment is independent of turnover). The difference between his model and ours is that since we are dealing with ‘competitive’ occupations, firms earn no rents, there are no externalities and neither party can distort the choice of training package. In his model, worker’s strategically invest in specific training in order to improve their outside options in the event that they move firms.

2.2 The Model

We assume that a finite number of tasks can be performed in the economy. In the case of GAT, these will refer to occupations, hence we use tasks and occupations interchangeably. Importantly, we assume that there are a large number of identical firms offering jobs within each of these occupations. These firms are price-takers in the output market, and we normalise the output price to unity.

This ‘identical firms’ assumption implies that there is no such thing as firm-specific human capital, and results in an equilibrium in which workers pay all of the costs of their training. Of course this may not be a reasonable description of certain occupations. However, as we will see in section 4, apprenticeship occupations are defined very narrowly. Moreover, we can provide an informal test of this assumption using the QaC data (described in section 5 and used more extensively in section 7). The survey asks apprentices how many of the apprenticeship skills are used on the current job. Comparing those in the training occupation and the training firm with those in the training occupation but a different training firm, the results are almost identical. For the former (latter) group, the responses are very few or none (0.78%, 0.95%); few (3.25%, 4.31%); some (10.53%, 11.04%); many (24.30%, 27.34%); and very many (61.14%, 56.36%) based on sample sizes of 893 and 951 respectively. Also consistent with this assumption is the fact that whilst firms subsidise training costs in certain occupations, in many occupations, workers do indeed pay for the training themselves.⁸

⁸Of course apprentices do not pay the costs of training up front. Instead, they earn wages lower than the value of their apprenticeship product during training. See Harhoff and Kane (1997) for a discussion.

Technology/Skill Space

Since we are interested in transferability, and since transferability suggests some asymmetries between occupations (a painter's skills may transfer easily to other kinds of decoration but not to electrical engineering), we need a way of representing occupations that makes this explicit. The obvious vehicle is a circle, since this avoids the endpoint problem. That is, the space refers to technologies, and occupations are located on the perimeter of the circle. Although our results generalise to this case, for ease of exposition, we consider the much simpler case of two occupations (A and B).

Untrained Workers

We assume a continuum of workers who supply labour inelastically. Like the firms in this economy, the workers are risk-neutral. We assume that workers are heterogenous in two dimensions. To see this, consider equations (1) and (2), which describe the product of untrained worker i in every firm in occupations A and B :

$$\Pi_{Ai}^{NOAPP} = m(q_i) + \xi_{Ai} \quad (1)$$

$$\Pi_{Bi}^{NOAPP} = m(q_i) + \xi_{Bi} \quad (2)$$

Equations (1) and (2) say that worker product is an additive function of 'general ability' $m(q_i)$ and the worker-occupation match ξ_{ji} , $j \in \{A, B\}$. General ability is a function of worker quality q_i , where $q_i \in (\underline{q}, \bar{q})$ and the worker-occupation match is assumed to be drawn randomly from a distribution $f(\xi)$ with support $(\underline{\xi}, \bar{\xi})$. The match is therefore independent of general ability.

Transferable Training

As already noted, the concept of transferability implies that there exist asymmetries between occupations. This suggests that the dichotomy traditionally used in the training literature between general skills (of use to all firms) and specific skills (of use to only one firm) is not appropriate for our purposes. In this spirit we adopt an alternative and, we believe, a more natural characterisation of workplace training. In particular, we suppose that training choices are made over *intensive* and *extensive* dimensions. Intensive training (denoted ' T ') is training in a particular occupation. Extensive training (denoted ' X ') involves learning about related tasks and acquiring a deeper understanding of the task being performed.⁹ The key characteristic of extensive training is that it does not improve a worker's performance within her current occupation, but does improve her performance in other occupations.

⁹This distinction was first introduced by Kim (1989) to analyse the connection between labour specialisation and market size.

Trained Workers

Now consider worker i after she has received the training package (T, X) as part of an apprenticeship. This worker has the following product when working in occupations A and B , and having received apprenticeship training in occupation A :

$$\Pi_{Ai}^{APP} = \alpha(q_i)T + m(q_i) + \xi_{Ai} \quad (3)$$

$$\Pi_{Bi}^{APP} = \alpha(q_i)T - L[\alpha(q_i)T, X] + m(q_i) + \xi_{Bi} \quad (4)$$

Comparing (1) and (3), worker i apprenticed and working in occupation A is more productive by $\alpha(q_i)T$ than untrained worker i working in occupation A . This is the value of intensive training - the annual return to the intensive training $\alpha(q_i)$ multiplied by the years of intensive training undertaken (T). We assume that the annual return to intensive training is an increasing function of worker quality: $\alpha'(q_i) > 0$.

Comparing (2) and (4), worker i trained in occupation A and working in occupation B is more productive by $\{\alpha(q_i)T - L[\alpha(q_i)T, X]\}$ than untrained worker i working in occupation B . The function $L[\alpha(q_i)T, X]$ is central to the analysis of this section as it represents the loss of skills involved with transferring intensive training from occupation A to occupation B . We assume that $L[\cdot]$ has the following properties:

$$L_1[\cdot] \geq 0; L_{11} \leq 0; L_X[\cdot] \leq 0 \quad (\text{P1})$$

$$L_1[., 0] = 1; L_1[., X] < 1; \lim_{X \rightarrow \infty} L_1[\cdot] = 0 \quad (\text{P2})$$

Property (P1) says that for given levels of extensive training, increased levels of intensive training increase the loss of skills incurred upon switching occupations, but at a decreasing rate. Similarly, these losses are decreasing in the level of extensive training. Property (P2) says that without any extensive training ($X=0$), all intensive skills are lost upon switching occupations so that human capital is occupation-specific. For positive levels of extensive training, not all intensive training is lost upon switching occupations, and in the limit, as $X \rightarrow \infty$, training becomes general. A simple functional form that satisfies both P1 and P2 is $L = (\frac{\alpha(q_i)T}{1+X})$.

Training is assumed to cost $C(X, T)$ where we assume that this function has the following properties:

$$C_X(\cdot) > 0; C_T(\cdot) > 0 \quad (\text{P3})$$

$$C_{XT}(\cdot) = 0; \lim_{T \rightarrow 0} C_T(\cdot) = 0; \lim_{X \rightarrow 0} C_X(\cdot) = 0 \quad (\text{P4})$$

Property (P3) says that training costs are increasing in the levels of both intensive and extensive skills. The assumption that $C_{XT}(\cdot) = 0$ simplifies the analysis, but it does not imply that the choice of training package is

separable in X and T , due to the presence of both T and X in the skill loss function $L[\alpha(q_i)T, X]$. The final pair of assumptions embedded in property (P4) ensure that the social planner will invest in some training.

Timing and Information

As is standard in the training literature, we assume that workers exist in the labour market for two periods: a training period and a working period. Prior to the first period, we assume that workers' optimal matches are unknown to all parties. After the first period, workers' optimal matches in both occupations are revealed to trainees and all firms. We assume that there is no uncertainty relating to any other aspect of the economy, and we assume that training levels are both observable and verifiable.

Under these assumptions, at the start of the first period, firms within each occupation will compete for new workers by offering training packages (W, T, X) , where W is the first-period wage, and these packages may depend on observed worker quality q_i . At the start of the second period, after the occupational matches have been revealed, firms compete for newly trained workers.

Labour Market Outcomes

Consider the labour market for trained workers, taking as given the training that these workers obtained during the first period. Since there are many firms within each occupation, and since occupational matches are common knowledge, workers will always be paid the value of their marginal product. Hence the probability that the worker will leave the training occupation A is:

$$\begin{aligned}
 P(\text{leave}) &= P(\Pi_{Bi}^T > \Pi_{Ai}^T) \\
 &= P(\alpha(q_i)T - L + m(q_i) + \xi_{Bi} > \alpha(q_i)T + m(q_i) + \xi_{Ai}) \\
 &= P(\xi_{Bi} > \xi_{Ai} + L) \\
 &= \int_{\underline{\xi}}^{\bar{\xi}-L} \left[\int_{\xi_{A+L}}^{\bar{\xi}} dF(\xi_B) \right] dF(\xi_A) \tag{5}
 \end{aligned}$$

In other words, workers leave when the gains in match exceed any loss of skills associated with switching occupations.

2.3 Social Planner's Problem

The social planner has two choices: for a given worker i she chooses the optimal level of intensive and extensive training to provide for this worker in the event that this worker is trained. She then decides whether this worker should be trained. Consider first the choice of training levels.

Choice of Training Levels

In this model, per capita social welfare is simply expected per capita output less the direct costs of training, $C(T, X)$:

$$\begin{aligned}
 SW &= \int_{\underline{\xi}}^{\bar{\xi}-L} \left[\int_{\xi_A+L}^{\bar{\xi}} [\alpha T - L + \xi_B] dF(\xi_B) + \int_{\underline{\xi}}^{\xi_A+L} [\alpha T + \xi_A] dF(\xi_B) \right] dF(\xi_A) \\
 &\quad + \int_{\bar{\xi}-L}^{\bar{\xi}} [\alpha T + \xi_A] dF(\xi_A) - C(T, X)
 \end{aligned} \tag{6}$$

To derive the socially optimal training programme for worker i , we obtain the first-order conditions for the social planner's maximisation problem by differentiating equation (6) with respect to T and X . These can be written as:

$$SR_T = \alpha - L_T \int_{\underline{\xi}}^{\bar{\xi}-L} [1 - F(\xi_A + L)] f(\xi_A) d\xi_A - C_T = 0 \tag{7}$$

$$SR_X = -L_X \left[\int_{\underline{\xi}}^{\bar{\xi}-L} [1 - F(\xi_A + L)] f(\xi_A) d\xi_A \right] - C_X = 0 \tag{8}$$

Since the term $\int_{\underline{\xi}}^{\bar{\xi}-L} [1 - F(\xi_A + L)] f(\xi_A) d\xi_A$ is simply the probability of the worker leaving occupation A from (5), these first-order conditions can be rewritten as:

$$SR_T = \alpha - L_T P(\text{leave}) - C_T \tag{9}$$

$$SR_X = -L_X P(\text{leave}) - C_X \tag{10}$$

We now have the following proposition:

Proposition 1 *Provided there is some occupational turnover, if the social optimum exists, it is characterised by positive levels of both intensive and extensive training.*

Proof. We show first that $T \neq 0$. This is established from SW_T using the fact that $L_T \leq \alpha$ and $P(\text{leave}) < \frac{1}{2}$. Since $C_T = 0$ when $T=0$, the first term in SR_T is positive whilst the second is zero, hence we have a contradiction of $SW_T = 0$. We now show that $X \neq 0$. Suppose $X=0$ and $T > 0$. Then since $L_X < 0$ and provided $P(\text{leave}) > 0$, the first term of SW_X is positive whilst the second term is zero since $C_X = 0$ when $X=0$. Hence $X=0$ represents a contradiction of $SW_X = 0$. *QED.*

Proposition 1 says that provided there is some occupational turnover, the social planner will take account of this by choosing a training package (T^*, X^*) that contains some extensive training. In other words, the socially

optimal training package will adapt to the expected degree of occupational mobility.

Comparative Statics

We now ask how the socially optimal package changes with q_i . The answer is straightforward, and summarised in the following proposition.

Proposition 2 $\frac{dT^*}{dq_i} > 0$

Proof. Taking differentials of the first-order conditions with respect to α gives us:

$$\frac{dT}{d\alpha} = \frac{SR_{TX}SR_{X\alpha} - SR_{XX}SR_{T\alpha}}{SR_{TT}SR_{XX} - SR_{TX}SR_{XT}}$$

From the second-order conditions for a solution to the social planner's problem, the denominator is positive and SR_{XX} is negative. We also have that $SR_{T\alpha} \geq 0$, since $L_{11} \leq 0$ from (P1) and $\frac{dP^{(leave)}}{d\alpha} < 0$. It is easily shown that SR_{TX} and $SR_{X\alpha}$ must take the same sign hence $\frac{dT}{d\alpha} > 0$. *QED.*

In other words, as we would expect, the optimal level of intensive training provided is increasing in the quality of the worker. Without being more specific about the cost function $C(\cdot)$ or the loss of skills function $L(\cdot)$, we can not unambiguously sign the effect of q_i on X^* . The ambiguity arises for the following reason. One effect of increased q_i is to increase the value of intensive training and so increase the potential loss of skills incurred upon switching occupations and so *increase* the value of extensive training. However, another effect is to reduce the probability of leaving the training occupation, making extensive investments less useful. It can be shown that for sufficiently positive C_{TT} the second effect dominates and the effect of increased q_i is to reduce X^* . Similarly, we can show under these conditions that an exogenous increase in occupational turnover increases the optimal level of extensive training and decreases the optimal level of intensive training. This is also as we would expect - more turnover induces the social planner to increase the transferability of training.

Number of Workers Trained

To analyse the optimal number of workers trained, we examine the net social return to training given optimal levels of T and X and given worker quality q_i :

$$SR(T^*, X^*; q_i) = R(\alpha(q_i)T^*, X^*) - C(X^*, T^*) \quad (11)$$

where R is the gross return to training and C the cost of training. Examining (3) and (4) and using the fact that $L_1 < 1$ from (P2), we have that $\frac{\partial R}{\partial q} \geq 0$, so that the value of training is increasing in q inside and outside of the training occupation. Using the envelope theorem, this implies that the social return to training is a strictly increasing function of q_i . Hence training will be socially optimal for any worker of quality $q \geq q^+$, where $SR(T^*, X^*; q^+) = 0$.

2.4 Private Optimum

When training levels are both observable and verifiable, firms are able to credibly commit to training packages (W, T, X) . In that case, firms will maximise profits subject to a constraint that the expected utility of the worker is at least as high as that offered by another firm. Competition among firms ensures that firms make zero profits, and so the problem is equivalent to firms maximising the utility of young school-leavers subject to a zero profit condition. Since the utility of a worker is just the sum of expected wages over the two periods, the firm's problem is:

$$\begin{aligned} \max_{T, X} U(W, T, X) &= W + P(\textit{stay})E[\Pi_{A_i}^{APP}|\textit{stay}] + P(\textit{leave})E[\Pi_{B_i}^{APP}|\textit{leave}] \\ \textit{s.t. } W &= -C(T, X) \end{aligned}$$

Since this reduces to precisely the problem faced by the social planner in (6), we have proved the following proposition:

Proposition 3 *In equilibrium, the training package offered by private firms is identical to that chosen by the social planner and an identical number of workers are trained.*

This result says that the privately optimal training package will be exactly that chosen by the social planner. Hence private firms will anticipate the possibility of occupational turnover and adapt the training package offered.

2.5 Training Market Failure

As yet, we have encountered no 'failure' in the market for training. This sits uneasily with the frequently heard suggestion that workplace training ought to be regulated, by for example standardising courses, and certification (see for example, Burtless (1994)). Is there a role for this kind of policy?

Notice first that we would expect firms to offer certificates as part of the training package, although we have no reason to suppose that the training courses offered in every occupation would conform to some national standard. A natural analogy is with language courses. These are typically unregulated, with institutions competing for students by offering both certificates and end-of-course tests. Even though the curriculum and the examining board differ between courses, it is not clear that there is a strong case for government intervention in this market.

A more significant problem would arise if training were non-verifiable. In this case, firms cannot commit to providing given levels of X and T and there is a potential moral hazard problem. If firms can not signal the levels of training offered, and if they can not build reputations for offering certain training levels, then they will have no incentive to provide any training

over and above a minimum quality training package (T^{MIN}, X^{MIN}) . In this case, there may be an argument for ‘licensing’ training providers, by ensuring for example that all trainers are themselves trained. This policy has been analysed by Shapiro (1986) and is operational in Germany (where all firms offering apprenticeships must employ trained ‘Meister’ or trainers).

In reality however, it is likely that firms can signal training levels and build reputations. Again, the analogy with language schools is apposite. In this market, schools signal quality by offering free trial periods, and by advertising the qualifications of their teachers. Since they are typically in the market for a long period of time, there is scope to build reputations. We return to these issues in our conclusions.

2.6 Summary

When firms compete to offer training packages to workers, they will offer socially optimal packages. These packages will provide more intensive training to higher quality workers, will adapt to the expected degree of post-training occupational mobility and will be paid for by workers. In our empirical analysis, we will examine the returns to GAT within the training occupation and the transferability of GAT across occupations. The next section outlines how we will identify these parameters.

3 Empirical Framework

In this section, we describe how we will identify the key parameters of interest: the return to intensive GAT and the cost of transferring intensive skills from the training occupation to other occupations. Again, we consider only the two-occupation case (A and B), although our results generalise to the n -occupation case. We start by describing the log-earnings of worker i in occupations A and B . This follows directly from equations (1), (2) (3) and (4):

$$\ln W_{Ai}^{NOAPP} = m(q_i) + \xi_{ai} \quad (1')$$

$$\ln W_{Bi}^{NOAPP} = m(q_i) + \xi_{bi} \quad (2')$$

$$\ln W_{Ai}^{APP} = \alpha(q_i)T(q_i) + m(q_i) + \xi_{ai} \quad (3')$$

$$\ln W_{Bi}^{APP} = \alpha(q_i)T(q_i) - L + m(q_i) + \xi_{bi} \quad (4')$$

where A is the training occupation and where ‘NOAPP’ (‘APP’) refers to worker i without (with) an apprenticeship qualification. From (5) we know that the worker *leaves* the training occupation when:

$$(\xi_B - \xi_A) > L \quad (5')$$

In other words, when the improved (occupational) match exceeds the cost of transferring intensive (occupational) skills.

3.1 Relation to the Existing Literature

We estimate the value of intensive training in a relatively standard fashion (see for example Lynch (1992)). Whilst we show that the presence of matching has implications for our estimates, we are able to derive an approximate lower bound to the value of apprenticeship training. Controlling for the presence of matching to identify the costs of transferring training across occupations is more difficult. Intuitively, workers leave the apprenticeship occupation to realise a better match (equation (5')). Hence estimates based on a comparison of movers and stayers is unlikely to reveal the true cost of transferring training between occupations.

The approach that we take - we use a sample of displaced workers to generate what are effectively 'exogenous' occupation changes - is similar to that taken by other papers that attempt to control for match-driven mobility. For example, Dustmann and Meghir (1999) focus on the returns to tenure and experience in the context of a model in which wages are also driven by worker-firm matching.¹⁰ As in our framework, the job change decision trades off match improvements with losses of human capital (firm-specific human capital in their case) and a sample of displaced workers is used to generate 'exogenous' firm changes.¹¹

The empirical paper closest to ours is Werwatz (1998). Werwatz also addresses the question of how transferable is GAT between occupations, finding that occupational 'movers' earn similar wages to occupational 'stayers', and concluding that GAT must be fairly transferable across occupations. Since Werwatz has only cross-sectional data, he controls for endogenous mobility by estimating a switching regression model. The fact that the selection terms are rarely found to be statistically significant could indicate that selection biases are not a problem for his results. However, it is more likely that the selection equation has been inadequately specified. In particular, it is not clear that a variety of 'quality of work' measures (such as standing up at work) capture what is driving occupational mobility (the search for better occupational matches) or that these do not belong in the earnings equation. Fortunately, the panel nature of the IAB data that we use enables

¹⁰In fact the model is more general, as the returns to tenure are allowed to vary across firms and individuals, and the returns to experience across individuals.

¹¹Returns to experience are estimated using a sample of workers starting a new job, and these estimates are used to calculate the within-firm wage growth which must be due to a combination of tenure and changes in match quality. Here, the selection problems are that workers observed working for a particular firm chose to join this firm and chose not to leave it. A sample of firm closures is used to control for the first problem and age is used as an instrument for tenure.

us to improve upon this strategy. The following subsections describe how we do this.

3.2 The Returns to Intensive Apprenticeship Training

We estimate the returns to intensive apprenticeship training under two assumptions: first, that school-leavers do not differ in quality and secondly that school-leavers do differ in quality.

Homogenous School-Leaver Quality

The first of these assumptions implies that $q_i = q$ for all i , so that $m(q_i) = m$, $T(q_i) = T$ and the return to one year's worth of intensive training (α) is a homogenous parameter. Suppose that we wish to estimate this parameter. We do this by estimating the total value of intensive training (αT) and then dividing by the average number of years spent in intensive training T (which we can measure from the data). Consider estimating by ordinary least squares an equation of the following form (ignoring covariates such as experience that have 'common effects'):

$$\ln W = a_0 + a_1 APPIN + \varepsilon \quad (12)$$

where $APPIN$ refers to an apprentice working inside of the training occupation, ε is a random disturbance term and the base group are those without any apprenticeship training. Then from (1'), (3') and (5'), the probability limits of \hat{a}_0 and \hat{a}_1 are given by:

$$\begin{aligned} p \lim \hat{a}_0 &= m_0 + E(\xi_a | \xi_a - \xi_b > 0) \\ p \lim \hat{a}_1 &= \{m_0 + \alpha T + E(\xi_a | \xi_a - \xi_b \geq -L)\} - \{m_0 + E(\xi_a | \xi_a - \xi_b > 0)\} \\ &= \alpha T + E(\xi_a | \xi_a - \xi_b \geq -L) - E(\xi_a | \xi_a - \xi_b > 0) \\ &\leq \alpha T \end{aligned}$$

since $L \geq 0$. Hence, using the estimated value of a_1 from (12), we will estimate a lower bound to αT and therefore a lower bound to α .

Heterogenous School-Leaver Quality

It is perhaps more plausible to assume that school-leavers differ in quality q_i . Hence as assumed in section 2, m and α are increasing functions of q_i , T is an increasing function of q , and α is a heterogenous parameter. Under this assumption, estimates of $\alpha(q_i)$ based on equation (12) represent lower bounds on the returns to intensive apprenticeship training *for those choosing to become apprentices*. Since we have shown in our discussion of equation (11) that only those school-leavers with $q > q^+$ will actually be apprenticed, it should be obvious that this is neither the population mean return to apprenticeship nor the return for those on the margins of apprenticeship and

non-apprenticeship (i.e. those with $q \simeq q^+$). Since we are often concerned with the effects of apprenticeship-type programs on hard-to-educate workers in other countries, this latter parameter may be of particular interest.

We provide an estimate of this parameter using data disaggregated by training firm size. It is often asserted that there is a clear ranking in both the quality of apprenticeship programs offered and the quality of applicants. Consistent with our model, it is certainly the case that large firms offer more intensive training than small firms, since their apprenticeship programs tend to last up to one year longer than those offered in smaller firms (as we will see in section 6). Moreover, as Steedman (1993) notes, “In the public mind in Germany, a definite and complex ranking of apprenticeship places exists linked to expected lifetime returns. As a general rule, *Industrie* apprenticeships are more highly sought-after than *Handwerk* apprenticeships” (p.1285). In fact, *Industrie* and *Handwerk* broadly correspond to large and small firms respectively, and Steedman’s claim is supported by evidence presented in Harhoff and Kane (1993), who find that the proportion of apprentices reporting good mathematics and good German scores in school are strongly increasing functions of apprenticeship firm size.

Since we have data on firm size, we can disaggregate the apprenticeship variable in equation (12) and use estimates of the returns to those trained in the smallest firms as an approximation to $\alpha(q^+)$, the return to apprenticeship training to those workers of lowest quality.¹²

3.3 The Costs of Transferring Intensive Apprenticeship Training

In order to investigate the worst-case scenario, we would like to identify an upper bound to the costs of transferring intensive apprenticeship skills.¹³ To see how we might do this, suppose that we had a sample of apprentices who are working in their training occupation in period $t-1$. Then consider their change in log earnings between periods $t-1$ and t . Since we know from equation (5’) that workers will only move out of the training occupation when $\xi_B - \xi_A > L$, comparing the wages of ‘movers’ and ‘stayers’ would cause us to under-estimate the costs of transferring training.

¹²In the model, there is no heterogeneity among firms and so we offer no formal explanation as to why the highest quality school-leavers train in the largest firms (as opposed to a situation in which every firm offers a range of programs to cater for different abilities). One possibility is that if training in large firms generates complementary firm-specific skills, the returns to these skills will be shared between trainee and firm, hence the expected lifetime utility of apprentices is higher in larger firms and larger firms earn higher rents on higher quality workers. Of course this assumes a fixed number of large firms.

¹³A lower bound to the costs of transferring training is easily derived by augmenting equation (12) to include a dummy for being an apprentice outside of the training occupation. A lower bound to the cost of transferring training then corresponds to the difference between the estimated coefficient on this dummy and $\widehat{\alpha}_1$.

We can, however, make some progress by basing this comparison on a sample of workers displaced for exogenous reasons (e.g. plant closure). To see this, consider estimating the following equation, where ‘MOVOUT’ is a dummy variable indicating whether the worker has left the training occupation:

$$\Delta \ln W = b_0 + b_1 MOVOUT + \varepsilon \quad (13)$$

Then we can estimate an upper bound to the costs of transferring training under the following assumption:

A1 Workers are randomly displaced from their firms. These workers accept the first job that they are offered and may decide to search for a more suitable position on the job.

In a formal model of search with offers arriving exogenously on and off the job, assumption A1 requires that the arrival rate of offers to unemployed workers is no greater than the arrival rate to workers in a job and that search costs for unemployed workers are at least as large as those employed workers.¹⁴ For workers accepting an offer within their training occupation, the expected change in log earnings is then zero. For those accepting a job outside of the training occupation, the change in log earnings is from equations (3’), (4’) and (5’):

$$\begin{aligned} E(\Delta \ln W | MOVOUT_i) &= -L + E(\xi_B | \xi_A - \xi_B \geq -L) - E(\xi_A | \xi_A - \xi_B \geq -L) \\ &= -L + E(\xi_B | \xi_B - \xi_A < L) - E(\xi_A | \xi_A - \xi_B \geq -L) \\ &< -L \end{aligned}$$

Hence we have that:

$$\begin{aligned} p \lim \widehat{b}_1 &= \{-L + E(\xi_B | \xi_A - \xi_B \geq -L) - E(\xi_A | \xi_A - \xi_B \geq -L)\} - \{0\} \\ &< -L \end{aligned}$$

This implies that estimates of b_1 derived from equation (13) for a sample of displaced workers are downward-biased under assumption A1 so that we *over-estimate* the costs of transferring skills. This can be explained as follows: The initial occupation was chosen in period ($t-1$) and the new occupation was not chosen in period ($t-1$). Hence the expected value of the match in the former is positive, the expected match in the latter negative and so a move outside of the training occupation entails an expected loss of match productivity as well as the cost of transferring skills.

¹⁴See for example Mortensen (1986).

This result may not hold under a less restrictive assumption regarding the job-search behaviour of displaced workers. A more plausible assumption is A2:

A2 There is some cost U to rejecting an offer from outside the training occupation. This represents the expected cost of remaining unemployed and waiting for an offer from inside the training occupation.

Then, displaced workers will only accept an offer from outside of the training occupation when $\xi_B - \xi_A > L - U$, so that behaviour will only mirror that under A1 for those workers with large U (in other words, when there is no option but to move). Under this more general assumption, an obvious strategy is to find proxies for U that will enable us to instrument the decision to leave the training occupation. These instrumental variables estimates could then be interpreted as if they were derived under A1. Another option is to use the answers to a survey question regarding the usefulness of skills learned during the apprenticeship that we interpret as abstracting from any matching considerations. We discuss both of these strategies in section 6.

4 German Apprenticeship Training

The model assumed a training technology with firms providing intensive training in a narrowly defined task or occupation and extensive training that enabled intensive training to be transferred to other occupations. It also assumed large numbers of firms competing for trained and untrained workers within every occupation. To see how this corresponds to the structure of GAT, we begin by outlining the 1969 Vocational Training Act, which remains the foundation stone of GAT.

This Act explicitly defined a number of occupations in which school-leavers could apprentice. Whilst these currently number 375, fewer than the 600 that could be apprenticed in the 1970s, they are defined very narrowly. For example, within the class of electrical occupations (a two-digit category) school-leavers can apprentice in 15 different occupations.¹⁵ This provides a rationale for our assumption that the skills acquired by apprentices are task-rather than firm-specific.

Although the length of apprenticeship depends on the apprenticeship occupation, GAT typically lasts between two and three and a half years. Importantly, the Act specifies the curricula to be followed in each of these occupations. For example, as Berg (1994) reports, training as a metalwork-

¹⁵These include occupations such as ‘electronic specialist, telecommunications’; ‘electronic specialist, communications (telecommunication systems)’; ‘electronic specialist, communications (information systems)’ and ‘electronic specialist, communications (radio engineering)’. See Federal Ministry of Education and Science (1992) for more details.

ing apprentice calls for a year of basic occupational training for all metals trades, a year of training in a general occupational group, and 1.5 years of training in a specialised area. This corresponds neatly to the training technology assumed in section 2, in that the first two years of training can be thought of as ‘extensive’, whilst the final 1.5 years can be thought of as ‘intensive’.

A crucial part of the curriculum for every training occupation involves training firms releasing their apprentices for one day per week to attend a local vocational school. These are organised around one of five vocational fields (industry, commerce, home management, agriculture and other occupations) and are designed to fill any gaps in general education and to prepare apprentices for the final examination. Steedman, Green, Bertrand, Richter, Rubin, and Weber (1997) describe the curriculum followed by an apprentice in industrial administration (*Industriekaufmann*):

“An apprentice in industrial administration (*Industriekaufmann*) goes to school $1\frac{1}{2}$ days per week. During the first year of apprenticeship, s/he takes 1 hour of German, some Sport and Religion/Ethics, 1-2 hours of English, 1-2 hours general economic and social studies, 3 hours of accounting and finance, and 3 hours of business studies. These courses amount to approximately 11 hours per week during the first year of apprenticeship and about 9 hours during the second year.” (p.69)

This type of training is clearly ‘extensive’, in that the 7-8 hours spent studying business-related courses will help trainees to transfer intensive occupation-specific skills to a wider range of business-related occupations. Training is completed when apprentices pass the final examinations. These typically consist of several written examinations in the subjects laid down by the training regulations, with many including an oral or practical component.

Having shown that the intensive/extensive training technology assumed in our model broadly corresponds to GAT, we turn to another key assumption, the competition for new trainees and new school-leavers. It has sometimes been suggested that the centralised German wage bargaining structure limits the degree of post- and pre-apprenticeship competition among firms. In fact though, the wages bargained centrally are more like minimum wages, with firms free to increase wages above these minimum levels. In any case, firms can increase wages by changing the job titles of workers and in the case of apprenticeships, by offering different fringe benefits. Casey (1991) reports evidence of this practise.

If we are correct in supposing that the GAT system broadly corresponds to our model, then we would expect to see some returns to apprenticeship within the training occupation and a degree of relatively costless occupational turnover facilitated by the extensive training component. The

remainder of the paper investigates these empirical issues in more detail, starting with a discussion of the data used.

5 Data Issues

The paper uses data from a 1% sample of German social security records (see the IAB Data Appendix for a fuller description of the data set). The data are available for the years 1975-1995, and are supplemented by data on the firms to which workers are attached. Importantly, this allows us to infer the occupation trained in and the size of the training firm. Generally speaking, these data are well suited to the task at hand. In particular, due to the administrative nature of the data, the wage information and timing of employment spells is very accurate. One problem with the data is that prior to 1984, firms were not obliged to report extra payments such as Christmas and holiday bonuses. Since these are an important part of compensation in Germany, all of our earnings equations are estimated using data from 1984 onwards. Also, the data do not cover the entire German labour force. Civil servants and the self-employed do not make social security contributions in Germany, and so they are not present in the data. Finally, although the data are top coded, the top coding affects only a tiny proportion of the young apprentices in our sample.

5.1 The Sample

Only German males are retained for analysis, and our sample consists of two groups: apprentices and non-apprentices. In order to exclude those engaged in short training spells, internships and the like, apprentices are defined as those having been observed training for greater than 450 days. We further restrict the sample of apprentices to those without the *Abitur* (usually completed by those that will eventually attend University) and those starting their apprenticeship aged 19 or under. The age restriction is designed to include those that take their military service after leaving school, but exclude those training after a spell in the labour market. We exclude those with an *Abitur* as the labour market for apprentices with this qualification will be significantly different to that for those without an *Abitur*. In any case, this group is relatively small.¹⁶ To make the sample of non-apprentices as comparable as possible, we include only those whose first spell is observed aged 19 or under and who do not have the *Abitur*.

Table 1 displays some descriptive statistics for our sample. The most noticeable feature of the Table is the gradual ageing of the sample. Since the maximum age for a person in the sample increases from 19 in 1975 to 39

¹⁶This group makes up less than 20% of all apprentices according to the author's calculations with the German Socio-Economic Panel (GSOEP).

in 1995, this is to be expected. Note also that the proportion of workers in the sample with apprenticeship training is increasing over the observation window. This reflects an increase in the proportions of young school-leavers undertaking apprenticeship training over this period. Due to the fact that the sample is relatively young, we do not observe many apprentices with the Meister certificate (an advanced vocational qualification typically undertaken by apprentices with several years of labour market experience).

5.2 Displacement

A key part of our empirical strategy involves the construction of a sample of ‘exogenously’ displaced workers. In this respect, the fact that plants are given a unique identifier in the IAB data helps, although we cannot assume that the disappearance from the data of a plant identifier implies that a plant has closed. This can happen for a variety of reasons, including closure, takeover or a merger. To deal with this problem, we construct three subsamples of ‘separations’:

‘Displaced’ First, we use a sample of workers who experience an unemployment spell after separation. We further restrict this unemployment spell to be greater than one month to avoid including those workers that quit their previous firm and exclude workers with unemployment spells of greater than one year to avoid problems regarding the scarring effects of unemployment. Although this upper limit is somewhat arbitrary, experiments suggest that it does not impact much on our results. Whilst this sample does not enable us to disentangle those workers displaced exogenously and those displaced for ‘cause’, displacements for cause are only a problem in equation (13) when they are based on unobserved and transitory components of earnings which are correlated with the decision to move out of the training occupation. Hence results based on this subsample are robust to dismissals for cause based on permanent components of earnings (observed or unobserved).

‘Close’ We can compare our results using the ‘displaced’ subsample to those obtained by further restricting this sample to those workers who separated, experienced an unemployment spell and whose plant identifier disappeared from the data. If we assume that plants that merge or reorganise lay workers off on a ‘last-in-first-out’ basis, this group will contain a higher proportion of workers displaced for exogenous reasons. Since we do not know the exact date at which the plant closed in the IAB data, we generate two samples of workers displaced because of ‘closure’: those whose plant identifier disappeared within one and two years of the separation date (‘close1’ and ‘close2’ respectively).

‘Quits’ Finally, for the purpose of comparison, we present the results for a sample of workers separating firms but not experiencing an intervening spell of unemployment.

Table 2 presents some descriptive statistics for the four different groups. Focusing on the pre-displacement characteristics, we find significant differences between these groups. In the second and third rows, we see that displaced workers are slightly younger than the other groups, although there are no significant differences between the ‘quits’ and the ‘close’ samples. The most marked differences occur with respect to pre-displacement tenure. The finding that this is lowest amongst the ‘displaced’ workers is consistent with a layoff policy of ‘last-in-first-out’. This is very important in Germany (see Bender, Dustmann, Margolis, and Meghir (1999) for details) and may also explain the differences between the ‘close’ and ‘quit’ samples, since those ‘close’ workers that were displaced prior to the actual closure of the plant will also have been subject to the ‘last-in-first-out’ rule. Pre-separation wages reflect these differences, and it is interesting to note that in a regression of pre-displacement wages on pre-displacement characteristics and dummy variables representing the groups ‘displaced’, ‘close2’ and ‘close1’, the estimated co-efficients (standard errors) on these variables were -0.00374 (0.00616) for displaced workers, and -0.0234 (0.0267) and -0.0251 (0.0438) for ‘close2’ and ‘close1’ respectively. Hence, controlling for pre-displacement characteristics, we can not reject the hypothesis that these workers represent a random sample of pre-displaced workers.

6 Occupational Mobility

Before analysing the returns to and transferability of apprenticeship, we begin by providing a brief overview of the occupational mobility of German apprentices. Specifically, we ask how the occupational mobility of German apprentices compares with that of a comparable group of untrained workers (here taken to be young males in the US). The point here is not to say that mobility is too low or too high in either country. As we argued in the Introduction, reduced mobility is an indirect cost of workplace training that needs to be weighed against the benefits that this form of training can provide. Instead, the purpose of this section is to account for the differences in cross-country mobility outcomes in terms of the nature of the workplace training that school-leavers receive.

To assess the mobility of apprentices out of their training occupations, Figure 1 plots Kaplan-Meier estimates of the nonparametric survival functions of post-apprenticeship spells in the 3-digit training occupation. That the probability decreases sharply upon completion of apprenticeship train-

ing implies that a significant proportion of apprentices leave the apprenticeship occupation immediately. Although the hazard decreases at a much slower rate after this point, it remains the case that after 20 years in the labour market, 75% of apprentices have left the training occupation at the 3-digit level. The graph for the probability of leaving the 1-digit occupation shows a similar pattern but a slightly higher proportion of workers remaining within the 1-digit occupation (about 35%). This implies that two-thirds of German apprentices eventually leave the apprenticeship occupation at the 1-digit level.

To compare this level of occupational mobility with that amongst a comparable group of workers without any formal workplace training, we use the results of Neal (1999), who focuses attention on occupational mobility in the US. He finds that the majority (55%) of firm changes amongst his sample of young men also involve changes of occupation and industry, and he interprets this finding as suggestive of young workers engaging in task-shopping. To obtain an equivalent estimate, we focus on the firm changes made by those initially working inside the training occupation (since those that have already left their training occupation are not constrained by the costs of transferring occupation-specific skills). Adopting Neal's definition of occupational mobility (moves involving a change of three-digit occupation and one-digit industry), we find that approximately 30% of all these job changes involve changes of occupation.

That this is much lower than Neal's estimate is consistent with our finding of significant costs to switching occupations at the 1-digit level (since 1-digit occupation switches are strongly correlated with 1-digit industry switches). However, this is not the same as saying that apprentices are trapped in their 1-digit occupation. Indeed, as we have already noted, after 20 years in the labour market, almost two-thirds of apprentices have left the 1-digit occupation. This suggests that whilst apprenticeship training reduces the occupational mobility of apprentices below that characterising young workers in the US, apprentices typically find sufficiently good match opportunities to induce them to leave the 1-digit apprenticeship occupation.

Table 3 analyses occupational and firm mobility for the ten most prominent apprenticeship occupations ranked by the number of observations in the pooled sub-sample. As one would expect firm mobility is still overall higher than occupational mobility. Also most of the apprenticeship occupation show the highest number of changes at the 1-digit level their is some variance among occupations. Occupations with no moves at the 3-digit level simply reflect the fact that there are no alternative occupations at the three digit level (e.g. toolmaker).

7 Estimates of Returns and Transferability

Our objective in this section is to estimate an approximate lower bound to the return to apprenticeship training within the training occupation $\alpha(q^+)$ and an upper bound to the costs of transferring training.

7.1 Returns to Apprenticeship

In order to obtain our approximate lower bound to the value of intensive skills (α), we begin by assuming that this value does not depend on school-leaver quality q , and simply split apprentices according to whether they are working inside or outside the training occupation. From the top panel of Table 4, we see that estimates of α based on equation (12) are approximately 0.15. With the conservative assumption that apprenticeship training lasts for an average of 2.75 years, and that an average of two-thirds of the apprenticeship is spent training, this gives an annualised average return of approximately 8.2%. This is comparable to estimates of the rate of return to schooling found in the literature. For example, Angrist and Krueger (1991), using US Census data, find estimates of the rate of return between 5% and 7% when estimating by OLS, and between 6% and 10% when estimating by IV.¹⁷ Our estimates of the return to additional years of experience is large, although this is likely to reflect the fact that we observe these workers over their first few years in the labour market, when much of the information regarding occupational match productivity is revealed. The low returns to tenure are consistent with those reported in Dustmann and Meghir (1999).

As noted in section 3, the value of apprenticeship is likely to depend on worker quality q . To obtain our approximate lower bound to the value of intensive skills $\alpha(q^+)$, we split the group of apprentices according to the size of their training firm. As a check that apprenticeship firm size proxies school-leaver quality, the first column of Table 5 presents self-reported school test scores (Harhoff and Kane (1997)). As commonly assumed in the GAT literature, we see a clear correlation between training firm size and school-leaver quality, although the group with the lowest scores are actually those trained in the second smallest firm size. The second column calculates the proportion of apprentices trained in firms of this size, whilst the third column

¹⁷For the purposes of comparing our results with the previous literature, the second panel of Table 2 presents the results of estimating equation (12), in which we pool the two types of apprentice (those working inside and outside the training occupation) to estimate the ‘return to apprenticeship training’ as commonly reported. This co-efficient averages roughly 0.19 over the sample period. Based on a similar specification, Winkelmann (1994), using the German Socio-Economic Panel (GSOEP), obtains co-efficients that average approximately 0.2 over the period 1985-1990. Werwatz (1998), using the Qualifications and Careers (QaC) data, obtains a co-efficient for 1985 of 0.137. That this is slightly lower than our estimate may be attributable to the fact that Werwatz (1998) includes additional covariates (such as industry dummies) in his earnings equation.

reports the co-efficients estimated from a version of equation (13) in which the apprenticeship variable is disaggregated according to firm size, and we pool across years (and include year dummies). To obtain estimates of the rate of return to an additional year of training, we again adjust these co-efficients according to the length of the training program (again assumed to be 2.75 years for each firm size group) and the average time spent training (column 5).

Rates of return across the different size groups are again in the broad range of estimates of the return to an additional year of schooling. Taking the return to one year of apprenticeship training for those workers trained in firms with between two and nine employees as the return to those on the margins of apprenticeship and work, this return is approximately 5.87%. As we would expect, this is lower than the return found for higher quality school-leavers training in larger firms, but it is still within the broad range of estimates of the rate of return to schooling. Moreover, since each of these estimates represent lower bounds on the true returns, we conclude that there is a significant amount of intensive training undertaken during GAT. We now attempt to estimate the transferability of this training.

7.2 Transferability

We now turn our attention to the transferability of apprenticeship training between occupations, and in Table 6, we present estimates of the costs of transferring training obtained from estimating equation (13). Before turning to the results for the groups of ‘displaced’ workers that we are interested in, we begin with our comparison group of ‘quits’. Looking at the left-hand column of the Table, we see that amongst the group of quits, the wage penalty associated with moving out of the training occupation is very close to zero. Whilst this might suggest that training is entirely transferable, these estimates are obviously biased because of the match-driven nature of this mobility. Turning then to column 2, we find that for displaced workers, the point estimate of the wage penalty is now negative and significant at the 10% level. This suggests that there are some costs to transferring occupations, although these are small in comparison to the total value of apprenticeship training (0.15 from Table 4). Moreover, in the final two columns, our estimates are not significantly different from zero.

The finding of only very small wage penalties associated with leaving the training occupation might suggest that training is transferable. However, it may be that training is more transferable between closely related occupations than between occupations which are a greater occupational ‘distance’ apart. To investigate this possibility, we disaggregate moves out of the training occupation and consider the wage penalties associated with different kinds of moves. We begin by measuring distance according to the occupational codes. That is, we say that a move at only the three-digit level

is a move into an occupation more closely related to the training occupation than one involving a move at the two-digit level. The penalties to moving out of the training occupation according to the distance moved are presented in the second panel of Table 6. Looking first at the left-hand column, for the quits, only the wage penalty associated with a move out of the training occupation at the 1-digit level is negative, although it is very small. Again, since the majority of these moves are selective, we would not expect to find large penalties to moving for this group.

For displaced workers, the wage penalty associated with moving out of the training occupation at the 1-digit level is slightly larger, and significant at the 1% level. However, it is still small when compared with estimates of the total value of apprenticeship training within the occupation (0.15 from Table 4). Amongst the ‘close’ samples, these estimates are also negative. For those leaving a plant that closed down within one year, the point estimate is larger in absolute value and approximately two-thirds of the value of intensive training. Taking this estimate as the worst-case scenario suggests that whilst training can be costlessly transferred within a 1-digit occupation, the costs of moving across a 1-digit occupation are large.

Since these codes may not be an adequate measure of occupational distance, we consider two means of improving upon them. First we look only at moves out of the training occupation that also involve changes in industry at the two-digit level (the classification is produced in Table A1). The idea here is that reported changes in occupation are less likely to be spurious if they are also accompanied by changes in reported industry. Looking at the third panel of Table 6, we find a similar set of results, with the exception that amongst the ‘close’ sample there is some evidence of wage penalties incurred for moves at the 2-digit level. Unfortunately, these estimates are not very precise.

An alternative strategy is to construct a measure of occupational distance from the data. We do this in the following way.¹⁸ For the full sample of firm separations, we construct an occupational transition matrix for which element (i,j) refers to the proportion of apprentices leaving occupation i for occupation j . Rather than use this as a measure of transferability however, we use this matrix to calculate occupational distances between occupations i and $i+1$ as $d_{i,i+1} = \frac{\sum_j |(i,j)-(i+1,j)|}{2}$. Hence distance lies between zero and one, with zero distance corresponding to a situation in which the pattern of transitions from occupation i to all other occupations is exactly the same as the pattern from occupation $(i+1)$ to all other occupations. We construct such a distance measure at the two-digit occupational level. Since the pair of two-digit occupations between which distance is estimated to be smallest (largest) are ‘farmer’ and ‘farm administrator’ (‘farm administrator’ and

¹⁸See Shaw (1987) for a detailed account of the construction of this type of distance matrix.

‘technician’ (chemist, physicist, etc.)) this gives us some confidence in our measure.

We use this measure to break down all moves at the two-digit occupational level into four equally sized groups, ranging from the smallest to the largest occupational distances moved. This serves as a classification of all two-digit moves according to the occupational distances moved. The bottom panel of Table 6 presents the results of this analysis. Looking again at the left-hand column, whilst none of the moves incur very large penalties, the pattern of wage differentials is as we might expect, with moves across the largest distances (quartile 4) incurring the largest penalties. Amongst displaced workers, we find significant penalties to moves across the second largest distances, and large wage penalties to moves across the largest distances. These are approximately equal to one-half of the value of apprenticeship training within the training occupation. Results for those leaving a plant that is about to close down suggest very high costs to leaving the training occupation and moving across the largest distances, with the point estimate (-0.28) greater than the value of apprenticeship training. Again, this suggests that roughly speaking, training is transferable within a broad occupational group but not outside of this group.

7.3 Results from a Question regarding Skill Use

It was observed in section 3.3 that if apprentices select the new occupation by trading off improved match values with the costs of transferring skills (as opposed to accepting the first offer and searching on the job), we may not be estimating an upper bound to the cost of transferring training. One solution to this problem is to instrument the decision to move out of the training occupation. Two instruments that we considered were whether the worker was married (assumed to be correlated with the value of leisure) and employment levels in the training occupation (assumed correlated with the arrival rate of offers from within the training occupation when unemployed). Whilst these variables entered the first stage regressions with the right sign, they were rarely significant and so the second-stage estimates were extremely unstable.

Hence as a final check on the robustness of our results regarding transferability, we use the answers from a question contained in the Qualifications and Careers Survey (QaC) data. The survey is cross-sectional, but it asks workers a number of retrospective questions that enable us to identify the training occupation of the worker. In particular, the survey asks workers: “How much of the occupational knowledge and skills you acquired during apprenticeship can you still apply in your current work?” The answer can be “very little or nothing at all; a little; some; quite a lot, a lot”. Since it is hard to see how workers could interpret this as anything other than a question concerning the actual value of apprenticeship skills in the new job,

it acts as a direct measure of transferability that is not affected by the value of the match in the new occupation.

Table 7 presents the answers to this question based on a sample of similar workers (German male apprentices without the *Abitur* under the age of 35). From the top panel, we see that overall, almost two-thirds respond that they are using ‘many’ or ‘very many’ of their apprenticeship skills, with the remaining third using ‘some’, ‘few’ or ‘very few or none at all’. When we split this group into those working inside and outside of the training occupation, the results are very interesting. Amongst those working inside their training occupation, almost 85% claim to be using ‘many’ or ‘very many’ of the skills acquired during apprenticeship. The figure for those outside of the apprenticeship occupation is just under 40%. Hence it is clear that apprenticeships are occupational. But it is interesting to note that even amongst those outside of the training occupation, only one-quarter claim to be using ‘very few or none’ of their skills.

In Panel B we break the movers down according to the distance moved. The results are very dramatic. Amongst those that move at a 3-digit level, only 1.4% claim to be using ‘very few or none’ of their skills, whilst 45.7% claim to be using ‘very many’. However, for those moving at the 1-digit level, the pattern is exactly reversed, with 30.31% claiming to use ‘very few or none’ and only 16.74% claiming to use ‘very many’. We find a similar pattern of results when interacting occupational moves with switches in two-digit industry (Panel C), although in Panel D, results based on the distance measure constructed from the IAB data are not as stark. This might suggest that the distance measure based on the data is not as accurate a measure of skill use as the occupational codes. Overall however, these skill use results reinforce the results based on earnings: apprenticeship training is transferable within a broad occupational group (e.g. a 1-digit occupation), but is not transferable outside of this group.

8 Conclusions

The paper began by stressing that workplace training has indirect as well as direct costs: namely, that it can prevent productive job-shopping. However, we showed in a theoretical model of workplace training that provided there is competition for trained and untrained workers within each occupation, the privately optimal training package will mirror the socially optimal one, which in turn will adapt training to the expected degree of occupational turnover. In the empirical part of the paper we painted a positive picture of GAT in which trainees receive an intensive training in a particular job skill or occupation and a sufficiently broad training to enable them to transfer these skills across a wide range of occupations. In line with these findings, patterns of occupational mobility suggest mobility from the training occupation is the

norm rather than the exception.

Given the picture we have painted, it would appear that workplace training is a promising alternative to traditional classroom-based routes to skills. This begs the question of why countries such as the UK and US do not have similar programs. One possible answer put forward in our theoretical discussion was the need for regulation to circumvent the moral hazard problem associated with the provision of non-verifiable training. Yet whilst this would account for the success of the heavily regulated German model, it is not clear why institutions such as free trial periods and reputations could not be used to overcome this problem without regulation.

Instead, a more plausible answer, and one suggested by Harhoff and Kane (1997), is that young school-leavers in other countries do not have the means to finance this kind of training since, for example, living in the parents' home into young adulthood is not the norm. In other words, our assumption that young workers are not credit-constrained and are willing and able to pay for their own training may need to be relaxed in other contexts. In that case, a wider examination of cost-sharing would also benefit from an investigation into why German firms are willing to share apprenticeship costs with workers in a minority of cases. Although the paper abstracted from this phenomenon, Acemoglu and Pischke (1998) have made an interesting start to answering this question¹⁹ and further analysis of GAT along these lines would be interesting.

¹⁹Acemoglu and Pischke (1998) present a model in which the training firm's superior information regarding the trainee's ability allows them to extract rents from trained workers and hence pay some of the training costs.

A Data Appendix

A.1 IAB Data

We use data from the German Institute for Employment Research (IAB) for the years 1975-1995. The basis of the IAB employment subsample is the integrated notifying procedure for health insurance, statutory pension scheme and unemployment insurance which is regulated through German legislation. The procedure requires that employers report all information of their employees registered by the social security system to the social security agencies. Employers have to notify the beginning and the end of an employment spell and have to give an annual notification for each employee. The employment statistics include all employees obliged to pay social insurance contributions. The employment statistics do not include, among others, civil servants, family workers, those in marginal employment, and students enrolled in higher education (Cramer (1985)). For 1995, the employment statistics cover nearly 79.4% of all employed persons in Western Germany (Bender, Haas, and Klose (2000)).

The notification provides information on individual characteristics as gender, year of birth, number of children and qualifications. Furthermore it reports information on the employment including information on the occupational code, the occupational status, the establishment number of the employer with information on the size and the industry of the employer, and finally the gross earnings of the employee over the past employment spell which served as the basis for social security contributions. This information is passed on from the social insurance agencies to the Federal Employment Services and collected in the so called *historic file*. The IAB employment subsample is an anonymised 1% sample from the historic file. Details of the anonymisation procedure are described in Bender, Haas, and Klose (2000). Due to the fact that the information for East Germany is only available for the time after unification we use only the information of notifications for people working in Western Germany. The employment subsample contains a total of 7,847,553 notifications with 6,711,153 notifications for Western Germany. On the basis of the final notifications in each case, the file provides information of 483,327 Western Germans (Bender, Haas, and Klose (2000), p.2).

Apart from information in the historic file the IAB employment subsample contains information from two other data sources. The benefits recipients file contains person-related information on periods in which the Federal Employment Service paid benefits like the status of the unemployed and the type of benefit payments (unemployment benefit, unemployment assistance or maintenance payments for participating in training or re-training programs). But not all spells of registered non-employment were covered (Bender, Haas, and Klose (2000)). The second file which adds information

to IAB employment subsample is the establishment file. The file provides additional information on the notifying establishment as the date of birth and death of the establishment as well as generated information on the pattern of skill levels of employees within the establishment.

A.2 QaC Data

We use the 1991/1992 Qualifications and Careers Survey (QaC) data of the Bundesinstitut fuer Berufsbildung (BiBB) and the Institut fuer Arbeitsmarkt- und Berufsforschung (IAB). This data set asks a random sample of the working population (excluding persons currently enrolled in an apprenticeship, people on military or civil service, and helping family members) about their qualification, job career, workplace conditions, job satisfaction as well as activities in formal and informal education. Similar surveys exist for the years 1979 and 1985/86, but all surveys are cross-sections. The data set collects in total information on 34277 individuals - 24090 for West Germany and 10187 for East Germany. We choose a sub-sample of male employees residing in West Germany, without an academic degree, who completed an apprenticeship, which lasts longer than 24 months. In addition we restrict our focus on employees younger than 35 years of age.

References

- ACEMOGLU, D., AND J.-S. PISCHKE (1998): "Why Do Firms Train? Theory and Evidence," *Quarterly Journal of Economics*, 113(1), 79–119.
- ANGRIST, J. D., AND A. B. KRUEGER (1991): "Does Compulsory Schooling Affect Schooling and Earnings?," *Quarterly Journal of Economics*, 106, 979–1014.
- BAILY, N. M., G. BURTLESS, AND R. LITAN (1992): *Growth with Equity: Economic Policymaking for the Next Century*. Washington DC: Brookings Institution.
- BASSI, L. J., AND J. LUDWIG (2000): "Schol-to-Work Programs in the United States: A Multi-Firm Case Study of Training, Benefits and Costs," *Industrial and Labor Relations Review*, 53(2), 219–239.
- BECKER, G. S. (1975): *Human Capital*. University of Chicago Press, Chicago, 2nd edn.
- BENDER, S., C. DUSTMANN, D. MARGOLIS, AND C. MEGHIR (1999): "Worker Displacement in France and Germany," The Institute for Fiscal Studies: Working Paper Series: No. W99/14.
- BENDER, S., A. HAAS, AND C. KLOSE (2000): "IAB Employment Subsample 1975-1995: Opportunities for Analysis Provided by the Anonymised Subsample," IZA Discussion Paper No. 117.
- BERG, P. B. (1994): "The German Training System," in *Britain's Training Deficit*, ed. by R. Layard, K. Mayhew, and G. Owen, pp. 282–313. Aldershot, Avebury.
- BURTLESS, G. (1994): "Comment on Heckman et al, "US Education and Training Policy: A Re-Evaluation of the Underlying Assumptions Behind the New Consensus"," in *Labor Markets, Employment Policy, and Job Creation*, ed. by L. C. Solomon, and A. R. Levenson, pp. 123–139. Milken Institute Series in Economics and Education. Boulder and London: Westview Press in cooperation with the Milken Institute for Job and Capital Formation.
- CARD, D. (1999): "The Causal Effect of Education on Earnings," in *Handbook of Labour Economics Volume 3A*, ed. by O. Ashenfelter, and D. Card. Elsevier, Amsterdam.
- CASEY, B. (1991): "Recent Developments in the German Apprenticeship System," *British Journal of Industrial Relations*, 29(2), 205–222.

- CHANG, C., AND Y. WANG (1996): "Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revisited," *Journal of Labor Economics*, 14(3), 505–519.
- CRAMER, U. (1985): "Probleme der Genauigkeit der Beschäftigtenstatistik," *Allgemeines Statistisches Archiv*, 69, 56–68.
- DUSTMANN, C., AND C. MEGHIR (1999): "Wages, Experience and Seniority," Institute for Fiscal Studies (IFS) Working Paper 99/1.
- FEDERAL MINISTRY OF EDUCATION AND SCIENCE (1992): *Vocational Training in the Dual System in the Federal Republic of Germany* Federal Ministry of Education and Science, Bonn.
- HARHOFF, D., AND T. J. KANE (1993): "Financing Apprenticeship Training: Evidence from Germany," NBER Working Paper No. 4557.
- (1997): "Is the German Apprenticeship System a Panacea for the US Labour Market?," *Journal of Population Economics*, 10, 171–196.
- HECKMAN, J. J. (1999): "Doing It Right: Job Training and Education," *The Public Interest*, Spring, 86–107.
- HECKMAN, J. J., R. L. ROSELIUS, AND J. A. SMITH (1994): "'U.S. Education and Training Policy: A Re-Evaluation of the Underlying Assumptions Behind the 'New Consensus.'", in *Labor Markets, Employment Policy, and Job Creation.*, ed. by L. C. Solmon, and A. R. Levenson, pp. 83–121. Milken Institute Series in Economics and Education. Boulder and London: Westview Press in cooperation with the Milken Institute for Job and Capital Formation.
- HUGHES, K. L., T. R. BAILEY, AND M. J. MECHUR (2001): "School-to-Work: Making a Difference in Education: A Research Report to America," Institute on Education and the Economy, Teachers College, Columbia University.
- KIM, S. (1989): "Labor Specialization and the Extent of the Market," *The Journal of Political Economy*, 97(3), 692–705.
- KRUEGER, A., AND C. ROUSE (1998): "The Effect of Workplace Education on Earnings, Turnover, and Job Performance," *Journal of Labor Economics*, 16(1), 61–94.
- LYNCH, L. M. (1992): "Private Sector Training and the Earnings of Young Workers," *American Economic Review*, 82(1), 299–312.
- MCCALL, B. P. (1990): "Occupational Matching: A Test of Sorts," *The Journal of Political Economy*, 98(1), 45–69.

- MILLER, R. (1984): "Job Matching and Occupational Choice," *The Journal of Political Economy*, 92(6), 1086–1120.
- MORTENSEN, D. (1986): "Job Search and Labour Market Analysis," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and R. Layard, pp. 849–919. Elsevier Science Publishers BV.
- NEAL, D. (1995): "Industry-Specific Human Capital: Evidence from a Sample of Displaced Workers," *Journal of Labor Economics*, 17(4), 237–261.
- (1999): "The Complexity of Job Mobility Amongst Young Men," *Journal of Labor Economics*, 17(2), 237–161.
- RYAN, P. (2001): "The School-to-Work Transition: A Cross-National Perspective," *Journal of Economic Literature*, 39(1), 34–92.
- SCOONES, D. (2000): "Matching and Competition for Human Capital," *Labour Economics*, 7, 135–152.
- SHAPIRO, C. (1986): "Investment, Moral Hazard and Occupational Licensing," *Review of Economic Studies*, 53, 843–862.
- SHAW, K. L. (1987): "Occupational Change, Employer Change, and the Transferability of Skills," *Southern Economic Journal*, 53, 702–719.
- STEEDMAN, H. (1993): "The Economics of Youth Training in Germany," *The Economic Journal*, 103(420), 1279–1291.
- STEEDMAN, H., A. GREEN, O. BETRAND, A. RICHTER, M. RUBIN, AND K. WEBER (1997): "Assessment, Qualifications and Standards: The UK Compared to France, Germany, Singapore and the US," London.
- STEEDMAN, H., F. GREEN, AND P. RYAN (1998): "Apprenticeship - a Strategy for Growth," London.
- STEVENS, M. (1994): "A Theoretical Model of on-the-Job Training with Imperfect Competition," *Oxford Economic Papers*, 46, 537–562.
- TOPEL, R. H., AND M. P. WARD (1992): "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 107(2), 439–479.
- WERWATZ, A. (1998): "Mobility After Apprenticeship - How Effective is German Apprenticeship Training?," SFB 373 Discussion Paper 97-55, Humboldt University, Berlin.
- WINKELMANN, R. (1994): "Training, Earnings Mobility and Skill Acquisition," CEPR Discussion Paper No. 982.

Table 1: Descriptive Statistics

| | 1985 | 1987 | 1989 | 1991 | 1993 | 1995 |
|-----------------------|--------|--------|--------|--------|--------|--------|
| Daily Wage (1995 DM) | 102.6 | 114.1 | 122.1 | 131.9 | 136.6 | 139.9 |
| Age | 23.1 | 24.2 | 25.4 | 26.6 | 27.9 | 29.2 |
| Experience | 4.82 | 5.70 | 6.74 | 7.76 | 8.96 | 10.17 |
| Tenure | 2.49 | 2.85 | 3.21 | 3.64 | 4.29 | 4.79 |
| Apprenticeship | 0.772 | 0.791 | 0.785 | 0.798 | 0.822 | 0.830 |
| Meister Qualification | .00690 | 0.0103 | 0.0141 | 0.0178 | 0.0221 | 0.0242 |
| N | 23725 | 28919 | 34279 | 37670 | 37981 | 38765 |

Table 2: Descriptive Statistics for Sample of Workers Inside 3-digit Apprenticeship Occupation Prior to Separation

| | 'Quits' | 'Displaced' | 'Close 2' | 'Close 1' |
|--------------------------------------|---------|-------------|-----------|-----------|
| Pre-Separation Tenure | 2.26 | 1.41 | 1.65 | 1.87 |
| Age | 24.61 | 24.34 | 24.54 | 24.68 |
| Experience | 5.47 | 5.23 | 5.46 | 5.74 |
| Pre-Separation Wage | 114.82 | 102.98 | 105.12 | 105.34 |
| Average Length of Unemployment (Yrs) | --- | 0.296 | 0.309 | 0.314 |

Notes: See text for definitions of 'Quit', 'Displaced', 'Close 2' and 'Close 1'. Each column is based on sample of workers observed in the apprenticeship occupation pooled across years 1984-1995.

Table 3: Occupational Mobility and Firm Mobility for Prominent Apprenticeship Occupations

| Rank by number of observations in sub-sample | Occupation | IABS 3-digit code | Observations in sub-sample 1984-1995* | Firm changes | Move out of apprenticeship occupation at | | |
|--|---|-------------------|---------------------------------------|--------------|--|-----------------|-----------------|
| | | | | | 3 - digit level | 2 – digit level | 1 – digit level |
| 1 | car mechanic | 281 | 29409 | 0.8111 | 0.0153 | 0.1129 | 0.5368 |
| 2 | electronic plumber | 311 | 23894 | 0.7402 | 0.0435 | 0.0211 | 0.3847 |
| 3 | clerk, office worker | 781 | 18138 | 0.6604 | 0.0124 | 0.1783 | 0.2808 |
| 4 | locksmith (without further specification) | 270 | 16222 | 0.6040 | 0.0960 | 0.1581 | 0.3869 |
| 5 | joiner | 501 | 14813 | 0.7296 | 0.0140 | 0.0364 | 0.3613 |
| 7 | plumber | 262 | 14734 | 0.7278 | 0.0324 | 0.0776 | 0.3642 |
| 8 | bricklayer | 441 | 11612 | 0.7193 | 0.0237 | 0.1020 | 0.3337 |
| 9 | toolmaker | 291 | 9637 | 0.5564 | 0 | 0.2036 | 0.3701 |
| 10 | painter, varnisher | 511 | 9571 | 0.7712 | 0.0476 | 0.0415 | 0.3706 |

Notes: The ranking as well as the percentage of respective movers remain mostly the same when we rank the frequency of all observations in the sample from 1975-1995.

Table 4: Returns to Apprenticeship Training

| | 1985 | 1987 | 1989 | 1991 | 1993 | 1995 |
|--|---------------|----------|----------|----------|---------|----------|
| Panel A: Return to 'Intensive' Apprenticeship Training | | | | | | |
| Apprenticeship Inside Occupation | 0.122 | 0.127 | 0.166 | 0.144 | 0.155 | 0.163 |
| Panel B: Estimates of 'Standard' Return to Apprenticeship Training | | | | | | |
| Experience | 0.0958 | 0.113 | 0.104 | 0.119 | 0.106 | 0.0988 |
| Experience ² | -0.00263 | -0.00468 | -0.00370 | -0.00443 | -0.0037 | -0.00325 |
| Tenure | 0.00419 | 0.009 | 0.00879 | 0.0104 | 0.0124 | 0.0140 |
| | | | | | | |
| Apprenticeship | 0.168 | 0.167 | 0.213 | 0.195 | 0.189 | 0.184 |
| Meister | <i>-0.019</i> | 0.159 | 0.152 | 0.176 | 0.144 | 0.155 |
| | | | | | | |
| N | 22591 | 27578 | 32516 | 35739 | 35941 | 37140 |

Notes: Estimates in the first panel are based on the same equation as those presented in the first panel, with the apprenticeship variable referring only to those with apprenticeship inside the occupation worked in. All estimates with the exceptions of those in italics significant at the 1% level, where t-ratios based on robust standard errors.

Table 5: Returns Broken Down by Training Firm Size

| Training Firm size | Propn with good Math. scores | Propn trained in this size of firm | Return Inside Training Occupation | Propn of Time spent Training | Estimated Annual Return to Training |
|--------------------|------------------------------|------------------------------------|-----------------------------------|------------------------------|-------------------------------------|
| 1 | 0.110 | 0.023 | 0.0726 | 0.56 | 4.71% |
| 2-9 | 0.107 | 0.225 | 0.0904 | 0.56 | 5.87% |
| 10-19 | 0.136 | 0.141 | 0.0900 | 0.61 | 5.37% |
| 20-49 | 0.136 | 0.148 | 0.109 | 0.61 | 6.50% |
| 50-99 | 0.136 | 0.089 | 0.154 | 0.69 | 8.12% |
| 100-499 | 0.154 | 0.168 | 0.202 | 0.69 | 10.65% |
| 500-999 | 0.160 | 0.063 | 0.190 | 0.81 | 8.53% |
| >1000 | 0.172 | 0.144 | 0.270 | 0.81 | 12.12% |

Notes: Data in the second column from Harhoff and Kane (1993), Table 7. Their table uses an identical size breakdown with the two exceptions: in their Table, rows 1 and 2 correspond to firms of size 1-4 and 4-9 respectively. Secondly, their data aggregates firm sizes 10-49. We assume the figures for sizes 10-19 and 20-49 are identical. Figures reported in the third and fourth columns are based on data pooled over the years 1984 to 1995. The return inside the apprenticeship occupation is derived from an equation identical to that used to estimate these returns in Table 3, except that the apprenticeship variable is interacted with firm size and year dummies are included (sample size 374,710). All estimates significant at the 1% level. Estimates of the proportion of apprenticeship time spent training are based on unpublished results from a study on the costs of apprenticeship for firms in West Germany in 1991, by the Federal Institute for Vocational Education (BiBB). We thank Ursula Beicht of the BiBB for making this information available. Estimates of the rate of return are calculated by dividing the estimated coefficient by the average numbers of years spent training and the estimated proportion of apprenticeship time spent training (inside or outside of the training firm). The average numbers of years spent training is assumed as 2.75 years.

Table 6: Wage Penalty to Moving Out of the Apprenticeship Occupation

| | ‘Quits’ | ‘Displaced’ | ‘Close 2’ | ‘Close 1’ |
|---|--------------------------------------|--------------------------------------|-------------------------------------|------------------------------------|
| Panel A: Wage Penalty to Leaving the Apprenticeship Occupation | | | | |
| | --- | -0.132 ^{***} (0.0318) | -0.107 (0.111) | -0.352 ^{***} (0.0134) |
| Change in Experience | 0.116 ^{***} (0.00714) | 0.137 ^{***} (0.00785) | 0.0899 ^{***} (0.0288) | 0.174 ^{***} (0.0552) |
| Change in Experience ² | -0.00488 ^{**} (0.000539) | -0.00691 ^{**} (0.000622) | -0.00373 ^{**} (0.00209) | -0.00579 [*] (0.00399) |
| Change in Tenure | 0.00268 ^{***} (0.000904) | 0.00787 ^{***} (0.00323) | -0.00149 (0.0114) | -0.0101 (0.0175) |
| Move Out | -0.000552 (0.00619) | -0.0165 [*] (0.0117) | 0.0300 (0.0381) | -0.0349 (0.061) |
| N | 14279 | 4893 | 408 | 161 |
| Panel B: Wage Penalty to Leaving the Apprenticeship Occupation by Distance (measured by Occupational Codes) | | | | |
| 3-digit only | 0.000641 (0.0128) | 0.0208 (0.0242) | 0.374 ^{***} (0.122) | 0.360 ^{***} (0.136) |
| 2-digit only | 0.0484 ^{***} (0.0124) | 0.0424 ^{***} (0.0216) | 0.0690 [*] (0.0632) | 0.033 (0.0965) |
| 1-digit | -0.0128 ^{**} (0.00738) | -0.0352 ^{***} (0.0132) | -0.0114 (0.0420) | -0.103 [*] (0.0708) |
| Panel C: Wage Penalty to Leaving the Apprenticeship Occupation and Switching 2-digit Industry (occupational distance measured by Industrial and Occupational Codes) | | | | |
| Move Out (All) | -0.0199 ^{***} (0.00730) | -0.0397 ^{***} (0.0124) | -0.0239 (0.0414) | -0.0619 (0.0705) |
| 3-digit only | -0.00235 (0.0196) | 0.00496 (0.0307) | 0.283 ^{***} (0.108) | 0.459 ^{***} (0.132) |
| 2-digit only | 0.0326 ^{**} (0.0164) | 0.0172 (0.0263) | -0.0359 (0.0738) | -0.106 (0.205) |
| 1-digit | -0.0320 ^{***} (0.00834) | -0.0541 ^{***} (0.0137) | -0.0400 (0.0449) | -0.0950 [*] (0.0764) |
| Panel D: Wage Penalty to Leaving the Apprenticeship Occupation where 2-digit Moves split according to Quartiles of data-generated Distance Measure | | | | |
| 2-digit – Quartile 1 | 0.0514 ^{***} (0.0110) | 0.0524 ^{***} (0.0220) | 0.0531 (0.0531) | 0.00153 (0.0721) |
| 2-digit – Quartile 2 | -0.00457 (0.0123) | -0.0000977 (0.0190) | 0.0217 (0.0646) | -0.0274 (0.114) |
| 2-digit – Quartile 3 | -0.0141 [*] (0.0120) | -0.0493 ^{***} (0.0194) | -0.0252 (0.0487) | -0.0515 (0.0811) |
| 2-digit – Quartile 4 | -0.0410 ^{***} (0.0134) | -0.0835 ^{***} (0.0210) | -0.0456 (0.0857) | -0.283 ^{***} (0.138) |

Notes: see text for definitions of ‘Quits’, ‘Displaced’, ‘Close 2’ and ‘Close 1’. Panel B, C and D estimated using the same equation as Panel A, with moves out of the apprenticeship occupation disaggregated by distance moved. See text for description of data-generated distance measure. Statistical significance at the 1% (5%, 10%) level denoted *** (**, *).

Table 7: How many Apprenticeship Skills used in the Current Job?

| | Very Few or None | Few | Some | Many | Very Many | N |
|------------------------------------|--|-------|-------|-------|-----------|------|
| | Panel A: Analysis of Skill Use according to whether Working in Apprenticeship Occupation | | | | | |
| All | 11.06 | 9.50 | 15.01 | 21.74 | 42.69 | 3317 |
| Working in Training Occupation | 0.87 | 3.80 | 10.79 | 25.87 | 58.68 | 1844 |
| Working Out of Training Occupation | 23.83 | 16.63 | 20.30 | 16.56 | 22.67 | 1473 |
| | | | | | | |
| | Panel B: Analysis of Movers according to Distance Measured by Codes | | | | | |
| 3-digit Level | 1.44 | 10.58 | 14.42 | 27.88 | 45.67 | 208 |
| 2-digit Level | 12.24 | 12.24 | 25.00 | 19.90 | 30.61 | 196 |
| 1-digit Level | 30.31 | 18.62 | 20.58 | 13.75 | 16.74 | 1069 |
| | Panel C: Analysis of Occupational Switchers and Industry Switchers (2-digit level) according to Distance Measured by Codes | | | | | |
| All | 33.41 | 20.26 | 21.14 | 12.92 | 12.27 | 913 |
| 3-digit Level | 1.43 | 24.29 | 17.14 | 30.00 | 27.14 | 70 |
| 2-digit Level | 19.15 | 17.02 | 26.60 | 15.96 | 21.28 | 94 |
| 1-digit Level | 38.18 | 20.29 | 20.83 | 10.95 | 9.75 | 749 |
| | Panel D: Analysis of Movers according to Distance Measured by Data | | | | | |
| 2-digit - Quartile 1 | 21.35 | 16.73 | 24.56 | 16.37 | 21.00 | 281 |
| 2-digit - Quartile 2 | 18.04 | 15.29 | 17.43 | 18.65 | 30.58 | 327 |
| 2-digit - Quartile 3 | 16.62 | 12.88 | 20.25 | 18.71 | 31.90 | 326 |
| 2-digit - Quartile 4 | 20.54 | 14.50 | 23.26 | 16.31 | 25.38 | 331 |

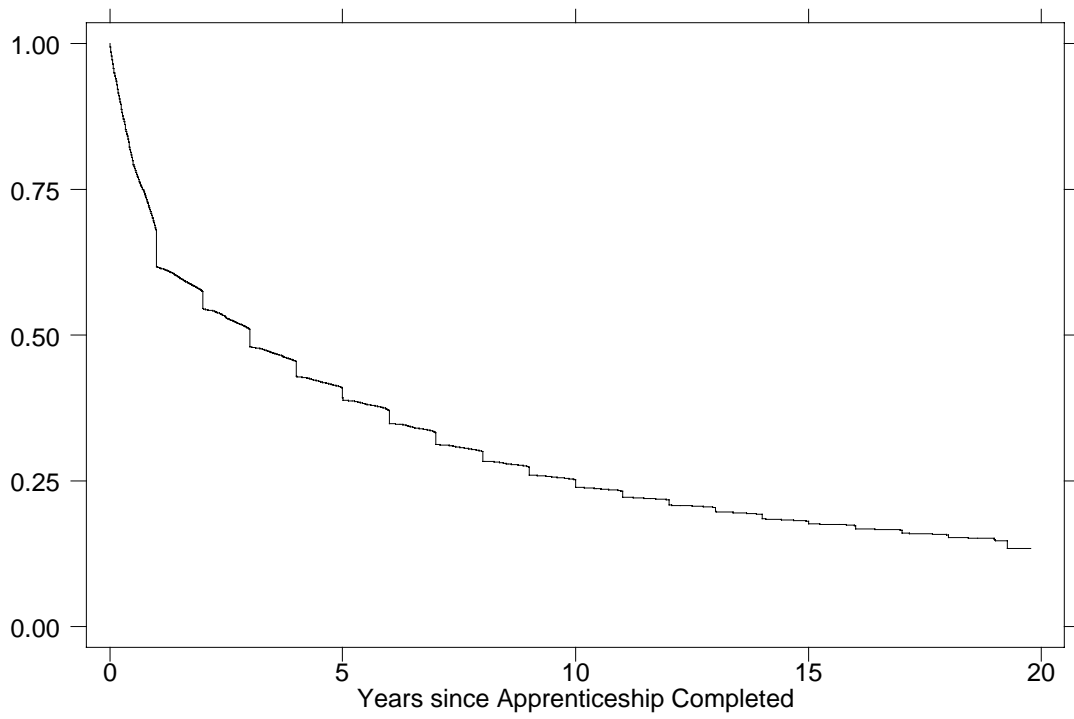
Notes: Column headings are answers to question ‘How much of your occupational knowledge and skills, which you have obtained in your apprenticeship, can you actually use in your current job?’ from QaC data. See text for details.

Table A1: Industry Classification as Used in Empirical Analysis

| | Industry | Two digit ES-classification* | PSID equivalent |
|----|---|------------------------------|-------------------|
| 1 | Agriculture, Forestry | 00-02 | 17-27 |
| 2 | Fishing | 03 | 28 |
| 3 | Energy | 04 | 377, 378, 467-479 |
| 4 | Mining | 05-08 | 47-57 |
| 5 | Chemical | 09 | 347, 357-369 |
| 6 | Synthetics | 10-13 | 348, 349, 379-387 |
| 7 | Earth/Clay/Stone | 14-16 | 119-138 |
| 8 | Iron/Steel | 17-21 | 139-169 |
| 9 | Mechanical Engineering | 22-32 | 177-198, 219-238 |
| 10 | Electrical Engineering | 33-39 | 199-209, 239-259 |
| 11 | Wood/Paper/Printing | 40-44 | 107-118, 328-339 |
| 12 | Clothing/Textiles | 45-53 | 307-327, 388-398 |
| 13 | Food industry | 54-58 | 268-299 |
| 14 | Construction/ Construction related | 59-61 | |
| 15 | Trade | 62 | 507-588, 607-698 |
| 16 | Train system | 63 | 407 |
| 17 | Postal system | 64 | 447-449, 907 |
| 18 | Other transport | 65-68 | 408-429 |
| 19 | Financial institutions | 69 | 707-709, 717 |
| 20 | Restaurants, Service Industry | 70-73 | 777-809 |
| 21 | Education/Sport | 74-77 | 857-869 |
| 22 | Health Service | 78 | 828-848 |
| 23 | Legal Services | 79 | 718, 849 |
| 24 | Other Services | 80-86 | 727-759, 888-897 |
| 25 | Non Profit (Voluntary/Church/Private Households) | 87-90 | 877-887, 769 |
| 26 | Public Institutions (Regional Authority/ Social Security) | 91-94 | 917-937 |

Notes: Classification in the list of industries used for the statistics of the Federal Employment Service in Germany (1973 edition).

Figure 1: Probability of Staying in Apprenticeship Occupation



IZA Discussion Papers

| No. | Author(s) | Title | Area | Date |
|-----|--|--|------|-------|
| 364 | M. Jääntti S. P. Jenkins | Examining the Impact of Macro-Economic Conditions on Income Inequality | 3 | 09/01 |
| 365 | H. S. Nielsen M. Rosholm N. Smith L. Husted | Qualifications, Discrimination, or Assimilation? An Extended Framework for Analysing Immigrant Wage Gaps | 1 | 09/01 |
| 366 | M. C. Regets | Research and Policy Issues in High-Skilled International Migration: A Perspective with Data from the United States | 1/5 | 09/01 |
| 367 | C. Dustmann | Parental Background, Primary to Secondary School Transitions, and Wages | 5 | 09/01 |
| 368 | J. Angrist | How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation | 5 | 09/01 |
| 369 | A. S. Kalwij | Individuals' Unemployment Durations over the Business Cycle | 3 | 09/01 |
| 370 | A. S. Kalwij | Individuals' Unemployment Experiences: Heterogeneity and Business Cycle Effects | 3 | 09/01 |
| 371 | S. C. Wolter A. Zbinden | Rates of Return to Education: The View of Students in Switzerland | 5 | 09/01 |
| 372 | J. Konings H. Lehmann | Marshall and Labour Demand in Russia: Going Back to Basics | 4 | 09/01 |
| 373 | S. J. Trejo | Does the Statutory Overtime Premium Discourage Long Workweeks? | 1 | 10/01 |
| 374 | G. J. van den Berg B. van der Klaauw | Counseling and Monitoring of Unemployed Workers: Theory and Evidence from a Controlled Social Experiment | 6 | 10/01 |
| 375 | J. A. Dunlevy W. K. Hutchinson | The Pro-Trade Effect of Immigration on American Exports During the Late Nineteenth and Early Twentieth Centuries | 1 | 10/01 |
| 376 | G. Corneo | Work and Television | 5 | 10/01 |
| 377 | S. J. Trejo | Intergenerational Progress of Mexican-Origin Workers in the U.S. Labor Market | 1 | 10/01 |
| 378 | D. Clark R. Fahr | The Promise of Workplace Training for Non-College-Bound Youth: Theory and Evidence from German Apprenticeship | 1 | 10/01 |