Matching, Screening and Firm Investment in General Training: Theory and Evidence

Damon Clark*

Nuffield College, Oxford and
Centre for Economic Performance (CEP)†

Abstract

When job matching is important, we show that firms will pay for general training under very weak conditions. The key ingredient in our model is the idea that it is more costly to screen skilled workers than it is to screen unskilled ones. In equilibrium, this 'softens' competition for trained workers, allowing firms to recoup training investments. We apply our model to a classic case of firm investment in general training - German Apprenticeship Training - and show that a key prediction of our model that is not shared by other models is strongly supported in the data.

Keywords: General Training; Human Capital; Auctions; Wages.

JEL Classification Numbers: J24, J41, D44, J31.

---

*This paper has greatly benefited from discussions with Steve Nickell and Margaret Stevens. Alan Manning, Steve Pischke and participants at the LSE and IZA Labour Market Workshops also provided extremely helpful comments. Thanks are also due to Rene Fahr for invaluable help with the IAB data.

†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

1 Introduction

Although it is difficult to determine precisely how the costs of training workers are shared between worker and firm, there are a number of prominent cases in which firms appear to bear at least some of the costs of 'general' training.[1] To take one celebrated example, German apprentices receive two to three years worth of general skills training within a firm, at substantial cost to the firms providing the training.[2] Other examples include the provision of general skills training by temporary help supply firms described by Autor (2001).

To economists brought up to believe that firms will not pay for general training of this type, these investments represent something of a puzzle. As Becker (1962) first pointed out, in competitive labour markets, any attempt to recoup training costs by paying trainees less than the value of their marginal product will result in them leaving to earn their full value in another firm. The simple fact that firms do subsidise general training is therefore of immediate theoretical interest: if the competitive paragidm is an inadequate description of the labour market, then how does this market operate?

The motives behind firm investment in general skills may also have important policy implications. In the case of German Apprenticeship, whilst it is often argued that other countries such as the US would benefit from apprenticeship-type schemes (see for example Baily, Burtless, and Litan (1992)), it is hard to make this case without a more complete understanding of the reasons why German firms choose to pay for these apprenticeships. As Harhoff and Kane (1997) point out "(T)he structure of in-

[1] See Bishop (1996) for an in-depth discussion of the evidence on this point.

[2] We will return to the issue of how general is GAT and look in more detail at the cost-sharing arrangements in section 4.

centives undergirding the German Apprenticeship System is not well understood - even in Germany" (p.172).

The objective of this paper is to understand why firms make investments in general training, and apply our theory to the case of German Apprenticeship. In fact, significant strides have already made in this area. In particular, Acemoglu and Pischke (1999) spell out the key condition under which firms pay for general training and describe a number of mechanisms that satisfy it. Termed 'wage compression', the condition says that for firms to invest in general training, the rents that firms earn on workers must be an increasing function of training.

The model set out in this paper meets this condition and does, we believe, rely on a much weaker set of assumptions than the other mechanisms so far proposed. The basic ingredients of the model are first, that workers are heterogenous with respect to their productivity in different firms, and secondly, that this idiosyncratic match between worker and firm is not immediately obvious to firms. Instead it must be inferred from the worker's CV, from interviews and so on. Since this 'screening' process is costly, in equilibrium, a finite number of outside firms will post offers for workers. This result, together with match heterogeneity, ensures that the worker's match product is highest in the 'incumbent' firm with strictly positive probability. In turn, this implies that the incumbent firm can pay workers a wage less than their product and earn rents on these workers. That these rents are increasing in the level of training is a byproduct of the fact that training increases the costs of screening for other firms, thereby increasing the incumbent firm's monopsony power and their expected rents.

The importance of matching in the labour market is well known, and perhaps best illustrated by Topel and Ward (1992). They find that 70% of the wage growth of young workers between the ages of 18 and 30 is accounted for by job shopping. The importance of hiring costs, and the fact that they are increasing in the skill level of workers is equally compelling. As noted in “expenses for recruiting, including agency fees or advertising, and for screening, are higher for the skilled employee” (p.184). Indeed, they report figures showing that the average hiring costs for professional, managerial and technical workers are five times as great as those for skilled workers and twelve times those for unskilled workers.

To illuminate the assumptions behind the mechanism that we suggest, we can nest the various models of general training investment within the following equation describing the product of worker i in firm j:

$$v_{ij} = \tau_j \eta_j + \xi_{ij}$$

Under this formulation, worker j’s expected product in firm i is a multiplicative function of his training τ_j and his ability η_j. His actual product is an additive function of expected product and a random zero-mean worker-firm match term ξ_{ij}.

Perhaps the most prominent model of firm investment in general training is Acemoglu and Pischke (1998). They assume that training τ_j is observed by all agents, but that ability levels η_j are not. There are no matching considerations in the model. The model’s key feature is that adverse selection on the outside labour market allows firms to earn rents on workers. The multiplicative relationship between training levels and ability then ensures that rents are increasing in training levels. A similar mechanism drives the model of Autor (2001), and allows temporary help supply firms to invest in the general training of their workers. The third model in the asymmetric information class is that of Chang and Wang (1996). Unlike Acemoglu and Pischke (1998) and Autor (2001) however, they consider asymmetric information regarding the level of training τ_j rather than ability η_j. In fact, Chang and Wang (1996) also give a role to match productivity ξ_{ij}, although it is not central to their results relating to general training.

A related paper that deals with matching is Scoones (2000).[3] Like Chang and Wang (1996), the match product is an 'experience good' for the incumbent firm and an 'inspection good' for other firms. That is, match product is initially unknown, is revealed to the incumbent firm after the worker has spent some time in the labour market and can then be inspected by all other firms. As regards training τ_j and ability η_j, Scoones (2000) implicitly assumes that these are common knowledge.

The matching technology employed in our model is identical to that of Scoones (2000) and we also assume that training levels τ_j and ability η_j are common knowledge. The difference however is that we assume that the match product can only be inferred by other firms at a cost k. This is the cost of screening workers. Since we model the competition for workers as a first-price auction, the screening cost acts as an entry cost to the auction. As the entry cost tends to zero, the labour market becomes competitive. The seller in the auction is the worker, but the reservation price is set

[3]The paper asks whether workers will invest the socially optimal levels of general and specific capital. The answers are 'yes' and 'no'. The intuition for the latter is that social welfare is maximised when workers invest in specific training according to the probability of them remaining with the training firm (greater than one-half). However, workers actually invest according to the probability of them leaving the training firm (less than one-half), since workers only benefit from specific training to the extent that it drives up their outside option when they leave the incumbent firm.

(strategically) by the incumbent firm. In this sense, the model is related to other auction models with endogenous entry[4] as well as the Milgrom and Weber (1982) limit-pricing model.

Having set out the model in the first part of the paper, the second part of the paper is devoted to an empirical examination of one of its key implications. A unique feature of our model is that trained workers *increase* their wages upon leaving the firm. This occurs because workers will always work for the winning bidder, whether this be the incumbent (training) firm, or another firm on the outside labour market. This implication is in stark contrast to that of the two most prominent explanations for firm investment in German Apprenticeship Training (GAT) - the asymmetric information model of Acemoglu and Pischke (1998) and models in which general and firm-specific skills are complementary.[5] In the Acemoglu and Pischke (1998) model, there is a unique, competitive outside wage offer which the incumbent firm will always more than match for those workers it chooses to keep. Hence movers are expected to earn less than stayers. Similarly, in all models of firm-specific capital, the rents generated by the firm-specific capital are shared between the training firm and the trainee, and are lost when workers leave the training firm.

To preview our results, we find strong support in favour of our model and against these other explanations. For a flavour of our results, Figure 1 tracks the regression-adjusted wages of five different groups of movers observed after the end of apprentice-

[4]Perhaps the most closely related is Fishman (1988) who considers the problem faced by the first bidder in a takeover bidding war. As in our model, the first bidder makes a bid that will deter other firms from entering the auction.

[5]The complementarity of general and firm-specific skills has been suggested as an explanation for firm investment in German Apprenticeship Training by, *inter alia*, Soskice (1994). Since apprenticeship training is observed by all parties in Germany, the model of Chang and Wang (1996) has never been considered in this context. In any case, its implications are identical to those of Acemoglu and Pischke (1998).

ship relative to the base group of stayers who remain with the training firm across all five post-apprenticeship observations (and whose regression-adjusted wages are normalised to zero). At the first post-apprenticeship observation (roughly one year after the end of apprenticeship), those workers that have already moved (MOVE01) earn similar wages to the stayers. The group that move between the first and second post-apprenticeship observation (MOVE12) initially earn significantly lower wages than both groups, although by leaving the training firm, they dramatically increase their relative wages. The same pattern is observed for other groups of movers.

We do not interpret our results as an outright rejection of all other models of firm investment, since we believe for example, that certain segments of the labour market (such as the market for low-skilled workers in which temporary help supply firms typically operate) are likely to be characterised by a strong form of asymmetric information. Instead, we argue that in more skilled labour markets such as the market for German apprentices, where training is certificated and trainees are fairly well educated, matching and screening may be a more plausible explanation for firm investment.

The paper is structured as follows: Section 2 sets out our model in more detail and section 3 contrasts the empirical implications of this model with those derived from the other models described above. Section 4 provides a brief overview of German Apprenticeship Training (GAT) and section 5 describes the data used. The tests are implemented in section 6, and our conclusions are presented in section 7.

2 A Model of Training

The purpose of the model is to explain why firms may invest in general training. The essential ingredient that allows firms to invest in general training is a screening cost that is increasing in the skill level of workers. Whilst we argued in the Introduction that this is true in practise, it is not clear theoretically why the cost of screening a skilled worker should exceed the cost of screening an unskilled worker. The point of the model is to present a setting in which it arises naturally. In particular, in a world in which there is match heterogeneity, the screening cost is viewed as the cost of inspecting the match product that must be incurred by firms on the outside labour market. In equilibrium, this 'softens' the competition for skilled workers faced by the incumbent firm, allowing the incumbent firm to earn rents on skilled workers and thereby pay for training.

2.1 Basic Set-Up

We begin by laying out the assumptions of the model:

A1 The economy consists of a large number of infinitely-lived firms. These firms produce homogenous output taking the price as given (normalised to unity) and produce according to a linear production technology.

A2 Workers are assumed to live for two periods and are characterised by ability level μ_j and training level τ_j. Ability is fixed across periods and workers are born untrained ($\tau_j = 0, \forall j)$.

A3 The productivity of worker j in firm i, v_{ij}, can be described as follows:

$$v_{ij} = \alpha_j + \zeta_{ij}$$

where $\alpha_j = \alpha(\mu_j, \tau_j)$ is the general skill level of worker j and ξ_{ij} is the quality of match between firm i and worker j. We assume that ζ_{ij} is a random draw from a uniform distribution with mean zero and support $(\underline{\xi}, \overline{\xi})$, and we assume that for worker j, these draws are independent across firms. The uniform assumption is not essential, but will simplify some of the analysis. Notice that training is 'general' in the sense that it increases the worker's general skill level α_j independently of the firm that the worker is matched to.

2.2 Timing

The model proceeds in two periods (with no discounting between periods):

1. Period 1

   i Firms have no information about school-leavers other than their expected productivity level α_j. Hence firms attract workers by posting entry wages indexed by α_j. By symmetry, the equilibrium entry wage will be identical across firms, and since workers are assumed to care only about wages, they choose the initial firm (which we will call the incumbent firm) at random.

   ii Training takes place.

   iii The incumbent firm (but no other agents) observe the quality of their match to worker j. They do not however observe worker j's potential match quality in other firms.

2 Period 2

i The incumbent firm makes a wage offer (w_0) to the worker

ii Before accepting the offer, the worker attempts to obtain a better offer from other firms. To do this, she sends a CV to every firm describing α_j, and the minimum wage offer that she prepared to accept, w^{MIN}.

iii Upon receiving the worker's CV, other firms have to decide whether or not to inspect it. The cost of doing so is k. Once inspected, they can decide whether to make an offer to the worker, and if so, what the offer should be.

iv After receiving offers from all firms, the worker moves to the firm offering the highest wage provided this wage exceeds w_0. Otherwise, she stays with the incumbent firm and earns w_0.

2.3 Computing Equilibria

Our strategy for computing the equilibria of the model is as follows. First, we show that the minimum acceptable wage demanded by the worker will be exactly the initial wage offer made by the incumbent firm. This simplifies the second period of the model so that there are now only three subperiods - the incumbent firm makes a wage offer, other firms make offers and the worker accepts the highest offer received.

We analyse the subgame-perfect equilibrium of the game. Hence we begin by analysing the entry process conditional on the incumbent's wage offer w_0 and the entry cost k. A zero profit condition gives the number of firms inspecting the match, n, as a decreasing function of both the incumbent's initial wage offer and the inspection cost, $n(w_0, k)$. We then compute the intial wage offer that maximises the profit of the incumbent. This will take account of $n(w_0, k)$, and hence will be a function of match productivity and the inspection cost, $w_0^*(v_0, k)$. Given this function describing the optimal wage offer, we then analyse the firm's incentives to train in terms of the impact of training on (equilibrium) expected profits $\Pi^*(v_0, k)$.

2.4 Worker's Minimum Acceptable Wage Offer

The worker is acting as a seller in an auction with entry, where her own value of the item is w_0 - her utility if there is no sale. Hence we can use the following result, derived in Mcafee and Mcmillan (1987):

Proposition 1 *(Mcafee and Mcmillan (1987)):The minimum acceptable wage reported by the worker will be the initial wage offer of the incumbent firm.*

Proof. The seller's (worker's) expected revenue is the winning bidder's valuation minus the expected profit of the n bidders that enter $E(v) - nk$. Hence for a given number of bidders, n, and an inspection cost k, the seller (worker) maximises revenue by setting a reserve price equal to his own valuation. See Mcafee and Mcmillan (1987), p.344. ■

Hence with free entry, and in contrast to the fixed-n case, the seller will not distort the reservation price. Or in this context, the minimum acceptable offer demanded by the worker will simply be the offer made by the incumbent firm. Since the incumbent firm is aware of this fact, they know that their initial wage offer will be perfectly transmitted to their potential rivals and so they will set this wage strategically.

2.5 Entry

From equation (A3) of Appendix A, the expected profit of an entrant *after* it has drawn its valuation v (ignoring subscripts i and j) can be written as:

$$\Pi(\xi) = \int_{\tilde{w}_0}^{\xi} G(x)dx$$

where $\tilde{w}_0 = (w_0 - \alpha_j)$ and $G(x)$ is the probability that the firm's match (x) is the highest of all the n potential bidder's matches. Hence $G(x) = F(x)^{n-1}$. Before it draws its valuation (and therefore, before it pays the cost of inspecting the match), the expected profit of a potential entrant is:

$$\begin{aligned}
\Pi^e &= E[\int_{\tilde{w}_0}^{\xi} G(x)dx] - k \\
&= \int_{\tilde{w}_0}^{\bar{\xi}} [\int_{\tilde{w}_0}^{\xi} G(x)dx] f(\xi)d\xi - k
\end{aligned}$$

Integrating by parts, this can be shown to be:

$$= \int_{\tilde{w}_0}^{\bar{\xi}} [F(\xi)^{n-1}(1 - F(\xi))]d\xi - k$$

Hence the free entry (pure strategy) equilibrium number of entrants (n) is determined by the following zero-profit condition:[6]

$$k = \int_{\tilde{w}_0}^{\bar{\xi}} [F(\xi)^{n-1}(1 - F(\xi))]d\xi \tag{1}$$

Equation (1) defines the number of entrants n as an implicit function of the incumbent's wage offer $\tilde{w}_0$ and the cost of inspecting the match, k. We assume that k is sufficiently small that for $\tilde{w}_0 = \underline{\xi}$, $n \geq 1$. In other words, the incumbent firm will always face at least one competititor when its wage offer net of general skills is equal to the support of the match distribution.[7]

2.5.1 Comparative Statics

We want to describe how n varies with the reservation price $\tilde{w}_0$ and the entry cost k. To do this, we take differentials using the previous equation to show that:

$$\begin{aligned} \frac{dn}{dk} &= \frac{1}{\int_{\tilde{w}_0}^{\bar{\xi}} [1 - F(\xi)] \log F(\xi) [F(\xi)]^{n-1} d\xi} \\ &< 0 \end{aligned} \tag{2}$$

[6] In fact we treat n as a continuous variable rather than an integer.

[7] We are assuming throughout that the outside firms must pay the cost k in order to inspect the match. Were these firms able to make 'blind' offers without having to inspect the match, competition amongst outside firms would ensure that these 'blind' offers were equal to the expected value of the match. Hence the equilibrium that we describe would refer only to the one-half of trained workers with matches better than expected match quality. Otherwise, the results are unchanged.

since $F(\xi) \leq 1$ hence $lnF(\xi) \leq 0$. In other words, as we would expect, a higher entry cost reduces the number of firms willing to inspect the match. Similarly:

$$
\begin{aligned}
\frac{dn}{d\tilde{w}_0} &= \frac{F(\tilde{w}_0)^{n-1}(1 - F(\tilde{w}_0))}{\int_{\tilde{w}_0}^{\bar{\xi}} [1 - F(\xi)] \log F(\xi) [F(\xi)]^{n-1} d\xi} \\
&< 0 \qquad (3)
\end{aligned}
$$

Intuitively, a higher wage offer by the incumbent firm reduces the number of firms willing to pay the inspection cost. This occurs first, because these firms realise that they are unlikely to be able to make an offer above this 'reservation price' and secondly, because even if they are, their expected profits from doing so are reduced.

2.6 Retention

When the firm sets the initial wage w_0 it trades off two forces: first, a higher wage reduces profits. Secondly, a higher wage reduces the probability of the worker leaving the firm. To understand the second effect, note that:

$$P[stay(\tilde{w}_0)] = F[\tilde{w}_0]^{n[\tilde{w}_0, k]}$$

Hence:

$$\frac{dP(stay)}{d\tilde{w}_0} = F[\tilde{w}_0]^n [\frac{dn}{d\tilde{w}_0} \ln F(\tilde{w}_0) + \frac{nF'(\tilde{w}_0)}{F(\tilde{w}_0)}] \qquad (4)$$

Equation (4) illustrates the two effects of an increase on $\tilde{w}_0$ on the probability of retaining workers: first, there is a direct effect which holds for fixed n (the second term inside the square brackets). Secondly, there is an indirect effect, in that a higher wage discourages entry of other firms and so reduces n. For future reference, we define:

$$
\begin{aligned}
H(\tilde{w}_0) &= \frac{P(stay)}{\frac{dP(stay)}{d\tilde{w}_0}} = \frac{1}{[\frac{dn}{d\tilde{w}_0}\ln F(\tilde{w}_0) + \frac{nF'(\tilde{w}_0)}{F(\tilde{w}_0)}]} \\
&> 0
\end{aligned}
$$

for $\tilde{w}_0 > \underline{\xi}$, since $\frac{dn}{d\tilde{w}_0} < 0$ and $\ln F(\tilde{w}_0) \leq 0$ and $\frac{F\prime(\tilde{w}_0)}{F(\tilde{w}_0)} > 0$.

2.7 Optimal Wage Offer

The profit of the incumbent firm can be written:

$$
\begin{aligned}
\Pi &= (v_0 - w_0)P(stay) \\
&= [(\alpha_j + \xi_0) - w_0]P(stay) \\
&= [\xi_0 - \tilde{w}_0]P(stay)
\end{aligned}
$$

The first-order condition for the maximisation of this function is:

$$\tilde{w}_0{}^* = \xi_0 - H(\tilde{w}_0{}^*) \tag{5}$$

where $H(\tilde{w}_0^*)$ is defined as above.

Proposition 2 *There exists a unique optimal wage offer for the incumbent firm* w_0^**, where* $\alpha_j + \underline{\xi} < w_0^* \leq v_0 \leq \alpha_j + \overline{\xi}$

Proof. We know that $H(w_0)$ is positive, and since all of the terms in $H(w_0)$ are continuous on the open interval $(\underline{\xi},\overline{\xi})$, $H(w_0)$ is also continuous on the open interval $(\underline{\xi},\overline{\xi})$. It can also be shown that as $w_0 \longrightarrow \underline{\xi}$ from above, $H(w_0) \rightarrow 0$. To prove this, note that $H(w_0) = \frac{1}{\frac{1}{F(w_0)}[(\frac{dn}{dw_0})\ln F(w_0)F(w_0)+nF'(w_0)]}$ and note that as $w_0 \longrightarrow \underline{\xi}$ from above, $F(w_0) \rightarrow 0$. From (3) the numerator of $(\frac{dn}{dw_0})$ tends to zero as $F(w_0) \rightarrow 0$, hence $(\frac{dn}{dw_0}) \rightarrow 0$ as $w_0 \longrightarrow \underline{\xi}$. Also, $\ln F(w_0)$F$(w_0) \rightarrow 0$ as $F(w_0) \rightarrow 0$, hence the term in square brackets tends to $nF'(w_0)$ as $w_0 \longrightarrow \underline{\xi}$. Since this is constant in the uniform case, and since $\frac{1}{F(w_0)} \rightarrow \infty$ as F$(w_0) \rightarrow 0$, we have that H$(w_0) \rightarrow 0$ as $w_0 \longrightarrow \underline{\xi}$. Hence the function defined by the right-hand side of the equation must intersect the function defined by the left-hand side of the equation at least once over this interval. To demonstrate that the functions intersect only once, it is sufficient to show that H'$(w_0) > -1$. We show this in Appendix A. Since H$(w_0) \geq 0$, ${w_0}^* \leq \xi_0$ so $w_0^* \leq \alpha_j + \xi_0$. Since $\xi_0 \leq \overline{\xi}$, the firm will never make in offer in excess of $\alpha_j + \overline{\xi}$. Since we have assumed that n≥ 1 when $w_0 = \underline{\xi}$, the firm will never make an offer less than $\alpha_j + \underline{\xi}$, since it is sure to lose all of its workers in that case. ■

2.7.1 Comparative Statics

The two exogenous variables in this set-up (given α_j) are the entry cost k, and the initial match ξ_0. Concentrating first on the entry cost k, we show that increases in k reduce the equilibrium wage offer of the incumbent firm. This result is entirely intuitive: increased inspection costs deter potential bidders, giving the incumbent firm more market power. As the entry cost disappears, then the wage tends towards its 'competitive' level.

Proposition 3 $\frac{dw_0^*}{dk} < 0$. *As* $k \longrightarrow 0$, $w_0 \longrightarrow v_0$

Proof. The first part of this proposition is proved by taking differentials with respect to k and $\tilde{w}_0$. Rewriting equation (5) as:

$$\tilde{w}_0 - v_0 + H(\tilde{w}_0) = 0$$

we have that:

$$\frac{dw_0}{dk} = \frac{d\tilde{w}_0}{dk} = \frac{-H_k(\tilde{w}_0)}{1 + H_{w_0}(\tilde{w}_0)}$$

From the Appendix, we know that the denominator is positive, and we have that:

$$H_k(\tilde{w}_0) = \frac{-\frac{dn}{dk}\frac{F'(\tilde{w}_0)}{F(\tilde{w}_0)}}{[\frac{dn}{d\tilde{w}_0}\ln F(\tilde{w}_0) + \frac{nf(\tilde{w}_0)}{F(\tilde{w}_0)}]^2}$$

since it is obvious from equation (1) that $\frac{d}{dk}(\frac{dn}{d\tilde{w}_0}) = 0$. $H_k(\tilde{w}_0)$ is now positive by virtue of the fact that the $\frac{dn}{dk} < 0$ (equation (2)). The second part of the proposition comes from the fact that from equation (1), as $k \longrightarrow 0$, $n \longrightarrow \infty$. As $n \rightarrow \infty$, the denominator of $H(w_0) \rightarrow \infty$ since $\frac{nF'(\tilde{w}_0)}{F(\tilde{w}_0)} \rightarrow \infty$, hence $H(\tilde{w}_0) \longrightarrow 0$ and so $w_0 \longrightarrow v_0$. ■

Turning next to the worker-firm match, we show that increased levels of match productivity in the training firm correspond to higher initial wage offers. Again this result is intuitively appealing.

Proposition 4 $\frac{dw_0^*}{d\xi_0} > 0$

Proof. This comes from taking differentials of equation (1) with respect to ξ_0 and $\tilde{w}_0$ and noting that $1 + H_{w_0}(\tilde{w}_0) > 0$. ■

Finally, we examine the effects of a change in the 'general skills' of worker j. We show that an increase in general skills induces a one-for-one shift in the equilibrium initial wage offer. The intuition for this result is that since this aspect of a worker's productivity is common to all firms, changes are perfectly reflected in the wage offers of potential competitiors, and therefore the wage offer of the incumbent firm.

Proposition 5 $\frac{dw_0^*}{d\alpha_j} = 1$

Proof. Since the expression for profits includes only the term $\tilde{w}_0$, where $\tilde{w}_0 = w_0 - \alpha_j$, any change in α_j must produce compensating changes in w_0. ∎

2.8 Training

We now wish to examine whether the incumbent firm will subsidise the cost of general training. Ultimately, this will depend on whether training increases the equilibrium profits of the incumbent firm. Hence we derive the necessary and sufficient condition for firms to subsidise the cost of general training as the third in a set of three results:

Proposition 6 *(i) A change in α_j has no effect on profit. .*

(ii) An increase in the costs of inspecting matches increases the equilibrium profits made by the incumbent firm.

(iii) If $\frac{dk}{d\tau}$=0, firms will not subsidise training. If $\frac{dk}{d\tau} > 0$, firms will subsidise training.

Proof. (i) We have shown that a change in α_j produces an exactly compensating change in w_0^*. Since the equilibrium probability of staying depends only on $\tilde{w}_0$, we therefore have that profits depend only on $\tilde{w}_0$ hence profits are not a function of α_j.

(ii) The second result follows from differentiating the profit function with respect to k, making use of the envelope theorem:

$$\frac{d\Pi(\tilde{w}_0^*)}{dk} = (\xi_0 - \tilde{w}_0)\frac{dP[stay(\tilde{w}_0^*)]}{dn}\frac{dn}{dk}$$

Since we have shown that $\frac{dn}{dk} < 0$ and since $\frac{dP[stay(\tilde{w}_0^*)]}{dn} < 0$, we have that $\frac{d\Pi(\tilde{w}_0^*)}{dk} > 0$. (iii) To prove the third result, consider the entry wage/training combination offered by firms. Competition among firms for new workers (school-leavers) ensures that the entry wage/training combination maximises the worker's utility subject to a break-even constraint. Normalising the value of the worker's output in the first period to zero, the break-even constraint can be written as:

$$w_e = \Pi(w_0^*) - C(\tau) \tag{6}$$

In other words, any profits/losses made by firms after the training period are passed back to the workers in the form of higher wages. Hence the issue of whether training is subsidised is one of whether firms increase profits by providing training. From results (i) and (ii), this only occurs when $\frac{dk}{d\tau} > 0$. ■

2.9 A Screening Technology Example

Since the question of whether firms pay for general training hinges on whether or not screening costs are increasing in the general skill level of workers, we propose two simple conditions under which this will hold. Suppose that a firm wishes to inspect the match between itself and a given applicant *j*, with general skill level α_j. To do this, we assume that it needs only to read the worker's CV. We assume that reading the CV takes time t, where t depends on both the skill level of the applicant j (α_j) and the skill level of reader k (α_k). Assuming constant returns to reading CVs, the cost of hiring a worker is then:

$$h(\alpha_j, \alpha_k) = \alpha_k t(\alpha_j, \alpha_k)$$

where α_k is the opportunity cost of reader k's time, and $t(.)$ is the length of time taken to read the CV. For a given general skill level of applicant, the firm chooses the skill level of the reader that will minimise this cost, α_k^*. The minimum hiring cost is then:

$$h^*(\alpha_j, \alpha_k^*) = \alpha_R^* t(\alpha_j, \alpha_k^*)$$

Then there are two simple cases in which hiring costs can be increasing in α_j :

1. It may be that the time taken to read a CV, conditional on the reader's general skill level, is increasing in the general skill level of the applicant. This would occur if the CVs of skilled applicants took longer to read (because they contained more information).

2. Suppose instead that the time taken to read a CV, conditional on the general skill level of the reader, is not a function of the general skill level of the applicant, but that only workers with skill levels higher than those of the applicant can read the CV ($\alpha_k \geq \alpha_j$). Then, since more highly skilled workers are required to read the CVs of more highly skilled applicants, opportunity costs and therefore hiring

costs are greater.

Either of these mechanisms would provide a rationale for screening costs increasing in the general skill level of workers, and we have already shown in our Introduction that the empirical evidence is strongly supportive of this assumption. Having shown how this assumption can account for the fact that firms invest in general training, the rest of the paper is devoted to an assessment of the empirical implications of this model.

3 Testable Implications and Empirical Strategy

In this section, we derive the testable implications of the model, and compare them with those derived from the model of asymmetric information proposed by Acemoglu and Pischke (1998) and the generic firm-specific training model. The principal implication will involve a comparison of the *wage changes* of those leaving the training firm ('movers') with the *wage changes* of those staying with the training firm ('stayers').

3.1 Wages

Focussing first on the *wage levels* of trained workers, it is obvious that the expected *wage levels* of stayers are lower than the expected *wage levels* of movers conditional on the match to the incumbent firm. To see this, consider two workers with match product v_0^+ within the incumbent firm. Both workers will receive wage offer $w_0^*(v_0^+)$ from the incumbent firm. Hence if one worker leaves the incumbent firm, it must be because she has received a wage offer in excess of $w_0^*(v_0^+)$. Hence conditional on v_0, the wages of 'movers' exceed those of 'stayers'. However, it can be shown that the probability of staying with the incumbent firm is an increasing function of match quality. This comes from the fact that the probability of staying with the the training firm is increasing in w_0^* (equation (4)) and the fact that the equilibrium wage is increasing in match ξ_0 (Proposition 4). Hence in general, we can not determine whether or not the expected *wage levels* of stayers will exceed the expected *wage levels* of movers.

We can however make predictions about the *wage changes* experienced by a worker leaving the incumbent firm. Supposing that we allow for some time to elapse between the worker receiving the wage offer from the incumbent firm and all of the wage offers from the outside labour market, then our model predicts that the wages of movers will exceed those of stayers, since workers will only ever leave the incumbent firm to earn higher wages. We summarise both of our results regarding the wages of trained workers as follows:

$$W^{MOVE} \lesseqgtr W^{STAY} \quad \text{(MS-L1)}$$

$$\Delta W^{MOVE} > \Delta W^{STAY} \quad \text{(MS-C1)}$$

We now compare these predictions with those derived from the other models.

3.2 Alternative: Asymmetric Information

The crucial assumption in the asymmetric information model of Acemoglu and Pischke (1998) is that workers' ability is revealed only to the incumbent firm. To keep things simple, suppose that there are only two types of ability - good workers and 'lemons'. Suppose also that a fraction of both types of worker leaves the firm at the end of the first period for exogenous reasons. Then the firm lays off the lemons and makes a wage offer to the good workers. This wage offer is determined by the outside wage offer, which in turn depends on the expected proportion of good and bad workers on the outside labour market. The firm's optimal wage offer is then slightly above this outside wage, ensuring that the firm keeps all of the remaining good workers. Since this wage will be below the productivity of good workers, the firm can earn rents on trained workers. That these rents are increasing in the level of training occurs because the productivity of workers is assumed to be a multiplicative function of ability and training.

In principle, there are at least three testable implications of the asymmetric information hypothesis. First, the model implies that in their first post-training job, the *wage levels* of the trainees that stay with the training firm should exceed the *wage levels* of those that move firms:

$$W^{MOVE} < W^{STAY} \tag{AP-L1}$$

Secondly, since workers that leave the firm for observable exogenous reasons (such as military service) should be paid the expected product of the entire population of trainees, the wages of these workers should exceed those of other movers:

$$W^{MOVE(EXOG)} > W^{MOVE(OTHER)} \tag{AP-L2}$$

Finally, the asymmetric information assumption implies that:

$$W^{MOVE(QUIT)} = W^{MOVE(LAYOFF)} \tag{AP-L3}$$

One problem with the asymmetric information story concerns its extension from two to multiple periods. In that case, workers can signal their ability by accepting an offer from the incumbent firm. Hence to extend the asymmetric information argument to multiple periods, it must be the case that ability is not perfectly observed by the incumbent firm at the end of the training period. Hence a stronger test of the asymmetric information hypothesis would involve a comparison of the *wage changes* of those trainees that stay with the training firm for a short time after training and then move, with the *wage changes* of those that remain with the training firm:

$$\Delta W^{MOVE} < \Delta W^{STAY} \tag{AP-C1}$$

$$\Delta W^{MOVE(EXOG)} > \Delta W^{MOVE} \tag{AP-C2}$$

$$\Delta W^{MOVE(QUIT)} = \Delta W^{MOVE(LAYOFF)} \tag{AP-C3}$$

3.3 Alternative: Firm-Specific Training

In the standard model of firm-specific training, firms and workers share the rents associated with the specific component of the training[8]. Hence we would expect the *wage levels* of stayers to exceed the *wage levels* of movers. Similarly, when workers leave the training firm, they lose their share of the specific human capital rents. Hence we

[8] See Acemoglu and Pischke (1999) for a model in which general and specific skills are complementary.

would also expect the *wage changes* of stayers to exceed the *wage changes* of movers:

$$W^{MOVE} < W^{STAY} \quad \text{(FS-L1)}$$

$$\Delta W^{MOVE} < \Delta W^{STAY} \quad \text{(FS-C1)}$$

3.4 Empirical Strategy

We test between our model and these two alternatives by comparing both the wage levels and wage changes of 'stayers' and 'movers'. However, we place more emphasis on the wage change results for three reasons. First, because our model is ambiguous as regards the expected wage levels of movers versus stayers. For a given level of match quality, movers earn higher wages, but the 'stayers' tend to be the workers with the best matches. The second problem with comparing the wage levels of movers and stayers is that we have argued that only a comparison of wage changes can serve as a test of the asymmetric information explanation once it has been extended from two to multiple periods.

Most importantly however, the third reason for focusing on wage changes is because we believe that tests based on wage levels will be biased because of unobserved heterogeneity. Implicitly, all of the empirical implications that we have so far derived refer to wages conditional on all those characteristics observed by the training firm and the outside labour market (in our model, α_j). Yet it is clear that the econometrician does not observe all these characteristics (in the German case, the obvious example is the test scores reported on the apprenticeship certificate). If these unobserved ability components are correlated with the decision to stay or leave the training firm (for example, if better workers are more likely to stay), then these tests may falsely accept the asymmetric information and/or the firm-specific capital hypotheses. Since first differencing wages will eliminate any permanent component of earnings correlated with the decision to leave the training firm, estimates based on first-differenced wages are robust to this problem.

In fact, rather than directly estimate the wage changes of movers and stayers, we estimate the wage differential between movers and stayers in two levels equations corresponding to observations before and after the group of movers leave the training firm. Then, comparing the estimated move-stay differentials in both periods (a regression-adjusted difference-in-difference procedure) is akin to estimating a model in first differences. The advantage of this approach is that it provides us with a neat way of comparing the wages of different groups of movers and stayers over a longer period - in our case, across five post-apprenticeship periods. Before turning to our data and empirical results however, we provide some contextual detail regarding the GAT system.

4 German Apprenticeship Training

In this section, we summarise the GAT system by way of seven stylised facts. Since the generality of apprenticeship skills and the fact that firms actually pay for this training are central to our story, we begin with a discussion of the evidence on these points. We then consider the recruitment and retention of trainees before discussing the evidence regarding screening costs.

4.1 The Nature of Apprenticeship Skills

The foundation of GAT remains the 1969 Vocational Training Act. This Act defined a number of occupations in which apprentices could be trained. These currently number some 375, with GAT lasting between two and three years, depending on the occupation trained in. Importantly, it also specified the curricula that would be followed within each of these apprenticeship occupations. The Act also required that training firms release their apprentices for one day a week to attend a local vocational college organised around one of five vocational fields (industry, commerce, home management, agriculture and other occupations). These colleges were designed to fill any gaps in general education and to prepare apprentices for the final examination. These examinations are heavily regulated, and typically consist of several written examinations in the subjects laid down by the training regulations, with many including an oral or practical component. Together, these regulations ensure that training is in principle 'general', at least within an occupation. Against this, it could be argued that these correspond only to minimum requirements, although as Acemoglu and Pischke (1998) report, firms typically do not have the time to train beyond these minimum requirements.

SF1 Although a worker apprenticed in a given occupation will not be able to transfer all of her training to every firm in the economy, she will be able to transfer all of her training to every firm operating jobs in the same occupation. In this sense, the training is 'general'

4.2 Estimates of Net Cost

Since the 1969 Vocational Training Act came into force, there have been three attempts to assess the extent to which German firms pay for apprenticeship training[9]. Each subtracts the production value of apprentices (P) from total training costs, where costs comprise direct training costs (apprenticeship wages (W^{APP}) plus materials cost (C^{M})) and the wage costs of training personnel ($C^{TRAINER}$):

$$NetCost = (W^{APP} + M) + C^{TRAINER} - P$$

The value of an apprentice's production (P) is calculated as the product of three factors: the time spent by apprentices in production as reported by supervisors (s), their estimated productivity relative to skilled workers (γ) and the productivity of skilled workers (as measured by their wages, $W^{SKILLED}$). The costs of training personnel are the time spent with apprentices (t) multiplied by the wage costs of trainers (W^{TRAIN})[10]. Hence:

$$NetCost = (W^{APP} + C^{M}) + tW^{TRAIN} - s\gamma W^{SKILLED}$$

The most contentious element of this calculation is the wage cost of trainers, W^{TRAIN}. In large firms, where full-time training personnel are employed, these costs are easily accounted for. In small firms by contrast, apprentices may simply watch the supervisor at work, or train at times when work is slack. In that case, it is not clear how large are

[9] These are described in more detail by Harhoff and Kane (1997).

[10] Since apprentices spend some of their time in the local vocational school, $s + t < 1$

training personnel costs.

The first two studies - undertaken for 1971/72 and 1980 - estimated training personnel costs in small craft firms by asking supervisors how much time they spent with trainees (t) and then multiply the supervisor's wage by this fraction. In the case of the second study, this produced a figure for net costs per apprentice per year of $9381 in the large firm (industry) sector and $5991 in the smaller firm (craft) sector (in 1990$)[11]. The third study - undertaken for 1991 - attempted to shed some light on the possible over-estimation of training firm costs in small firms, by excluding the wage costs of any trainers not engaged in training on a full-time basis. Using this method, estimated costs are $5485 for large industrial firms and $240 for smaller craft firms.

SF2 The net costs of training apprentices for large industrial firms are substantial. For smaller craft firms, they are much smaller, and perhaps zero.

4.3 Recruitment and Retention

We turn next to the provision of training places and the retention of new trainees. The first point to note is that whilst the majority of firms in Germany train apprentices, not all do so. Using data from an establishment panel of more than 4000 firms, Bender and Schwerdt (2000) find that in 1993, whilst over 85% of firms with more than 500 employees trained apprentices, the figure among firms with less than 50 employees was less than 30%.

SF3 The vast majority of large industrial firms train apprentices. Less than one-third

[11] Reported by Harhoff and Kane (1997) and converted at a rate of 1.62DM.

of small craft firms train apprentices.

In terms of the quality of apprentices recruited, there seems to be a clear correlation between school-leaving ability and firm size. Using self-reported Maths and German test score data, Harhoff and Kane (1993) show that those taking apprenticeships in the largest firms scored higher than those taking apprenticeships in the smallest firms.

SF4 The most able school-leavers apprentice in large industrial firms. The least able apprentice in small craft firms.

Finally, Harhoff and Kane (1997) provide evidence on the propensity of firms to train according to local labour market conditions. In probit and tobit estimates of the incidence and extent of training, they find a significant negative impact of the number of other firms operating in the same industry in the local county.

SF5 Firms are more likely to train when there are fewer local competitors operating in the same industry.

Turning to the retention of apprentices, another important fact concerns the high degree of post-apprenticeship mobility. Whilst Harhoff and Kane (1997) find that amongst a sample of apprentices graduating in 1980, the average retention rate of newly completed apprentices was 75%, this falls to 60% for those within one year of completion and 40% for those within five years of completion. Among smaller craft firms, retention rates are lower. Only 65% remain with the training firm upon completing apprenticeship, and this falls to 50% for those within one year of apprenticeship and 30% for those within five years of apprenticeship. Amongst larger industry firms, the figures are 80%, 70% and 50%. We provide our own evidence on this point in section 7.

SF6 Retention rates of new apprentices are higher in larger industrial firms than smaller craft firms.

4.4 Screening Costs

Wagner (1998) discusses the evidence relating to the costs of screening newly apprenticed workers in Germany. She argues that the costs of screening workers may be up to \$1,200-\$1,800 per hire[12]:

> "Companies that hire skilled workers from the market have to pay costs for advertisements, screening applicants, doing interviews at different levels and hiring costs. These have been estimated by a large German company to reach 2000DM-3000DM per hire" (p.6).

Moreover, she argues indirectly that these costs are likely to be far lower for unskilled workers. Referring to the costs of screening untrained school-leavers she claims that it is less costly to test their skills, since "(T)hey will learn these on the job, and one person is sufficient to conduct the interviews" (p.6). Since the same argument would presumably apply to older unskilled workers, we draw the following conclusion, in line with the evidence summarised in the Introduction:

SF7 Screening costs are higher for skilled workers (apprentices) than unskilled workers (non-apprentices).

[12] In \$1990 and converted at a rate of 1.62DM.

5 Data Issues

The paper uses two sets of data to implement the tests discussed in section 3: the IAB data, a 1% sample of German social security records, and the German Socio-Economic Panel (GSOEP), a much smaller survey of German households[13]. These data are complementary in the sense that the former contains detailed wage histories for a huge number of individuals, but very little other information. Whilst the latter is much smaller in size, there are some interesting questions in the survey that we will exploit in our analysis.

5.1 IAB Data

The IAB 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 firm trained in, and we have some limited information on this firm (including its size and the industry in which it operates). 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. The data are top coded, although the top coding affects only a tiny proportion of the young apprentices in our sample.

5.1.1 The Basic Sample

Only German male apprentices are retained for analysis. 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

[13] See the Data Appendix for a more detailed discussion of the two data sets.

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.[14]

5.1.2 Post-Apprenticeship Observations

The data is organised as an event history, and since changes of employer and spells of unemployment are notifiable events, we can record them accurately. However, there is a problem with these data in that some events are non-notifiable, in particular, changes of wages and apprenticeship completion. The fact that wage changes are not notified means that wages in a spell actually refer to mean daily wages. For apprentices completing training and then moving to another firm or into non-employment, the change of firm is notified, and so we can pin down the end of apprenticeship precisely. However, for apprentices staying with the training firm after apprenticeship however, no notification is made. To understand the problem that this creates, and our solution to it, consider the following fragment of the event history of fictional person number 500, which is typical of these data:

- Spell x - person number 500; firm number 1000; spell start Jan 1 1983; spell end

[14]This group makes up about 10-15% of all apprentices (author's calculations with the German Socio-Economic Panel (GSOEP)). See also Euwals and Winkelmann (2001).

Jan 1 1984; employment status: trainee

- Spell x+1 - person number 500; firm number 1000; spell start Jan 1 1984; spell end Jan 1 1985; employment status: skilled worker
- Spell x+2 - person number: 500; firm number: 1000; spell start Jan 1 1985; spell end Jan 1 1986; employment status: skilled worker

This tells us that on January 1 1984, this person was still working as an apprentice. Since we know that by 1 January 1985 this person was no longer an apprentice, we can infer that apprenticeship finished some time between 1 January 1984 and 1 January 1985, although we do not know exactly when. The problem is that the wage recorded for this spell will be an average of the apprenticeship wage and the post-apprenticeship wage and will therefore be of little use to us. Our solution is to define the first post-apprenticeship observation as that in progress on the 1 July in the calendar year *after* the last notification of apprenticeship was given. Hence the first post-apprenticeship spell for this worker would be spell number x+2. To take another example, someone who was notified as an apprentice on 1 January 1984 but then moved firms and became a skilled worker on 15 May 1984 would have as their first apprenticeship observation the spell in progress on 1 July 1985.

This implies that the first post-apprenticeship observation could in principle have started anywhere between 1 day (last apprenticeship notification recorded 31 December 1985, first post-apprenticeship observation spell started 1 January 1986) and 18 months (last apprenticeship notification 1 January 1985, first post-apprenticeship observation spell started 31 June 1986) after the end of apprenticeship. Notice also that this is not necessarily the first post-apprenticeship spell, since the apprentices could have any number of spells of work, unemployment or time out of the labour force in between finishing apprenticeship and the first post-apprenticeship observation.

In order to obtain a more detailed picture of post-apprenticeship wage patterns, we concentrate on a five-observation post-apprenticeship window. To do this, we classify spells in progress on 1 July in each year as either the first post-apprenticeship observation (defined as above), or the second, third, fourth or fifth post-apprenticeship observation.[15] Second (third, fourth, fifth) post-apprenticeship observations are defined as those in progress in the second (third, fourth, fifth) calendar year after the last apprenticeship notification.

5.1.3 Descriptive Statistics

The key information for our purposes is first, whether the first post-apprenticeship observation involves full-time work, and secondly, whether the first post-apprenticeship observation is a spell with the training firm. On the first question, we only analyse those apprentices in full-time work at the post-apprenticeship observations of interest. Although this may be a selective group, we obviously have no wage information on the non-employed. Regarding the second issue, if the apprentice is still with the apprenticeship firm, we define this person a ‘STAYER’. Otherwise, we define her as a ‘MOVER’.[16]

The top panel of Table 1 uses this information to calculate the propotion of appren-

[15]We discard spells corresponding to spells of employment more than five calendar years after the end of apprenticeship.

[16]In fact, firm numbers can change due to firm reorganisations and mergers, so there may be a few misclassifications. These will however be a tiny fraction of the total sum of firm changes.

tices still with the training firm at each of the five post-apprenticeship spells. From the Table, we see that for the first post-apprenticeship observation, approximately 56% of all apprentices in full-time employment are still working in the apprenticeship firm. The typical figure for the proportion remaining with the apprenticeship firm upon completing apprenticeship is 70%, but since we are looking at apprentices in the calendar year after the end of apprenticeship we would expect to find a lower figure. Our estimates of the proportion of movers between one and five observations after apprenticeship (roughly, between one and five years after apprenticeship) are similar to other estimates. For example, Euwals and Winkelmann (2001) find that the proportion of workers staying for at least 5 years is 26%.

Suppose next that we look at the whole sample of apprentices observed in full-time employment at all five post-apprenticeship observations. This sample will be smaller than the sum of the sample sizes for each post-apprenticeship observation since some of those observed in full-time employment at the first post-apprenticeship spell may not be employed at the second post-apprenticeship spell (and vice-versa). We can then classify this sample into six mutually exclusive and exhaustive groups (the sample size and fraction of the total sample of workers observed in full-time employment across all five post-apprenticeship observations are given in parentheses).

1. STAYER (2725, 35.03%): those observed working for the training firm across all five post-apprenticeship observations

2. MOVE01 (3564, 45.81%): those who left the training firm some time between finishing the apprenticeship and the first post-apprenticeship observation

3. MOVE12 (566, 7.28%): those who left the training firm some time between the first and second post-apprenticeship observations

4. MOVE23 (377, 4.85%): those who left the training firm some time between the second and third post-apprenticeship observations

5. MOVE34 (313, 4.02%): those who left the training firm some time between the third and fourth post-apprenticeship observations

6. MOVE45 (235, 3.02%): those who left the training firm some time between the fourth and fifth post-apprenticeship observations

In order to assess how these groups differ, Table 2a presents some descriptive statistics for four of the six subsamples: STAYER, MOVE01, MOVE12 and MOVE45. From the first rows, we see dramatic differences in training firm size between the different samples. Whilst a significant proportion (27.21%) of the STAYER group were apprenticed in the largest firms, the figure for those moving between the end of apprenticeship and the first post-apprenticeship spell (MOVE01) is only 9.35%. On the other hand, these apprentices are far more likely to have trained in firms with less than ten employees. In this respect, those moving between the first and second post-apprenticeship spell (MOVE12) look more similar to the MOVE01 group, whilst those leaving much later (MOVE45) lie somewhere in between this group and the STAYER group.

In the next rows, we describe some facts regarding the period between the final apprenticeship notification and the first post-apprenticeship observation (recall that for the group that stay with the training firm, the last apprenticeship notification may not correspond exactly to the end of apprenticeship). First, we see that those workers that move between the end of appprenticeship and the first post-apprenticeship spell (MOVE01) are slightly older than those other three groups of workers still with the training firm at this point. This confirms the findings in Euwals and Winkelmann (2001). Next, we look at the time gap between the final apprenticeship notification and the start of the spell classified as the first post-apprenticeship observation. For the three groups of workers still with the training firm at this point (STAYER, MOVE12 and MOVE45), the mean lag is close to 365 days. This simply says that in the vast majoity of cases, the data look like the example presented above. The exceptions are those with spells of unemployment or time out of the labour market between finishing apprenticeship and starting work.[17] For those that have moved, the average gap is smaller, since for many of these, the reported end of apprenticeship may be in the middle of the calender year prior to the first post-apprenticeship observation.

In the next rows, we present the cumulative number of days of unemployment or non-employment experienced by apprentices. As we would expect, for the three groups still with the training firm (STAYER, MOVE12 and MOVE45), the average number is less than one week of unemployment and one month out of the labour force in all cases. For the group of movers however, the average time spent unemployed is more than two weeks, with an average of one month out of the labour force. Also of interest is the fact that almost 40% of the movers have already moved again by the first post-apprenticeship observation. By the fifth post-apprenticeship observation, the group MOVE01 have an average of roughly five weeks of unemployment, five weeks out of

[17] Note that our classification into five post-apprenticeship observations and the sample restriction that we have all five obseravations for every person eliminates those with long breaks from the labour force, such as those completing military service after apprenticeship (which typically lasts for 15 months).

the labour force and have worked for an average of 2 new employers.

Turning to wages, we present mean daily wages at each of the five spells for each of the four groups studied. At the first observation, it is interesting to note that the STAYER group have higher mean wages than the group that have already moved (MOVE01). However, the MOVE01 group have higher mean wages than the other groups of workers who are still with the training firm at this point but will leave the training firm within the observation window. Although these are only raw wage levels, these differences suggest that a preoccupation with the mover-stayer differential at the first post-apprenticeship observation (i.e. aggregating these different groups of stayers into one group and comparing their wages with the group of workers that have already moved (MOVE01)) may mask interesting and important features of the data.

One such feature can be seen by comparing the wages of the MOVER12 group before and after they leave the training firm (the first and second post-apprenticeship observation respectively). Whilst this group have the lowest average wages at the first post-apprenticeship observation, by the second post-apprenticeship observtaion, their average wages are similar to the MOVE01 group. Again, these are only raw wage differentials. However, this pattern is line with the empirical implications of our model as they relate to wage changes (MS-C1) and goes against the predictions of the asymmetric information and firm-specific human capital models ((AP-C1) and (FS-C1)). Moreover, this phenomenon is repeated for those moving between the second and third, third and fourth, and fourth and the fifth post-apprenticeship observation. The next section investigates whether the same pattern holds for regression-adjusted wages. First though, we present some descriptive statistics based on the GSOEP data.

5.2 GSOEP Data

5.2.1 Basic Sample and Post-Apprenticeship Observations

The GSOEP data set is much smaller than the IAB data, and is based on an annual survey, rather than being organised along 'event history' lines. We therefore proceed by generating a sample of apprentices observed in apprenticeship at one interview and reporting having finished apprenticeship training at the next interview. We then define this interview as the first post-apprenticeship observation. We construct similar samples for each pair of years from 1984-1985 to 1995-1996 (inclusive). In order for sample sizes not to become too small, our samples include all apprentices (men and women, German and non-German) without an *Abitur* certificate. Pooling our eleven pairs of years together, this create an overall sample of about 1000 individuals. By tracking these apprentices for another year (the second post-apprenticeship interview), we create another, albeit smaller sample of around 800 individuals.

The major advantage of the GSOEP lies in the fact that we can classify the group of workers that have left the training firm according to the reasons why they left. This will help us to test some of the predictions of the asymmetric information model, in particular (AP-L3) and (AP-C3). We classify movers into two groups - those that quit the training firm and those that left for other reasons. Since we would expect an over-reporting of quits and an under-reporting of moves for other reasons, we include as quits only those workers who gave 'quit' as their *only* response to the question of why they changed firms. Any other responses - or a combination of 'quit' and other responses - we classify as 'moves for other reasons'. These other reasons change across the survey waves and are listed in Table A1.

5.2.2 Descriptive Statistics

In the first row of Panel B of Table 1, we present mobility statistics based on the GSOEP data. We see that of the 922 interviewees in full-time employment at the first post-apprenticeship interview, approximately one-third have already left the training firm. Since our sample are observed a maximum of twelve months after apprenticeship (since GSOEP interviews are typically twelve months apart), it is unsurprising that the proportion of stayers is slightly higher than that found using the IAB.

The Table also shows that of the 341 movers, 63 could be classified as have 'quit' whilst 278 moved firms for other reasons. The relatively small fraction of 'quitters' is certainly due in part to our classification of 'quit' and 'other'. However, it may also be due to propensity of trainees to work for the training firm for a short time after the end of apprenticeship, in order to signal their ability (see our discussion of section 3.2). This is consistent with the fact that by the second post-apprenticeship observation, the proportion still with the training firm has fallen to 49%, and, splitting into those that quit and those that left for other reasons, the fraction of quitters has now increased to more than one-third.

Table 2b investigates further the differences between quitters and other types of mover further. Looking first at training firm size, it seems that as found using the IAB data, stayers are far more likely to have trained in larger firms, although as with all of these results, there is a caveat regarding the small sample sizes. It is also interesting to note that a smaller proportion of those quitting between the end of apprenticeship and the first post-apprenticeship observation were trained in large firms than those moving for other reasons. Among those moving between the first and second post-apprenticeship interview, differences are not so apparent.

Turning to the age at the first post-apprenticeship observation, we find that those workers staying across both post-apprenticeship observations are on average the oldest at the first post-apprenticeship observation. The differences between these results and those found for the IAB data will in part be accounted for by the fact that we now inlcude women and non-Germans in our sample. It should also be remembered that whilst these samples refer to those in full-time employment at both post-apprenticeship observations, the results reported in Table 2a based on the IAB data refer to those in full-time employment at five post-apprenticeship observations.

Looking next at the daily wages of apprentices at the first post-apprenticeship oberavation, we see that wage levels for all groups of workers are lower than those reported in Table 2a using the IAB data. This will in part be due to the fact that the apprentices in the GSOEP are typically observed at an earlier stage in their career than those in the IAB data. Again, the fact that the IAB sample is selected on the basis of a five- rather than a two-observation atachment to the labour force will also account for some of the difference.

The relative wages of movers and stayers look similar to those found using the IAB data in that at the first post-apprenticeship observation, the STAYER group enjoys the highest average wages, followed by the group MOVE01 and then MOVE12. Disaggregating according to whether the movers quit or left the training firm for other reasons, we see that the average wages of quitters moving between the end of training and the first post-apprenticeship interview are lower than those of apprentices moving for other reasons, and similarly for those moving between the first and second post-apprenticeship interviews. As was the case with the IAB data, the group leaving the training firm between the first and second post-apprenticeship observations enjoy the largest increase in mean wages. Again, this is consistent with (MS-C1) and inconsistent with (AP-C1) and (FS-C1). However, these refer only to raw wage differentials, and since we have shown that these different groups have different characteristics, we need to ask whether these differences persist after controlling for these characteristics. We address this question in the next section of the paper.

6 Empirical Results

This section describes our empirical results. The key test of our model involves a comparison of the *wage changes* of movers and stayers. Recall that our model predicts that the *wage changes* of movers should exceed those of stayers (MS-C1), whilst the asymmetric information and firm-specific human capital models predict precisely the opposite ((AP-C1) and (FS-C1))

Before we turn to our comparisons of the wage changes of movers and stayers, we briefly review the existing literature that compares the *wage levels* of movers and stayers. We argued in section 3 that whilst this comparison has been used to test the asymmetric information and firm-specific human capital theories, the comparison may be biased towards accepting implications (AP-L1), (AP-L2) and (FS-L1) if workers with low levels of unobserved ability are more likely to leave the training firm.

6.1 Movers versus Stayers: Wage Levels

Of the few papers that compare the wages of movers and stayers, all compare the *wage levels* of stayers with those of movers. We describe the results of these papers before presenting our own results based on the two datasets.

Previous Evidence

This set of papers can be classified into those that attempt to control for the selective nature of moves out of the training firm, and those that do not. Table 3a summarises the results of those studies that do not attempt to correct for selection when comparing wages. In column (1), we reproduce the results of Acemoglu and Pischke (1998). Their estimates are designed to test both (AP-L1) and (AP-L2), using those leaving the training firm to join the military as a proxy for an 'exogenous' move out of the training firm. As can be seen from the Table, they find that both stayers and military quitters earn slightly more than movers. Although they interpret these results as supportive of (AP-L1) and (AP-L2), the associated standard errors are large. Morever, there is a major problem with the data used. Whilst the survey questions relating to apprenticeship are retrospective, the wage measure used is the current wage. Hence there are in the sample some 55 year olds for whom earnings aged 55 are being used in conjunction with events that occurred when the respondents were 20.

Looking across the other columns of this Table, it is clear that the results are mixed. For example, as seem in column 2, Harhoff and Kane (1997) use the same data and find that the wages of movers *exceed* those of stayers. Of course, since the data is the same as that used by Acemoglu and Pischke (1998), it is subject to the same difficulties.

Also, only one wave of this data is used, rather than the three waves used by Acemoglu and Pischke (1998) and no training firm controls are included.[18]

In the third column, we present the results of Euwals and Winkelmann (2001). Reversing those of Harhoff and Kane (1997) these would appear to support the findings of Acemoglu and Pischke (1998), in that the stayers now earn more than the movers, particularly amongst those workers trained in large firms. Of course these estimates may suffer from selection biases, and so we turn briefly to two studies that address the selection issue.

Euwals (1998) addresses the problem by estimating a switching regression model. However, the variable used in his switching regression model (marriage) does not seem able to adequately capture the driving force behind mobility from the firm and his results are mixed. In the regression controlling for training firm size, he finds that movers earn more than stayers (the estimated coefficient (standard error) on the 'stayer' variable (movers are the base group) is -0.029 (0.012) for those trained in firms with over 50 employees and 0.006 (0.009) for those trained in all firms).

Werwatz (1996) also estimates a switching regression model, this time using age at the end of apprenticeship and a dummy variable for school type as variables intended to capture mobility. Werwatz (1996) finds that male movers earn around 9% more than male stayers, whilst female movers earn 11% less than female stayers. Again however, the selection equation is not well determined and results are mixed. For these reasons,

[18] Harhoff and Kane (1997) interpret their results as evidence in favour of a mobility cost explanation of firm investment. We have ignored this possibility, since we see no reason why mobility costs should be increasing in the level of training offered. That is, whilst mobility costs would enable firms to earn rents on all workers (since mobility costs would enable firms to pay workers less than their marginal product), there seems no reason to suppose that these costs - and therefore firm rents - are increasing in the level of training offered.

we prefer to control for selection out of the training firm using longtitudinal data. We present the results of this analysis in the next subsection. First though, we present our own estimates of the wage differences in levels.

Evidence based on IAB data

Table 4a presents the results of our wage comparisons at the first post-apprenticeship observation using the IAB data. In column (1), we present a specification including only training firm size dummies, age and age squared as expanatory variables, in addition to a dummy variable indicating whether or not the worker is still with the training firm. Notice that we follow the previous literature in not distinguishing stayers according to when they eventually left the training firm. For example, we aggregate together those with the training firm that stay across all five post-apprenticeship observations (STAYER) and those that leave between the first and second post-apprenticeship observation (MOVE12).

Focusing first on the other variables included in the equation, we see that age has a small but positive impact on earnings, although the age squared term is never significant. More dramatic are the training firm size coefficients. Although these are not included in the Table, the estimated coefficients increase sharply (single employer firms are the excluded category) between firms with between two and nine employees (coefficient 0.24; standard error 0.015) and those with more than 1000 employees (coefficient 0.479, standard error 0.015). This pattern is to be expected, since it has been shown elsewhere that the most able school-leavers typically train in the largest firms (see especially Harhoff and Kane (1993) who present self-reported test scores by apprenticeship firm size). We also include year dummies in the specification. As we would expect, coefficients are large for later years, picking up the general increase in real wages over the period.

Turning now to the mover-stayer differential, we find that stayers earn slightly more than movers. Our finding is similar to that of **?** who use the same dataset and a slightly different sample. In Table 2a, we showed that the MOVE01 group typically experienced more time in unemployment and more time out of the labour force between apprenticeship completion and the first post-apprenticeship observation, hence we include these variables in specification (2). As we would expect, they are negatively correlated with earnings, with one month in unemployment associated with roughly a 0.3% drop in earnings and one month out of the labour force associated with roughly a 0.5% drop in earnings. Including these terms results in a slight fall in the stayer earnings advantage.

We do not wish to include any current employer characteristics such as firm size in our specification, since we want to compare wage levels without conditioning on these variables. For example, if large firms pay higher wages (for whatever reason) and workers can earn higher wages by leaving small training firms to work for larger firms, we want to allow our estimates to capture this. The suspicion that controlling for employing firm characteristics would over-estimate the earnings advanatage of stayers is confirmed in specification (3), where we include employing firm size dummies. Specification (4) restricts the sample to the period after 1983, since it can be argued that the wage data are more reliable over this period[19]. In fact, it does not have a great impact on our results. Finally, in specification (5), we restrict the sample to those work-

[19]Prior to 1984, firms were not obliged to report extra payments such as Christmas and holiday bonuses, which are an important part of compensation in Germany.

ers trained in firms of over 100 workers. Consistent with some of the other literature (especially **?**) we now find a substantial wage advantage for stayers over movers, of the order of 8%.

To summarise, when comparing wages at the first post-apprenticeship observation, we find results similar to some of those estimated elsewhere. In particular, we find that stayers earn slightly more than movers across most of the specifications, and substantially more when concentrating only on large training firms. Whilst these results support (AP-L1), it should be remembered that some other papers find that movers earn more than stayers. More importantly, since these estimates do not adequately control for selection out of the training firm, the next subsection analyses the wage dynamics of movers and stayers over a longer period. First though, we perform a similar analysis using the GSOEP data.

Evidence based on GSOEP data

Table 4b presents the results of comparing the wages of movers and stayers at the first post-apprenticeship interview using the GSOEP data. In column (1), we estimate a specification similar to that estimated in Table 4a, except that we now include dummy variables for being 'male' and 'German'. Estimates for age and the training firm size dummies are similar in sign and magnitude to those estimated using the IAB data, and whilst the 'male' estimate is large and positive, the 'German' estimate is not sigificantly different from zero. In contrast to our estimates, we now find that movers earn slightly more than stayers. That this result is different to that estimated using the IAB data is in line with the mixed evidence presented in Table 3. Once again however, we find that stayers earn more than movers when we focus on large firms only (specifications (5)).

Including employing firm size dummies (specifications (3)) does not affect our results substantially.

The main motive for using the GSOEP data is to break down the group of movers into those that quit and those that moved for other reasons. This will allow us to test (AP-L3). In the basic specification (2) and the extended specification (4), we see that the differences between the two groups of movers are not large. However, focussing only on those apprentices trained in large firms (specification (6)), we see that apprentices quitting large firms enjoy positive wage differentials, whilst those leaving for other reasons have negative wage differentials. This is evidence against the presence of asymmetric information in the market for new apprentices (AP-L3), since this theory predicts that the oustide labour market should not be able to tell these two groups apart.

6.2 Movers versus Stayers: Wage Changes

We now turn to a comparison of the wages of those that stayed with the training firm and those that left some time after the apprenticeship was completed. We begin with the results based on our IAB sample, before turning to the GSOEP data.

IAB Data

Recall that we can use the IAB data to generate a sample of workers each with five post-apprenticeship observations. In section 5, we described how we split these into six mutually exclusive and exhaustive groups: STAYER, MOVE01, MOVE12, MOVE23, MOVE34 and MOVE45. In the previous section, we pooled the five groups of workers still with the training firm at the first post-apprenticeship observation (STAYER, MOVE12, MOVE23, MOVE34 and MOVE45). Now however, we split these groups up and compare wages at all five post-apprenticeship observations.

The results are presented in Panel A of Table 5a. Focus first on the first column of the Table, which refers to the first post-apprenticeship observation. Looking down the rows of the Table we see the earnings differentials of all five groups of mover, relative to the stayers (the base category). The specification is identical to column (2) of Table 4, and the estimated age effects are similar. Notice however that the sample size is much smaller, since we are now focussing only on those workers in full-time employment at all five post-apprenticeship observations.

The interesting feature of column (1) is that whilst those workers that have already left the firm earn slightly less than those workers that eventually stay with the firm across all five post-apprenticeship observations, they earn significantly more than the other group of stayers that will eventually leave the training firm. Turning then to column (2) of the Table, we see the results of the same equation estimated at the second post-apprenticeship observation, *after* the group MOVE12 have left the training firm. Now, we see that whilst the MOVE23, MOVE34 and MOVE45 groups still earn less than the STAYER and MOVE01 group, the wages of this MOVE12 group are no longer significantly different to the STAYER group. This suggests that these workers can increase their (relative) wages by leaving the training firm. This is strong evidence in favour of (MS-C1) and against (AP-C1) and (FS-C1).

Looking across the remaining three columns of the Table, we see this pattern repeated for the other groups of movers (MOVE23, MOVE34 and MOVE45). This can be seen even more clearly in Figure 1, which plots the wage differentials of all movers relative to the base group of stayers (whose wages are normalised to zero). In Panel B of the Table, we estimate the same equations for only those workers trained in firms of more than 100 employees (the wage differentials are plotted in Figure 2). Whilst the pattern is now less clear, increases in relative wages are still experienced by each group upon moving.

We interpret these patterns as further evidence in favour our claim that comparing the wages of movers and stayers in levels at the first-post-apprenticeship observation masks interesting patterns in the wage dynamics of movers and stayers. Moreover, the increase in wages experienced by all types of mover is strongly suggestive of a matching-screening explanation for human capital investment and strongly against asymmetric information or specific human capital explanations for firm investment. In order to examine the asymmetric information argument further, we now present a similar analysis based on the GSOEP.

GSOEP Data

We use the GSOEP to compare the wage dynamics of movers and stayers over a short time period (two post-apprenticeship observations) and the results are plotted in Table 5b. We find that at the first post-apprenticeship observation, the earnings of the group moving between the end of apprenticeship and the first post-apprenticeship observation are not significantly different to those of the stayers, whilst the earnings of the group moving between the first and second post-apprenticeship observations are much lower. At the second post-apprenticeship observation however, after this group has left the training firm, this conclusion no longer holds, and the earnings of this group are now insignificantly different to those of the stayers.

Looking at the movers according to whether they quit the training firm or left for other reasons, we see that this pattern holds for both groups, although the effect is more dramatic for those quitting between the first and second post-apprenticeship observations. This is evidence against asymmetric information implication (AP-C3), since the theory again predicts that the wage changes of these two groups should (*ceteris parabus*) be identical. Finally, in the bottom panel of the Table, we restrict attention to only the largest training firms. We see familiar patterns of wage dynamics for movers and stayers when the different types of movers are aggregated together, but diaggregating the movers, we find that the increase in relative wages between the two post-apprenticeship observations is accounted for entirely by the group of movers quitting the training firm. Again, this is strong evidence against the asymmetric information hypothesis.

7 Conclusions

The paper presented a model of the labour market in which a combination of match heterogeneity and screening costs gave incumbent firms some monopsony power and enabled them to earn rents on retained workers. Without matching, the screening problem would be trivial, since workers would essentially be the same. Without screening costs, the labour market would be competitive, in the sense that an infinite number of firms would bid for workers, with workers leaving to work at the firm where they were most productive. The crucial ingredient in our model is the assumption that training increases the costs of screening, and therefore increases the monopsony power of workers. Although it is a well known fact that screening costs are a sharply increasing function of a worker's skill level, we believe that this is the first model to account for and exploit this fact.

We also showed that wage patterns of trained workers are consistent with our model, at least for the case of German apprentices. Although we also interpreted our results as evidence against the asymmetric information model of Acemoglu and Pischke (1998) and the generic firm-specific human capital model, we would not claim that asymmetric information arguments are never relevant to the decision of firms to invest in general training. Instead, it would appear to us that different motives may drive the training decisions for different types of workers. Hence for the case of low-skilled workers with tenuous attachments to the labour force, asymmetric information may well be the driving force behind training decisions, and Autor (2001) presents convincing evidence to this effect. For more educated workers however, where ability is likely to be known, or easily signalled, we believe that a combination of matching and screening costs may drive training decisions. Empirical analyses of other training institutions would help to confirm whether or not this was in fact the case.

A Proofs of Propositions

In this Appendix, we derive the results ommitted from section 2. We begin by deriving the optimal bidding strategy of bidder i after he has paid the cost k of inspecting the match and has a valuation for worker j of v_{ij}.

A.1 Optimal Bidding

We proceed by guessing that the optimal bid satisfies certain properties, deriving the optimal bid under these assumptions, and then showing that that this is in fact an equilibrium bid[20]. In particular, we guess that the optimal bidding strategy is to bid:

$$B(v_{ij}) = \alpha_j + b(\xi_{ij})$$

if v_{ij} is greater than some marginal type v_0 and to not bid if $v_{ij} < v_0$, where $b(.)$ is a strictly increasing function.

To derive the optimal bid under these guesses, we use the fact that the expected profit of a bidder after he has drawn his valuation v_{ij} can be written as:

$$\Pi(v_{ij}) = Q(v_{ij})v_{ij} - P(v_{ij}) \tag{A1}$$

where $Q(.)$ is the equilibrium probability of winning function, and P is the expected payment function. From a standard version of the revenue equivalence theorem, we

[20]This subsection closely follows the discussion in Matthews (1995).

can show that this is equivalent to:

$$\Pi(v_{ij}) = \Pi_i(v_0) + \int_{v_0}^{v_{ij}} Q(y)dy \tag{A2}$$

where $\Pi_i(v_0)$ is the profit of the marginal type v_0 (see for example {matthews:1995}). Provided that the auction awards the object to the highest bidder (as most standard auction types do), then given our guess that $b(.)$ is strictly increasing, the equilibrium probability of winning is:

$$Q(v_{ij}) = G(v_{ij}) = G(\xi_{ij})$$

if $\xi_{ij} > v_0 - \alpha_j = \xi_0$, and zero otherwise. In other words, the equilibrium probability-of-winning function $G(.) = F(.)^{n-1}$ is the probability that bidder i has the best match among n potential bidders. Hence from (A2), the expected profit of a bidder after he has drawn ξ_{ij} is:

$$\Pi(v_{ij}) = \Pi_i(\xi_0) + \int_{\xi_0}^{\xi_{ij}} G(\xi)d\xi \tag{A3}$$

Under the assumptions on the bid function $B(.)$, the expected payment of a bidder with valuation v_{ij} in a first-price auction is:

$$P(v_{ij}) = [\alpha_j + b(\xi_{ij})]G(\xi_{ij})$$

Substituting this into equation (A1) and using (A3), we have that:

$$\Pi_i(\tilde{\xi}_0) + \int_{\tilde{\xi}_0}^{\xi_{ij}} G(\xi)d\xi = G(\xi_{ij})\xi_{ij} - b(\xi_{ij})G(\xi_{ij})$$

Since the marginal and non-participating types must have zero profit, we have that $\Pi_i(\tilde{\xi}_0) = 0$ and:

$$b(\xi_{ij}) = \xi_{ij} - \int_{\tilde{\xi}_0}^{\xi_{ij}} \frac{G(x)}{G(\xi)}dx$$

Hence:

$$B(\xi) = \alpha_j + \xi_{ij} - \int_{\tilde{\xi}_0}^{\xi_{ij}} \frac{G(x)}{G(\xi)}dx \qquad \text{(A4)}$$

Since the marginal type wins with probability $G(\tilde{\xi}_0)$ with any bid not less than the reservation bid price w_0, this type must bid exactly w_0. For type $\tilde{\xi}_0$ to make zero expected profit when bidding w_0, it must be the case that:

$$w_0 = \alpha_j + \tilde{\xi}_0$$

Hence for the marginal type, $\tilde{\xi}_0 = w_0 - \alpha_j = \tilde{v}_0$. Substituting this into the optimal bid function (A4), we have that:

$$B(v_{ij}) = \alpha_j + \xi_{ij} - \int_{\tilde{v}_0}^{\xi_{ij}} \frac{G(x)}{G(\xi)}dx \qquad \text{(A5)}$$

Hence the optimal bid function $B(.)$ that we have found satisfies our guesses. Proving that it is actually an equilibrium is standard, since the presence of α_j makes no difference to the 'pseudo-concavity' or second-order condition (see Matthews (1995)).

A.2 $1+H_{w_0}(w_0) > 0$

Using the definition of $H(w_0)$, we can write $H_{w_0}(w_0)$ as:

$$H_{w_0}(w_0) = \frac{-\frac{\partial}{\partial w_0}[\frac{dn}{dw_0}\ln F(w_0) + \frac{nF'(w_0)}{F(w_0)}]}{[\frac{dn}{dw_0}\ln F(w_0) + \frac{nF'(w_0)}{F(w_0)}]^2}$$

Hence:

$$1 + H_{w_0}(w_0) = \frac{[\frac{dn}{dw_0}\ln F(w_0) + \frac{nF'(w_0)}{F(w_0)}]^2 - \frac{\partial}{\partial w_0}[\frac{dn}{dw_0}\ln F(w_0) + \frac{nF'(w_0)}{F(w_0)}]}{[\frac{dn}{dw_0}\ln F(w_0) + \frac{nF'(w_0)}{F(w_0)}]^2}$$

For this term to be positive, we require that:

$$[\frac{dn}{dw_0}\ln F(w_0) + \frac{nF'(w_0)}{F(w_0)}]^2 > \frac{\partial}{\partial w_0}(\frac{dn}{dw_0}\ln F(w_0) + \frac{nF'(w_0)}{F(w_0)}) \qquad \text{(A6)}$$

We examine both sides of equation (A6) in turn.

A.2.1 Left-Hand-Side of (A6)

We can evaluate the left-hand-side of equation (A5) as follows:

$$[\frac{nF'(w_0)}{F(w_0)}]^2 + 2n\frac{F'(w_0)}{F(w_0)}[\frac{dn}{dw_0}\ln F(w_0)] + [\frac{dn}{dw_0}\ln F(w_0)]^2$$

where all three terms are positive.

A.2.2 Right-Hand-Side of (A6)

The right-hand-side of (A6) can be written:

$$\frac{d^2 n}{d\hat{w}_0} \ln F(\hat{w}_0) + \frac{dn}{d\hat{w}_0} \frac{F'(\hat{w}_0)}{F(\hat{w}_0)} + n[\ln F(\hat{w}_0)'']$$

where the first term is:

$$\frac{d^2 n}{d\hat{w}_0} \ln F(\hat{w}_0) = [\frac{N(\hat{w}_0) M_{\hat{w}_0}(\hat{w}_0) - M(\hat{w}_0) \frac{d}{d\hat{w}_0} N(\hat{w}_0)}{N(\hat{w}_0)^2}] \ln F(\hat{w}_0)$$

and:

$$\begin{aligned}
N(\hat{w}_0) &= \int_{\hat{w}_0}^{\overline{\xi}} L(\hat{w}_0) d\xi \\
L(\hat{w}_0) &= M(\hat{w}_0) \ln F(\hat{w}_0) \\
M(\hat{w}_0) &= F(\hat{w}_0)^{n-1} [1 - F(\hat{w}_0)] \\
M_{\hat{w}_0}(\hat{w}_0) &= [(n-1) \frac{F'(\hat{w}_0)}{F(\hat{w}_0)} - \frac{F'(\hat{w}_0)}{1 - F(\hat{w}_0)}] M(\hat{w}_0)
\end{aligned}$$

Using the fact that:

$$\frac{\partial}{\partial w_0} [N(\hat{w}_0)] = -L(\hat{w}_0)$$

it can be shown that:

$$
\begin{aligned}
\frac{d^2n}{du_0}\ln F(u_0) &= \frac{M_{u_0}(u_0)\ln F(u_0)}{N(u_0)} + \frac{L(u_0)M(u_0)\ln F(u_0)}{N(u_0)^2} \\
&= [(n-1)\frac{F'(u_0)}{F(u_0)} - \frac{F'(u_0)}{1-F(u_0)}]\frac{L(u_0)}{N(u_0)} + (\frac{L(u_0)}{N(u_0)})^2 \\
&= [(n-1)\frac{F'(u_0)}{F(u_0)} - \frac{F'(u_0)}{1-F(u_0)}][\frac{dn}{du_0}\ln F(u_0)] + [\frac{dn}{du_0}\ln F(u_0)]^2
\end{aligned}
$$

where $\frac{L(u_0)}{N(u_0)} = [\frac{dn}{du_0}\ln F(u_0)]$ from (3). We now have that:

$$
\begin{aligned}
RHS(A6) &= [(n-1)\frac{F'(u_0)}{F(u_0)} - \frac{F'(u_0)}{1-F(u_0)}][\frac{dn}{du_0}\ln F(u_0)] + [\frac{dn}{du_0}\ln F(u_0)]^2 \\
&\quad + \frac{dn}{du_0}\frac{F'(u_0)}{F(u_0)} + n[\ln F(u_0)'']
\end{aligned}
$$

Subtracting RHS(A6) from LHS(A6) gives:

$$
\begin{aligned}
&= [\frac{nF'(u_0)}{F(u_0)}]^2 + 2n\frac{F'(u_0)}{F(u_0)}[\frac{dn}{du_0}\ln F(u_0)] + [\frac{dn}{du_0}\ln F(u_0)]^2 \\
&\quad -[(n-1)\frac{F'(u_0)}{F(u_0)} - \frac{F'(u_0)}{1-F(u_0)}][\frac{dn}{du_0}\ln F(u_0)] - [\frac{dn}{du_0}\ln F(u_0)]^2 \\
&\quad -\frac{dn}{du_0}\frac{F'(u_0)}{F(u_0)} - n[\ln F(u_0)'']\} \\
&= [\frac{nF'(u_0)}{F(u_0)}]^2 + [2n\frac{F'(u_0)}{F(u_0)} - (n-1)\frac{F'(u_0)}{F(u_0)} + \frac{F'(u_0)}{1-F(u_0)}][\frac{dn}{du_0}\ln F(u_0)] \\
&\quad -\frac{dn}{du_0}\frac{F'(u_0)}{F(u_0)} - n[\ln F(u_0)''] \\
&= [\frac{nF'(u_0)}{F(u_0)}]^2 + [(n+1)\frac{F'(u_0)}{F(u_0)} + \frac{F'(u_0)}{F(u_0)[1-F(u_0)]}][\frac{dn}{du_0}\ln F(u_0)] \\
&\quad -\frac{dn}{du_0}\frac{F'(u_0)}{F(u_0)} - n[\ln F(u_0)'']
\end{aligned}
$$

Hence we can write:

$$
\begin{aligned}
\frac{1+H_{\tilde{w}_0}}{H(\tilde{w}_0)^2} &= \frac{F'(\tilde{w}_0)}{F(\tilde{w}_0)}\frac{dn}{d\tilde{w}_0}\{(n+1)\ln F(\tilde{w}_0)+\frac{\ln F(\tilde{w}_0)}{[1-F(\tilde{w}_0)]}-1\} \\
&+\{[\frac{nF'(\tilde{w}_0)}{F(\tilde{w}_0)}]^2-n[\ln F(\tilde{w}_0)'']\} \\
&= \frac{F'(\tilde{w}_0)}{F(\tilde{w}_0)}\frac{dn}{d\tilde{w}_0}\{(n+1)\ln F(\tilde{w}_0)+\frac{\ln F(\tilde{w}_0)}{[1-F(\tilde{w}_0)]}-1\} \\
&+\{[\frac{nF'(\tilde{w}_0)}{F(\tilde{w}_0)}]^2-n[\ln F(\tilde{w}_0)'']\}
\end{aligned}
\qquad \text{(A6)}
$$

Since $\frac{F'(\tilde{w}_0)}{F(\tilde{w}_0)} > 0$, $\frac{dn}{d\tilde{w}_0} < 0$ and $\ln F(\tilde{w}_0)<0$, every term in this expression is positive,hence we have that $[1+H_{\tilde{w}_0}(\tilde{w}_0)] > 0$.

B Data Appendix

B.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.

B.2 GSOEP Data

The GSOEP is a panel dataset from 1984 to the present consisting of some 13,500 individuals and roughly 7000 households living in West and East Germany. The international 'public use' version of data is used here, and this contains approximately 5% fewer observations. See Burkhauser (1991) for more details on the public use version.

Aside form the classification of movers discussed in section 5, the other important issue is the measurement of earnings and the weighting procedure used. We use reported gross monthly earnings, and where we provide a daily wage, we derive this by dividing the monthly figure by 30, in order to make it as comparable as possible to the measure used in the IAB data. As regards weighting, although the first wave of the survey (excluding immigrants) is representative of the non-immigrant German population, since we include immigrants in many of our estimations we weight all of our cross-sectional samples using cross-section sampling weights. Moreover, since attrition out of the survey is non-random, we generate longtitudinal weights based on the cross-section sampling weights and the attrition probabilities available in the GSOEP in all of our samples involving more than one year of data for the same individuals.

References

ACEMOGLU, D., AND J.-S. PISCHKE (1998): "Why Do Firms Train? Theory and Evidence," Quarterly Journal of Economics, 113(1), 79–119.

——— (1999): "The Structure of Wages and Investment in General Training," Journal of Political Economy, 107, 539–572.

AUTOR, D. H. (2001): "Who Do Temporary Help Firms Provide Free General Skills Training?," MIT Working Paper 01-03.

BAILY, N. M., G. BURTLESS, AND R. LITAN (1992): Growth with Equity: Economic Policymaking for the Next Century. Washington DC: Brookings Institution.

BECKER, G. S. (1962): "Investment in Human Capital: A Theoretical Analysis," Journal of Political Economy, 70(5), 9–49.

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.

BENDER, S., AND W. SCHWERDT (2000): "Why Do Firms Train?," Paper presented at 'Matched Employee-Employer Data' conference, IZA, Bonn, December 2000.

BISHOP, J. H. (1996): "What We Know About Employer-Provided Training: A Review of the Literature," CAHRS Working Paper 96-09, Cornell University.

BURKHAUSER, R. (1991): "Introduction to the German Socio-Econmoic Panel for English Speaking Researchers," Program Project Paper No. 1, Cross-National Studies in Ageing, New York, Syracuse University.

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 Beschaftigtenstatistik," Allgemeines Statistiches Archiv, 69, 56–68.

EUWALS, R. (1998): "Wages in the First Job After Apprenticeship: Movers versus Stayers," phD thesis, Tilburg University.

EUWALS, R., AND R. WINKELMANN (2001): "Why Do Firms Train? Empirical Evidence on the First Labour Market Outcomes of Graduated Apprentices," IZA Discussion Paper No. 319, IZA, Bonn.

FISHMAN, M. J. (1988): "A Theory of Pre-Emptive Takeover Bidding," Rand Journal of Economics, 19, 88–101.

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.

MATTHEWS, S. (1995): "A Technical Primer on Auction Theory I: Independent Private Values," Discussion Paper No. 1096, Northwestern University.

MCAFEE, R. P., AND J. MCMILLAN (1987): "Auctions with Entry," Economics Letters, 23, 343–347.

MILGROM, P. R., AND R. J. WEBER (1982): "Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis," Econometrica, 50, 443–459.

SCOONES, D. (2000): "Matching and Competition for Human Capital," Labour Economics, 7, 135–152.

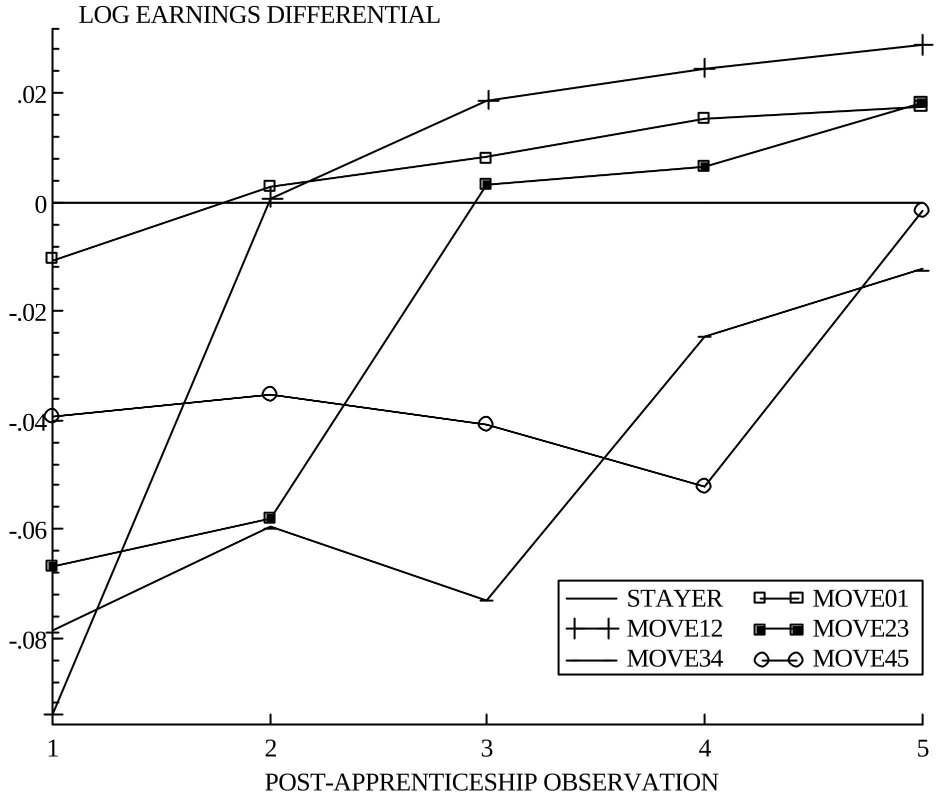
SOSKICE, D. (1994): "Reconciling Markets and Institutions: The German Apprenticeship System," in Training and the Private Sector - International Comparisons, ed. by L. M. Lynch, pp. 25–60. The University of Chicago Press, Chicago.

TOPEL, R. H., AND M. P. WARD (1992): "Job Mobility and the Careers of Young Men," Quarterly Journal of Economics, 107(2), 439–479.

WAGNER, K. (1998): "The German Apprenticeship System After Unification," WZB Discussion Paper FS I 98-30.

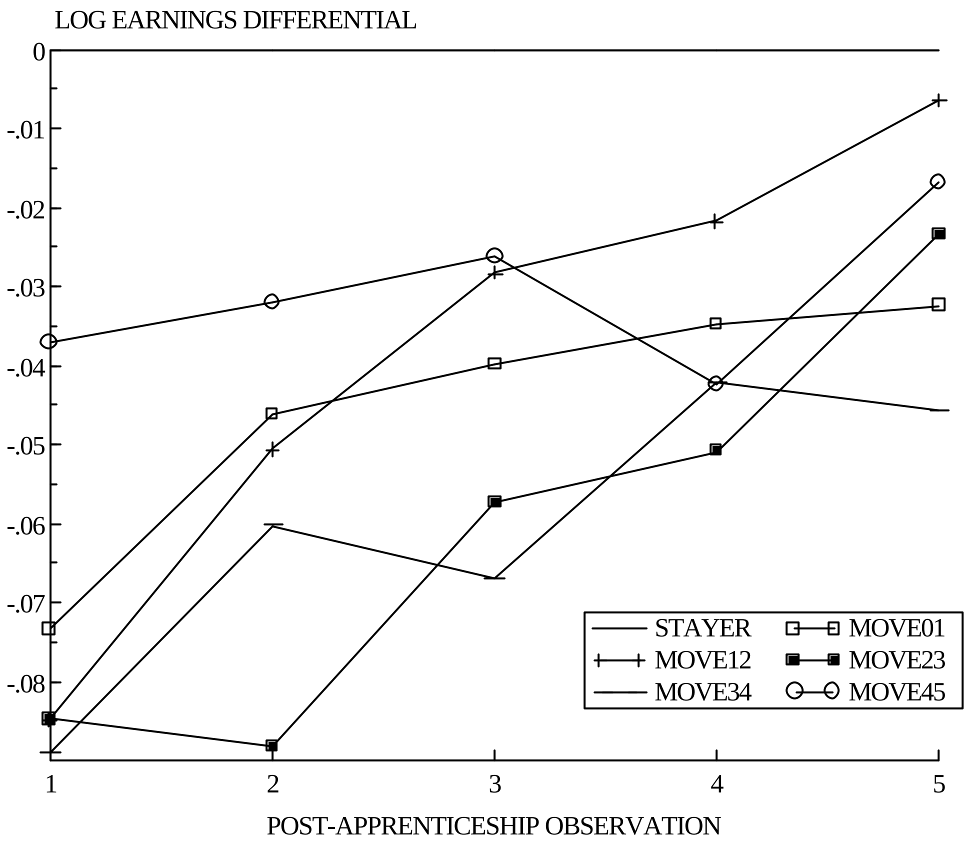
WERWATZ, A. (1996): "How Firm-Specific is German Apprenticeship Training?," Discussion Paper No. 96-12, SFB 373, HU Berlin.

Figure 1: Wage Dynamics of ‘Movers’ and ‘Stayers’: All Training Firms



Notes: Data drawn from results presented in Table 5a. Log earnings differentials refer to estimated coefficients in log earnings equations estimated at each of the five post-apprenticeship observations.

Figure 2: Wage Dynamics of ‘Movers’ and ‘Stayers’: Large Training Firms



Notes: Data drawn from results presented in Table 5a. Log earnings differentials refer to estimated coefficients in log earnings equations estimated at each of the five post-apprenticeship observations. See section 5 of the text for definitions of the different groups.

Table 1: Mobility from the Training Firm

	Sample Size	Number that Stay	Total Quits	Total Other	Total that Move	Proportion that Stay
Panel A: IAB Data						
At First Post-Apprenticeship Observation	22747	12781	N/A	N/A	9966	56.19
At Second Post-Apprenticeship Observation	21722	10087	N/A	N/A	11635	46.44
At Third Post-Apprenticeship Observation	21375	8616	N/A	N/A	12759	40.31
At Fourth Post-Apprenticeship Observation	20573	7294	N/A	N/A	13279	35.45
At Fifth Post-Apprenticeship Observation	19808	6247	N/A	N/A	13561	31.54
Panel B: GSOEP Data						
At First Post-Apprenticeship Observation	922	581	63	278	341	63.02
At Second Post-Apprenticeship Observation	811	397	116	298	414	48.95

Notes: See section 5 for a discussion of how the data are organised into post-apprenticeship observations

Table 2a: Descriptive Statistics: IAB Data

	STAYER	MOVE01	MOVE12	MOVE45
Training Firm Size				
1	2.05	2.57	2.10	1.31
2-9	13.05	28.90	24.67	21.40
10-19	10.11	14.52	16.63	15.28
20-49	12.00	14.79	18.55	13.10
50-99	8.92	7.71	9.56	11.35
100-499	18.68	16.14	15.30	18.34
500-999	7.99	6.01	4.40	5.68
1000+	27.21	9.35	8.80	13.54
First Post-Apprenticeship Observation				
Age	20.669	20.995	20.564	20.443
Days between Apprenticeship and Spell	356.242	287.684	362.497	349.401
Cumulative Days of Unemployment	2.480	17.530	6.034	6.432
Cumulative Days of Non-Employment	12.120	30.204	22.101	9.149
Cumulative Number of Firms	0	1.397	0	0
Daily Wage	**114.932**	**109.294**	**99.765**	**107.485**
Second Post-Apprenticeship Observation				
Cumulative Days of Unemployment	4.816	25.318	24.710	10.332
Cumulative Days of Non-Employment	13.189	34.430	30.731	10.345
Cumulative Number of Firms	0	1.631	1.189	0
Daily Wage	**122.920**	**118.730**	**117.063**	**115.926**
Third Post-Apprenticeship Observation				
Cumulative Days of Unemployment	7.429	30.714	31.607	11.413
Cumulative Days of Non-Employment	13.898	37.417	33.347	10.770
Cumulative Number of Firms	0	1.812	1.375	0
Daily Wage	**129.0362**	**125.2807**	**124.683**	**121.547**
Fourth Post-Apprenticeship Observation				
Cumulative Days of Unemployment	10.906	35.959	36.819	14.319
Cumulative Days of Non-Employment	14.346	39.534	34.538	12.660
Cumulative Number of Firms	0	1.959	1.522	0
Daily Wage	**133.943**	**130.754**	**130.656**	**125.088**
Fifth Post-Apprenticeship Observation				
Cumulative Days of Unemployment	13.425	40.827	40.863	31.474
Cumulative Days of Non-Employment	14.980	41.699	38.225	19.171
Cumulative Number of Firms	0	2.076	1.644	1.192
Daily Wage	**138.784**	**135.974**	**135.678**	**136.358**
N	2591	3346	523	229

Notes: see section 5 of the text for definitions of the different groups

Table 2b: Descriptive Statistics: GSOEP Data

	STAYER	MOVE01-QUIT	MOVE01-OTHER	MOVE12-QUIT	MOVE12-OTHER
Training Firm Size					
<20	0.238	0.365	0.330	0.412	0.433
20-200	0.351	0.382	0.355	0.401	0.373
200-2000	0.200	0.216	0.222	0.128	0.150
>2000	0.211	0.037	0.090	0.059	0.044
Sex	0.579	0.477	0.497	0.518	0.496
Nationality	0.938	0.934	0.933	0.932	0.917
At First Post-Apprenticeship Observation					
Age	20.705	19.979	20.102	19.104	19.728
DM	92.875	87.503	89.867	71.114	78.744
At Second Post-Apprenticeship Observation					
DM	107.792	101.839	107.788	97.914	93.494
N	519	51	257	52	70

Notes: see section 5 of the text for definitions of the different groups

Table 3: Previous Estimates of Mover-Stayer Wage Differential

	Acemoglu and Pischke (1998)		Harhoff and Kane (1997)		Euwals and Winkelmann (2001)	
Data	3 waves of QaC pooled		1985 wave of QaC		1% Social Security Waves 1975-1990 pooled	
Sample Selection	German men aged 23-59 employed full-time; finished education after 9 or 10 years; training firm and current employer in private sector		Men with more than 5 years of experience		Men in first job after apprenticeship; born 1960-1965	
Dependent Variable	Current Gross Monthly Wage		Current Gross Monthly Wage		Daily wage in first job after apprenticeship	
	SPECIFICATIONS		SPECIFICATIONS		SPECIFICATIONS	
	(1)	(2)	(1)	(2)	(1)	(2)
Training Firm Chars	>50	All	Industry	Craft	>100	All
	ESTIMATES		ESTIMATES		ESTIMATES	
Base Group	Movers		Movers		Movers	
Military Quitters	0.045 (0.025)	0.011 (0.014)				
Stayers	0.012 (0.015)	0.027 (0.008)	-0.019 (0.023)	-0.025 (0.020)	0.0406 (0.0144)	0.0003 (0.0079)
Training Firm Controls	SIZE & SECTOR	SIZE			SIZE & SECTOR	SIZE & SECTOR
Sample Size	5355	13051	2302	3711	2659	6451

Table 4a: Wages at First Post-Apprenticeship Observation: IAB Data

	(1)	(2)	(3)	(4)	(5)
Move	-0.0222*** (0.00564)	-0.0110*** (0.00447)	-0.0201*** (0.00430)	-0.0238*** (0.00523)	-0.0841*** (0.00659)
Age	0.0615* (0.0105)	0.0272*** (0.0105)	0.0233*** (0.0100)	0.000215 (0.0119)	0.0105 (0.0146)
Age^2 * 100	0.0270 (0.0236)	-0.0191 (0.0236)	-0.0116 (0.0225)	0.0371 (0.0260)	0.0103 (0.0324)
Days Unempl. * 1000		-0.199** (0.0562)	-0.0837*** (0.0539)	-0.276*** (0.0640)	-0.712*** (0.0942)
Days out Lab Force * 1000		-0.317*** (0.0283)	-0.246*** (0.0271)	-0.339*** (0.0330)	-0.286** (0.0402)
Training Firm Size Dummies	YES	YES	YES	YES	YES
Employer Firm Size Dummies	NO	NO	YES	NO	NO
Time Dummies	YES	YES	YES	YES	YES
N	17963	17963	17963	12866	6955

Notes: see text of section 6 for a discussion of the different specifications

Table 4b: Wages in First Post-Apprenticeship Observation: GSOEP Data

	(1)	(2)	(3)	(4)	(5)	(6)
Move	0.0620^{*} (0.0408)		0.0616^{*} (0.0416)		-0.0592 (0.0631)	
Move-Quit		0.0871 (0.0915)		0.0802 (0.092)		0.0743 (0.1176)
Move-Other		0.0569^{*} (0.0434)		0.0578^{*} (0.0443)		-0.0735 (0.0673)
Age	0.0420^{**} (0.0240)	0.0417 (0.0240)	0.0430^{**} (0.0246)	0.0428^{**} (0.0246)	0.0616^{**} (0.0316)	0.0572^{**} (0.0325)
$Age^2 * 100$	-0.00704 (0.0381)	-0.00671 (0.0381)	-0.0122 (0.0392)	-0.0119 (0.0392)	-0.0288 (0.0480)	-0.0221 (0.0494)
Male	0.175^{***} (0.0388)	0.176^{***} (0.0386)	0.158^{***} (0.0396)	0.159^{***} (0.0393)	0.104^{**} (0.0565)	0.100^{**} (0.0557)
German	-0.0156 (0.0572)	-0.0150 (0.0574)	-0.0164 (0.0587)	-0.159 (0.0589)	-0.0887^{*} (0.0618)	-0.0886^{*} (0.0618)
Training Firm Size Dummies	YES	YES	YES	YES	YES	YES
Employer Firm Size Dummies	NO	NO	YES	YES	YES	YES
Time Dummies	YES	YES	YES	YES	YES	YES
N	777	777	761	761	331	331

Notes: see section 5 of the text for a definition of the different groups, including the two types of mover (‘quit’ and ‘other’). See text of section 6 for a discussion of the different specifications

Table 5a: Wage Dynamics of Movers and Stayers: IAB Data

	Obsn 1	Obsn 2	Obsn 3	Obsn 4	Obsn 5
	Panel A: All Training Firm Sizes				
MOVE01	-0.0106* (0.0070)	0.00295 (0.00635)	0.00819* (0.00625)	0.0152*** (0.00604)	0.0177*** (0.00602)
MOVE12	-0.0940*** (0.0122)	0.000779 (0.0112)	0.0187** (0.00889)	0.0246*** (0.0106)	0.0288*** (0.0106)
MOVE23	-0.0788*** (0.0155	-0.0581*** (0.0131)	0.00336 (0.0130)	0.00636 (0.0126)	0.0182* (0.0125)
MOVE34	-0.0788*** (0.0129)	-0.0596*** (0.0141)	-0.0731*** (0.0139)	-0.0245** (0.0135)	-0.0124 (0.0134)
MOVE45	-0.0395*** (0.0148)	-0.0354*** (0.0158)	-0.0409*** (0.0156)	-0.0521*** (0.0151)	-0.00163 (0.0151)
Age	0.0113 (0.0169)	-0.0109 (0.0162)	-0.0116 (0.0167)	-0.0131 (0.0165)	0.00491 (0.0170)
Age^2	0.000155 (0.000385)	0.000685* (0.000352)	0.000533* (0.000347)	0.000539* (0.000331)	0.000204 (0.000327)
Days Unemployed	-0.294*** (-0.0859)	-0.200*** (0.0547)	-0.177*** (0.0433)	-0.176*** (0.0361)	-0.175*** (0.0321)
Days Out of Labour Mkt	-0.245*** (0.0418)	-0.284*** (0.0364)	-0.243*** (0.0349)	-0.206*** (0.0330)	-0.190*** (0.0321)
Training Firm Size	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
N	7292	7292	7296	7299	7303
	Panel B: Large Training Firms Only (>100 employees)				
MOVE01	-0.0734*** (0.00904)	-0.0462*** (0.00837)	-0.0399*** (0.00814)	-0.347*** (0.00797)	-0.0324*** (0.00816)
MOVE12	-0.0846*** (0.0182)	-0.0504*** (0.0169)	-0.0283** (0.0164)	-0.0217*** (0.0161)	-0.00641 (0.0164)
MOVE23	-0.0847*** (0.0194)	-0.0883*** (0.0179)	-0.0572*** (0.0175)	-0.0509*** (0.0172)	-0.0233* (0.0176)
MOVE34	-0.0890*** (0.0212)	-0.0603*** (0.0195)	-0.0670*** (0.0190)	-0.0422*** (0.0187)	-0.0460*** (0.0191)
MOVE45	-0.0370* (0.0234)	-0.0319* (0.0216)	-0.0261* (0.0210)	-0.0423*** (0.0206)	-0.0168 (0.0212)
N	2907	2909	2909	2908	2912

Notes: In all cases, the base group are the 'STAYERS' group. See section 5 of the text for definitions of the different groups

Table 5b: Wage Dynamics of Movers and Stayers: GSOEP data

	Observation 1		Observation 2	
	Panel A: All Training Firm Sizes			
MOVE01	-0.00373 (0.0442)		0.0396* (0.0301)	
Quit		0.0134 (0.0861)		-0.00015 (0.0475)
Other		-0.0096 (0.0485)		0.0504* (0.0332)
MOVE12	-0.179*** (0.071)		-0.0260 (0.0398)	
Quit		-0.230*** (0.0890)		-0.0113 (0.0541)
Other		-0.110 (0.107)		-0.0527 (0.0455)
Male	0.212*** (0.0434)	0.213*** (0.0431)	0.227*** (0.0276)	0.226*** (0.0272)
German	0.0376 (0.0748)	0.0397 (0.0746)	0.0118 (0.0327)	0.0103 (0.0325)
Age	0.0181 (0.0230)	0.0175 (0.0230)	-0.0093 (0.0184)	-0.0088 (0.0183)
Age^2	0.00023 (0.000364)	0.000245 (0.000364)	0.000619 (0.000297)	0.000612 (0.000298)
Training Firm Size	YES	YES	YES	YES
Year	YES	YES	YES	YES
N	509	509	519	519
	Panel B: Large Training Firms Only (>200 employees)			
MOVE01	-0.118** (0.0648)		0.0422 (0.0419)	
Quit		-0.0355 (0.122)		-0.122* (0.104)
Other		-0.130** (0.0717)		0.0670** (0.0398)
MOVE12	-0.170** (0.010)		-0.0560 (0.0664)	
Quit		-0.188 (0.164)		0.005 (0.0765)
Other		-0.144* (0.0966)		-0.135* (0.084)
N	211	211	217	217

Notes: See the text of section 5 for a discussion of the different groups and a definition of `quit' and 'other'.

Table A1: Coding of Different Types of Firm Change in GSOEP

	1985-86	1987-1990	1991-1995
'QUIT'	Quit	Quit	Quit
'OTHER'	Fired	Fired	Fired
	Mutual	Mutual	Transfer
	Transfer	Transfer	
	Contract End	Contract End	Contract End
	Training End	Training End	Training End
	Downsize	Downsize	
	End Self-Employment	End Self-Employment	End Self-Employment
	Other	Early Retirement	Retirement
		Other	Early Retirement
			Other
	Company Bankruptcy	Company Bankruptcy	Company Closed