WHY DO FIRMS TRAIN? THEORY AND EVIDENCE*

DARON ACEMOGLU AND JÖRN-STEFFEN PISCHKE

This paper offers a theory of training whereby workers do not pay for the general training they receive. The superior information of the current employer regarding its employees' abilities relative to other firms creates ex post monopsony power, and encourages this employer to provide and pay for training, even if these skills are general. The model can lead to multiple equlibria. In one equilibrium quits are endogenously high, and as a result employers have limited monopsony power and provide little training, while in another equilibrium quits are low and training is high. Using microdata on German apprentices, we show that the predictions of our model receive some support from the data.

I. INTRODUCTION

A large portion of human capital accumulation in the form of training and on-the-job learning takes place inside firms. Becker [1964] and Mincer [1974] provide a systematic explanation of training investments and the associated wages of workers. This standard theory draws a crucial distinction between general and firm-specific training. General training will increase a worker's productivity in a range of employment opportunities, and therefore will translate into higher earnings in a competitive labor market. Thus, it is the worker who has to pay for general training. The firm should pay only for the firm-specific component of training that does not help the worker receive higher wages elsewhere. However, these predictions seem to be at odds with reality.

In Germany, firms voluntarily offer apprenticeships to workers entering the labor market. Although general skills are an important component of these programs, much of the financial burden is borne by the firm (see Section IV for evidence on this). Similar apprenticeship programs also exist in other countries. In

* We thank Robert Gibbons, Lawrence Katz, Kevin Lang, Lones Smith, and two referees for useful comments. This research has been supported by a grant from the German Federal Ministry for Research and Technology to the World Economy Laboratory at the Massachusetts Institute of Technology. Acemoglu also acknowledges financial support from the National Science Foundation Grant SBR-9602116. Some of the data used in this paper have been obtained from the German Zentralarchiv für Empirische Sozialforschung at the University of Köln (ZA). The data for the study "Qualifikation and Berufsverlauf" were collected by the Bundesinstitut für Berufsbildung and the Institut für Arbeitsmarkt- und Berufsforschung and documented by the ZA. Neither the producers of the data nor the ZA bear any responsibility for the analysis and interpretation of the data in this paper.

The Quarterly Journal of Economics, February 1998

this paper we offer a theory in which, although skills are completely general in the sense that they can be used as effectively in other firms, the employer bears the costs of training.

In order to explain firms' investments in general skills, some labor market imperfection must exist so that the mobility of workers is restricted and that employers earn rents on trained workers. The mechanism we propose in this paper is the adverse selection story also analyzed by Waldman [1984], Greenwald [1986], Lazear [1986], and Gibbons and Katz [1991] among others. The abilities and skills of young workers who arrive in the market are in general unknown. Some will be more productive and abler at tasks in which they choose to specialize than will others. The early years of the worker's career constitute the period in which a substantial amount of information is accumulated about his ability. However, most of this information is gathered by the employer and is not transmitted to outsiders. The crucial step we take relative to the previous literature is to note the link between the firm's ex post informational *monopsony power* and its incentives to finance general training.

Given that the firm is able to obtain part of the marginal product of the worker, it also has an interest in increasing this marginal product by investing in the worker's human capital. Workers, on the other hand, may not be willing to pay for general skills themselves because they realize that a large fraction of the returns will be appropriated by the firm. As a result, in stark contrast to the standard Beckerian model of training, a firm that wants to attract workers to its apprenticeship program may have to pay not only for the training but also offer apprentices a training wage above their initial productivity! There is no need to resort to credit constraints of workers for this to be true.

The failure of the U. S. economy to generate as much training as Germany or Japan is sometimes blamed on the higher turnover in the United States (e.g., Blinder and Krueger [1996]). However, if training is general to a large extent, this statement cannot be reconciled with the standard model of training: with or without quits, firms will have to pay the full marginal product of the worker in the outside markets. Thus, exogenous differences in turnover alone do not explain the differences in the amount of general training. In our model quit rates determine the composition of the secondhand market for workers, and the wages that outside firms are willing to pay. Therefore, they influence the degree of monopsony power that the current employer has over the workers and, through this channel, its incentives to provide training. Moreover, because quit decisions are endogenous, our model is able to generate multiple equilibria. For example, in one equilibrium the economy could be characterized by high training and low quits: because workers who quit receive low wages, the apprenticeship firm will have substantial monopsony power that supports a high level of training. In another equilibrium the same economy will have low training and high quits: quitters would receive relatively high wages, leave their firms with a high frequency, and the firm would only enjoy moderate monopsony power and invest little in training. Paradoxically, the equilibrium with high quits, which achieves a better allocation of workers to jobs, may be less efficient since the level of training is lower.

While our theory, at least to us, has obvious appeal, a real test involves deriving new and different predictions and confronting them with data. We do this in the second part of the paper. Our test does not look directly at training but rather investigates the presence of adverse selection among those receiving firm-sponsored training. In particular, we look at the wages of German apprentices who stay in their firm, who are laid off, or quit their firm voluntarily, and compare them with the wages of apprentices who quite for *exogenous* reasons; that is, to go into the military. West Germany has had a draft system since 1957, but not all males get drafted. There is some screening mainly on physical fitness, but mostly the selection into the military is random. Our model predicts that military quitters should have higher wages than voluntary quitters (and laid-off workers), which is confirmed in the data. This test has an obvious similarity to Gibbons and Katz's [1991] empirical investigation of adverse selection in the U. S. labor market. However, we also draw a stronger prediction from our theory that is not shared by other models. Because the firm has monopsony power over workers who stay, they are paid below their marginal product. In contrast, military quitters, thanks to their exogenous reason of separation, are freed from this monopsony power. While they will be *less able* than stayers on average, they can have *higher wages*. In fact, we find that their wages are above those of stayers.

Apart from the literature on asymmetric information in the labor market, our paper relates to a number of other studies. Waldman [1990] shows that workers will not invest in general training when the incumbent employer has private information on whether such investment has taken place. His paper does not consider the investment incentives of the firm. Katz and Ziderman [1990] first proposed the idea that the firm may pay for training if the amount of the training investment is not observed by outsiders. Chang and Wang [1995, 1996] analyze a model with this feature in more detail. Imperfect information about the training level limits the wages that a trained worker can obtain in the outside market, and it gives a monopsony power to the employer that supports firm-sponsored investments in training. This is different from the mechanism we investigate where it is the *ability* of workers, not their training, that is unobserved. At least in the context of the German apprenticeship system, this appears a more plausible assumption. The curricula of apprenticeship programs are regulated, and apprentices take standardized exams. Thus, there would be relatively little uncertainty about the amount of training investments that a worker has received. In contrast, since these workers are relatively young, there will be considerable uncertainty about their abilities which will unravel slowly, and which will be observed mostly by the current employer. Also, despite the similarity between the models, the empirical implications we will test do not follow from the models of Katz and Ziderman and Chang and Wang.

Other papers have also obtained multiple equilibria in turnover and wages in the presence of informational imperfections but with considerably different formalizations and without considering the implications for general training (see Abe [1994], Prendergast [1992], and Glaeser [1992]). Acemoglu [1997] discusses search-induced monopsony as a different mechanism that would support firm-sponsored investment in general training. In Acemoglu and Pischke [1996b] we generalize these ideas and show that all models in which firms pay for general training rely on similar principles: that rents exist in the labor market and that these rents increase in the skill level of workers. Nevertheless, the predictions we test here are specific to asymmetric information.

The remainder of the paper is organized as follows. In the next section we lay out the basic model. Section III discusses a number of extensions. Empirical implications of the model and corroborating evidence are presented in Section IV. Section V describes the data, and Section VI presents the empirical results. In Section VII we conclude by drawing the implications of our model for the different labor market outcomes and institutions of the United States and Germany.

II. The Basic Model

We start with the simplest model to obtain the sharpest predictions. The next section deals with extensions aimed at demonstrating the robustness of the model.

A. Description of the Environment

The world lasts for two periods. Firms can hire workers at the beginning of either period. We assume that firms have access to a technology that is linear in the number of workers they hire, and each firm employs a large number of workers. The productivity of an individual worker is never observed by outsiders, and is therefore noncontractible. All firms and workers are risk neutral. Also, there is no discounting between periods. During the first period, firms can offer training to the workers they hire. The amount of this training is denoted by τ. With or without training, workers produce nothing during the first period (a normalization).

In the second period each worker produces

$$y = \alpha(\tau)\eta, \tag{1}$$

where $\alpha(\tau)$ is the general human capital and η is the ability of the worker.

The assumption that all human capital is general is of course extreme, but it will demonstrate that none of our results are driven by firm-specific human capital. It is important for our results that ability and training are complementary; the multiplicative specification of (1) captures this in a simple fashion. The cost of training each worker is given by τ, and this amount is incurred by the firm. Regarding the return from training, $\alpha(\tau)$, we make the following assumption:

ASSUMPTION 1. $\alpha(\tau)$ is twice continuously differentiable, strictly increasing, and concave with $\lim_{\tau\to 0} \alpha'(\tau) = \infty$ and $\lim_{\tau\to\bar{\tau}} \alpha'(\tau) = 0$ for some $0 < \bar{\tau} < \infty$.

The Inada-type condition embedded in this assumption implies that it is always socially optimal to have a positive amount of training.

Ability is distributed according to a distribution function $F(\eta)$ with support $[0,\eta^{sup}]$. The distribution from which these abilities are drawn is common knowledge, but in the first period no one knows the ability of a worker (our analysis would apply exactly as it is if workers knew their ability). At the end of the first period, the firm learns the ability of each worker it employs. It is important to note that this is independent of whether and how much training is offered to the worker. *Training neither facilitates nor leads to learning.* The reason firms train will *not* be because training enables them to learn better.

In our economy there are also reasons for quits: workers may not get along with their colleagues or supervisor or wish to move to a different city for family reasons. To capture these and other reasons for voluntary quits, we assume that with probability λ each worker receives a disutility shock θ. If the worker stays with the same firm, he suffers disutility θ and avoids this cost if he quits and finds a new job. The disutility shock θ has a distribution function $G(\theta)$. This shock is never observed by any agent other than the worker in question.

We now make some assumptions about the distribution functions.

ASSUMPTION 2. (i) $G(x)$ is twice differentiable and is log-concave in x. (ii) $F(x)$ is twice differentiable and $\int_x \eta \, dF(\eta)$ is log-concave in x.

Log-concavity ensures well-behaved second-order conditions and comparative statics results [Burdett 1979]. It simply requires that the log of the function in question is concave, and many well-known distribution functions including uniform and normal satisfy this requirement. These assumptions are by no means necessary for our results, but they simplify the exposition and the proofs.

The exact sequence of events in this economy can be summarized as follows.

1. In period 1 firms decide how many workers to hire and how much training to offer to each worker. At this point firms do not know the ability of workers. Thus, *they cannot offer more training to abler workers.*
2. At the end of period 1 the incumbent firm finds out about the ability η of each worker.
3. At the beginning of period 2 the incumbent firm chooses which workers to lay off and offers a uniform wage $w(\tau)$ to all retained workers.[1]

1. Since the output of an individual worker is not verifiable, neither the first employer nor future employers can write wage contracts contingent on the output of the individual worker, and given that each firm has a large number of workers, conditioning on the total output of the firm is not useful. This implies that the only

4. Next, workers learn their disutility shock θ, and the workers who are retained choose whether to stay ($\hat{q} = 0$) or to quit ($\hat{q} = 1$).
5. Outside firms make wage offers $v(\tau)$ to those workers who are in the secondhand market. These workers either are laid off or have quit voluntarily, but firms in the secondhand labor market cannot distinguish layoffs from quits.

A couple of features need to be noted. First, $v(\tau)$, the offer of the outside firms, is conditioned on τ because outside firms observe the training level of workers in the secondhand market, and workers with different levels of training will have different productivities. Since outside firms do not know whether the worker has quit or has been laid off, $v(\tau)$ cannot be conditioned on this information. This seems to be a plausible assumption. Suppose, on the contrary, that firms could distinguish quits from layoffs. Then a worker who anticipated being laid off would jump the gun and quit immediately, destroying the informational content of the distinction between quits and layoffs.[2] Finally, the sequence of events allows no poaching: outside firms cannot make offers to employed workers. This restriction will be relaxed in Section III.

To close the model, we also impose a free entry condition on firms at all points in time. Thus, no firm will earn positive profits in equilibrium. As we will see, this may imply that apprentices (workers in period 1) need to be paid a positive wage even though they are not productive during this period. We denote this wage by W. We also assume that firms only compete for workers by offering a training wage W, which is not contingent on the level of training, and then *unilaterally* decide the level of training to maximize profits. This basically corresponds to assuming that contracts on training levels are hard to enforce: for instance, the amount of training is not perfectly verifiable (see Section III for an alternative assumption).

B. Equilibrium with Symmetric Information

The model we have is noncompetitive due to the informational asymmetries. We first want to derive the equilibrium of the

type of contracts that are allowed are flat wage contracts. Contracts contingent on ability will be discussed in Section III and will be shown not to affect our results.

2. In part of our sample we can distinguish between workers who report having quit voluntarily and those who report having been laid off. There is no systematic difference between the future wages of these two groups.

economy in the absence of these informational asymmetries in order to clarify our contribution.

Suppose that the ability level η of each worker is still unobserved in the first period, but in the second period it becomes common knowledge (or alternatively that $F(\eta)$ is degenerate at some level η^*). This implies that the secondhand labor market is perfectly competitive. Since ability is observed, the secondhand wage is now conditional on ability; i.e., $v(\eta,\tau) = \alpha(\tau)\eta$. Any worker can quit and obtain this wage. Therefore, no firm will be able to retain its trained workers unless it pays the full marginal product (the full contribution of the worker to the firm which is y). This immediately implies that the training firm will enjoy no rents in the second period, and there will be no training unless the worker is willing and able to pay for his training. If we impose the credit constraint that $W \geq 0$, the model delivers $\tau^0 = 0$. On the other hand, if W is unconstrained, then we have the optimal amount of training: $\tau^c = \arg\max \int \eta\alpha(\tau)\, dF(\eta) - \tau$, and the corresponding training wage is $W = -\tau^c$. Therefore, in the absence of credit problems, the worker pays for his general training by taking a wage sufficiently below his marginal product to cover the cost of training. Note that because in the first period abilities are still unknown, even in this model, all types receive the same amount of training, τ^c. The most important finding for future comparison is that in the absence of information asymmetries, workers are paid their full marginal product in the second period, and therefore, *firms never pay for general training.* We summarize these results in Propsition 1.

PROPOSITION 1 (BECKER). In the absence of information asymmetries, firms never pay for general training. If there are credit constraints such that $W \geq 0$, then there is no training: $\tau = \tau^0 = 0$. If there are no credit constraints, then $W = -\tau^c$ and $\tau = \tau^c$, where τ^c is the efficient amount of training in this economy.

C. Equilibrium with Adverse Selection

In the rest of the paper we will analyze the case in which $F(\eta)$ is nondegenerate and η is not observed by firms other than the current employer. Also, to expose the differences between our approach and the standard one, we will consider the case where workers are severely credit constrained such that $W \geq 0$.

The equilibrium concept we will use is Perfect Bayesian.[3] An equilibrium is a set of functions and numbers $\{\hat{q}(\tau), v(\tau), w^*(\tau), \hat{\eta}^*(\tau), \tau^*, W^*\}$ such that

1. All workers take the optimal quit decisions $\hat{q}(\tau)$ given the wage offer function of their firm, $w^*(\tau)$, and the outside wage $v(\tau)$.
2. All apprenticeship firms (firms with trainees) choose the optimal wage policy $w^*(\tau)$, layoff strategy $\hat{\eta}^*(\tau)$, training τ^*, and training wage W^*, taking as given the outside wage offers $v(\tau)$ and the workers' optimal quit decisions $\hat{q}(\tau)$.
3. All outside firms make optimal offers, $v(\tau)$, to workers looking for a job in the second period, given their beliefs regarding the ability of these workers.
4. The beliefs of outside firms are derived by Bayes' Rule from the quit behavior of workers, $\hat{q}(\tau)$, and the layoff behavior of firms, $\hat{\eta}^*(\tau)$.

We can now characterize the equilibria. A worker with training τ who quits receives $v(\tau)$, whereas if he stays he will get $w(\tau)$, but he will also suffer disutility θ. The optimal decision is to quit, $\hat{q}(\tau) = 1$, iff

$$v(\tau) - w(\tau) + \theta \geq 0,$$

and $\hat{q}(\tau) = 0$ otherwise, where θ will take a value from the distribution $G(\theta)$ with probability λ and is equal to zero with probability $1 - \lambda$. Therefore, as long as $w(\theta) \geq v(\tau)$ (which will be the case in equilibrium), the equilibrium probability of quitting, conditional on the wage offer of the firm and the outside wage, is

$$q[w(\tau),v(\tau)] = \lambda[1 - G[w(\tau) - v(\tau)]].$$

Because the firm is forced to pay the same wage to all remaining workers of the same training level, its layoff decision will take a simple cutoff form whereby all workers above a skill level $\hat{\eta}(\tau)$ will be kept.

The maximization problem of the firm, given optimal quit

3. In this model, all Perfect Bayesian Equilibria satisfy the Intuitive Criterion of Cho and Kreps [1987] and thus do not need to be refined. See Acemoglu and Pischke [1996a] for details.

behavior and outside wages, can be written as

$$(2)\quad \max_{w(\tau),\hat{\eta}(\tau),\tau} \Pi = [1 - q[w(\tau),v(\tau)]] \int_{\hat{\eta}(\tau)}^{\infty} [\alpha(\tau)\eta - w(\tau)]\, dF(\eta) - \tau - W.$$

In words, the firm takes the quit behavior of its workers as given and maximizes its profit by choosing the cutoff ability for layoffs $\hat{\eta}^*(\tau)$, wages to be paid to workers of different training levels, $w^*(\tau)$, and the training level τ^*. In equilibrium we only observe $\hat{\eta}^*(\tau^*)$ and $w^*(\tau^*)$. These functions still need to be determined as part of the equilibrium strategies in order to pin down the off-the-equilibrium-path behavior.

We will first characterize $\hat{\eta}^*(\tau)$ and $w^*(\tau)$ and then return to τ^*, and finally, the training wage W will be determined from the zero-profit condition. The first-order conditions are[4]

$$(3)\quad \hat{\eta}^*(\tau) = \frac{w^*(\tau)}{\alpha(\tau)},$$

$$(4)\quad -[1 - q(w^*(\tau),v(\tau))][1 - F(\hat{\eta}^*(\tau))] - \frac{\partial q[w^*(\tau),v(\tau)]}{\partial w(\tau)} \int_{\hat{\eta}^*(\tau)}^{\infty} [\alpha(\tau)\eta - w^*(\tau)]\, dF(\eta) = 0,$$

for all τ. These two first-order conditions are fairly intuitive. Equation (3) says that the firm will lay off all workers on whom it would lose money. Equation (4) determines the optimal wage by weighing two factors: a higher wage leads to lower profits (the first term), but it also enables the firm to reduce voluntary quits (the second term).

The final first-order condition is with respect to τ. Differentiating (2), making use of the other first-order conditions and the fact that $\partial q/\partial w = -\partial q/\partial v$, we can write this condition as

$$(5)\quad [1 - q[w^*(\tau^*),v(\tau^*)]] \int_{\hat{\eta}^*(\tau^*)}^{\infty} [\alpha'(\tau^*)\eta - v'(\tau^*)]\, dF(\eta) = 1.$$

Next, if profits are positive, some other firm will offer a slightly higher level of the training wage, and attract all trainees in period

4. The second-order conditions are satisfied, given Assumption 2; see Acemoglu and Pischke [1996a].

1. In equilibrium we need to have $\Pi = 0$. Thus, using (2),

$$(6) \quad W^* = \max\{0; [1 - q[w^*(\tau^*), v(\tau^*)]] \int_{\hat{\eta}^*(\tau^*)}^{\infty} \cdot [\alpha(\tau^*)\eta - w^*(\tau^*)]\, dF(\eta) - \tau^*\}.$$

Finally, outside wages for quitters need to be determined. Due to free entry, these workers will be paid their expected marginal product, which by Bayes' rule gives

$$(7) \quad v(\tau) = \frac{q[w^*(\tau), v(\tau)]\alpha(\tau)\bar{\eta} + [1 - q[w^*(\tau), v(\tau)]] \int_0^{\hat{\eta}^*(\tau)} \alpha(\tau)\eta\, dF(\eta)}{q[w^*(\tau), v(\tau)] + [1 - q[w^*(\tau), v(\tau)]]F(\hat{\eta}^*(\tau))}.$$

The secondhand market consists of workers who have quit and others who have been laid off. The numerator in (7) has two terms: the first is the probability that he has voluntarily quit (q) times the average productivity ($\alpha(\tau)\bar{\eta}$), and the second is the probability that he has not quit but was laid off times the productivity of an average worker who is laid off. The denominator is the total probability that a worker is found in the secondhand market.

PROPOSITION 2. (i) An equilibrium, with $w^*(\tau)$, $v(\tau)$, and $\hat{\eta}^*(\tau)$ continuously differentiable, exists and is characterized by (3)–(7). (ii) Denote the set of equilibrium training levels by $\mathcal{T}$. Then $\forall \tau^a \in \mathcal{T}$, $\tau^0 = 0 < \tau^a < \tau^c$, where τ^0 and τ^c are the training levels in the economy without informational asymmetries, respectively, with and without credit constraints.

The proof of this proposition, like all others that follow, is in the Appendix. Part (i) establishes the existence of a well-behaved equilibrium. For the purposes of this paper, part (ii), which characterizes some key properties of the equilibrium allocation, is more important. It establishes that the amount of training is always positive but less than first-best. Therefore, in contrast to the case of no information asymmetries, the equilibrium with asymmetric information always has a *positive level of training sponsored by the firm.*

Expressed differently, because in our economy the firm will have an ex post monopsony power over its employees, it will be able to capture a portion of their product, and has therefore an incentive to increase the productivity of its workers. In contrast, recall that in the competitive equilibrium the firm always paid the marginal product and thus had no incentive to invest in skills. As remarked above, the multiplicative structure of ability and training is not crucial for this result. Instead, the important feature is that $\alpha(\tau)$ and η are complements. To see this, note that if (1) were additive, i.e., $y(\tau) = \alpha(\tau) + \eta$, then the firm would still have ex post monopsony power, but the amount of rents it could obtain in the second period due to this monopsony power would be completely independent of $\alpha(\tau)$. All increases coming from $\alpha(\tau)$ would be reflected in the worker's outside wage, i.e., $\alpha'(\tau) = v'(\tau)$, yielding an equilibrium training level of $\tau^* = 0$. However, with the complementarity between training and ability, outside wages are determined by the productivity of low ability workers who are abundant in the secondhand pool (e.g., imagine the case with $q \approx 0$). Then the wages that the initial firm has to pay to retain workers are also closely related to the marginal product of low ability workers, but their revenue is proportional to the marginal product of high ability workers who are retained. This wedge enables the firm to increase profits by training. This is an important insight as it implies that asymmetric information per se is not sufficient to make firm-sponsored training profitable. Rather it is the interaction between the asymmetry of information and the complementarity of ability and training that distorts the equilibrium wage structure and induces firms to invest in general skills.[5]

It is also intuitive that the equilibrium level of training is less than the first-best amount τ^c. The firm lays off some workers, and other workers quit. Since training these workers is costly but the firm only reaps the benefits from workers who stay with the firm, incentives to provide training are suboptimal. In particular, the firm is bearing the cost of training, but an increase in training also benefits the workers who quit or are laid off. Next,

PROPOSITION 3. (i) There can exist multiple equilibria. (ii) Suppose that there exist multiple equilibria. Then for any two equilibria, a and b, we have that $\forall\tau$, $v^a(\tau) > v^b(\tau) \Leftrightarrow q^a[w^a(\tau),v^a(\tau)] > q^b[w^b(\tau),v^b(\tau)] \Leftrightarrow \tau^a < \tau^b$.

Our model can therefore generate multiple equilibria: one with high training and low turnover, and the other with low

5. See Section IV for evidence on the complementarity of ability and training. Alternatively, our model would also generate firm-sponsored general training if training is not complementary to ability but facilitates learning about ability.

training and high turnover. It is relatively easy to construct sensible examples of multiple equilibria and verify the results of Proposition 3 (see Acemoglu and Pischke [1996a]).

Let us explain the intuition for this proposition. The right-hand side of (7) is increasing in the probability of quitting, q. That is, at a given level of training τ, $\partial v(\tau)/\partial q > 0$. More voluntary quits imply that the quality of the labor pool in the secondary market is higher. To see this, consider the case in which $q = 0$. Then all workers who are looking for a job must have been laid off, and thus they must have a level of ability less than $\hat{\eta}$. In contrast, when $q > 0$, some of the quitters may be of very high ability but will nevertheless leave voluntarily because they are not happy with their jobs. Hence, the higher the quit probability, the higher the average ability in the secondhand pool, and the higher the outside wage.

Furthermore, from the definition of q we see that $\partial q[w^*(\tau),v(\tau)]/\partial v(\tau) = \lambda g[w^*(\tau) - v(\tau)] > 0$. The same relation holds when the indirect effect of $v(\tau)$ is taken into account (see the Appendix). In other words, higher outside wages imply that workers who are unhappy can leave more easily, and as a result voluntary quits are higher. Therefore, at a given level of training, equilibrium wages and quit behavior can be determined as the intersection of two upward sloping curves—hence the multiplicity. Next, with higher quits, profits from training are lower because the firm has less ex post monopsony power over trained workers. This induces less training in the equilibrium with high quits, which explains the second part of the proposition.

Given the multiplicity, it is natural to ask whether an equilibrium with a higher level of training Pareto dominates one with lower training. The answer is ambiguous. To see why, note that since $\tau < \tau^c$, increasing the level of training is in general beneficial. However, an equilibrium with higher training also has lower quits. Therefore, more workers who have received negative utility shocks and who would be better off finding a new job are forced to stay in their old firm because of the stigma of quitting. This means that the allocation of workers to jobs is more distorted. In general, it is not possible to ascertain which of these two forces will dominate. So, when there are multiple equilibria, the equilibrium with high training has better investment but a worse allocation of workers to firms.

We see next that even when workers can contribute to the financing of training, the firm may end up bearing all the costs.

PROPOSITION 4. There exists an open set of parameter values such that W^* is strictly positive.

This proposition implies that the constraint $W \geq 0$ may not be binding. Therefore, even when workers are not credit constrained, in stark contrast to Becker's theory, *they may not pay for their training!* Hence, we obtain similar results without assuming that workers are credit constrained. The underlying reason is once more the firm's ex post *monopsony* power over trained workers. Depending on the equilibrium beliefs, posttraining wages may be sufficiently low so as to necessitate an additional payment to workers ex ante; that is, $W^* > 0$. More explicitly, hiring a worker in period 1 is to obtain the rights to the informational monopsony in period 2, and given free entry, firms have to pay a positive price for this right to future profits. They pay this price partly by bearing the costs of training, but this may not be enough, and positive wages could also arise. Further, it is useful to note at this point that irrespective of whether $W^* = 0$ or > 0, compensation is *front-loaded* in this economy compared with the allocation without the information asymmetries, in the sense that the earnings profile is flatter over time than is the productivity profile.

III. SOME EXTENSIONS

In this section we discuss four possible extensions of the basic model. We present the basic results and intuition here; more formal statements of these extensions can be found in the working paper version of this article [Acemoglu and Pischke 1996a].

A. Different Wages for Different Abilities

A simplifying assumption that we used so far is that the wage offer of the apprenticeship firm in the second period is not conditional on ability. We now allow the wage to be of the form $w(\eta,\tau)$. In the spirit of our previous analysis, we assume that outside firms cannot observe the wage that a worker is paid. There will now be three differences as compared with our baseline analysis:

1) the firm chooses $w(\eta,\tau)$ such that $\partial w(\eta,\tau)/\partial\eta > 0$;
2) the firm lays off no worker since it can pay as low a wage as it wishes;
3) the probability that a worker will quit is no longer independent of his ability.

The main similarity is that the outside wage $v(\tau)$ is still not dependent on ability since firms in the secondhand labor market observe neither ability nor the wage offer that a quitter received before leaving. Then, because high ability workers receive higher wage offers from their employer but face the same outside wage, q will be decreasing in η. Therefore, high ability workers will quit with a lower probability, implying again that the secondhand labor market has a disproportionate number of low ability workers. Hence $v(\tau)$ does not increase linearly with $\alpha(\tau)\overline{\eta}$, and initial employers will have an incentive to invest in the skills of their employees. In fact, exactly the same argument as in the proof of Proposition 2 establishes that the equilibrium level of training is always strictly positive. Also, multiple equilibria are again possible due to the same reasons as above. Therefore, our results are not sensitive to the assumption that the firm makes a unique wage offer. What is important is that outsiders do not observe ability or the wage that would reveal it.

B. Equilibrium with Verifiable Training

In our main analysis the firm offered only a training wage to attract workers and decided unilaterally about the training level. We can alternatively allow firms to offer a contract $[W,\tau]$ to attract workers to their training programs. Suppose that $\{v(\tau),w(\tau),\eta(\tau),\tau^*,W\}$ is an equilibrium of Section II (i.e., when firms offer wages to attract workers and then decide about training), then it can be shown that there exists an equilibrium with verifiable training $\{v(\tau),w(\tau),\eta(\tau),\hat{\tau},\hat{W}\}$ such that $\tau^c > \hat{\tau} \geq \tau^*$. Therefore, when firms offer wages and training packages to attract workers, the second stage of the game (layoff, quit, and hiring decision *functions*) are unchanged, but firms offer more training. Intuitively, a firm can attract workers more cheaply by offering higher training and a lower W because higher training increases the expected future earnings of the worker, and he accepts employment at a lower initial wage. Nonetheless, this does not restore the level of training to the first-best, τ^c. This is first because, with the constraint $W \geq 0$, the firm may not be able to reduce W. Second, training and quits interact in the second stage. More training raises the wedge between the inside and the outside wage and therefore induces workers not to quit, even when they are unhappy in their job. This reduction in quits reduces the expected return from training for the workers relative to the case of perfect information where training and quit decisions were perfectly separated.

C. Raids on Trained Workers

So far, we have only allowed outside firms to hire quitters which can be thought of as *passive poaching*. The alternative is of course *active poaching* or *raids* whereby a firm makes an offer to a currently employed worker (see for instance Lazear [1986]). The presence of raids does not destroy the informational advantage of the initial employer because of a *winner's curse*. If an outside firm makes an offer to a very productive worker, the initial employer is likely to make a counteroffer (as long as the worker can credibly demonstrate that he has an outside offer). Only when the outside offer is above the productivity of the worker, will the firm prefer to let him go. To see the implications of this winner's curse on training, first consider the case with $\lambda = 0$. Thus, no worker receives a disutility shock. In this case, it is a dominated strategy for an outside firm to make a raid on trained workers. Suppose that it offers a wage $v^R > w$. Then, the incumbent employer will make a counteroffer to all workers with productivity greater than v^R, but not to those who produce less than v^R. Therefore, the raider will certainly lose money, and the equilibrium characterized in Section II remains unchanged. Things are little more involved when $\lambda > 0$ because now the raider may attract some of the workers with $\theta > 0$ (i.e., those who are unhappy in their current firm). It can be shown in this case that there is a *unique* equilibrium which is the same as the one with the lowest amount of training among the equilibria characterized in Proposition 2. In other words, the possibility of raids destroys the low-quit/high-training equilibrium of our baseline model. Intuitively, raiders can always offer the highest possible outside wage and induce unhappy workers to quit, whereas the low-quit/high-training equilibrium relies on the fact that very few workers quit. Thus, the outside wage remains low. This result also shows that the institutional structure of the labor market, e.g., whether raids are possible or not, may have important consequences for training.

D. Multiple Periods

It is also straightforward to extend our analysis to multiple periods. The exact results depend on how the information structure evolves over time. The most natural starting point seems to be the one where training takes place in the first period and there are multiple periods thereafter. If the firm discovers worker ability at the end of the first period, then a worker can stay with the firm for one period and then quit at the end of period 2, signaling to the outside market that he is not low ability. In this case, the monopsony power of the initial employer will disappear after the second period. However, in a more realistic model the firm would learn slowly about workers' abilities. For example, with a certain probability χ_t the firm learns the ability of the worker at the end of period t. Then, as long as $\chi_1 < 1$ and $\chi_2 > 0$, even workers who leave their firms at the end of period $t = 2$ could be lemons just discovered and laid off by the employer. If $\chi_t = \chi$, then the informational monopsony power of the firm would remain unchanged over time. In contrast if $\chi_t > \chi_{t+1}$, then this monopsony power would decline but not necessarily disappear. The implications of this observation are important for our empirical work below, since we use data on workers of all ages. Thus, we are assuming that the monopsony power present in the market never completely disappears; i.e., $\chi_t > 0, \forall t$.

Alternatively, firms may also be able to delay apprenticeship until they have learned about the ability of the workers. This is, however, costly for two reasons: first, workers would not be realizing their full productive potential before apprenticeship, and in many lines of business there will only be a limited amount they can do without the required skills. Second, by offering different amounts of training to workers of different abilities, firms would reduce their informational monopsony, and thus lose rents. A full analysis of this case is beyond the scope of this paper.

IV. Testable Implications and Empirical Strategy

We now draw implications from the model that we will test below with data from Germany. Our test does not directly involve the incidence or the amount of training that workers receive. We use data on wages of workers who have gone through an apprenticeship program to test whether this segment of the labor market is indeed characterized by adverse selection, the mechanism responsible for training in our model.

There are three main reasons why apprentices leave their firm in Germany: (i) their employer does not offer a permanent contract after the apprenticeship, (ii) workers quit voluntarily to find other employment; (iii) men are drafted to do military service. We assume that firms are unable to distinguish between voluntary quits and layoffs, but can observe whether a worker has just come back from military service. The fact that there is another group of workers which separates for military service does not change any of our theoretical analysis, but enables a sharp test of our mechanism. In our data we can distinguish workers who stayed with their training firm initially, those who left for military service and then work for a different employer, and workers who leave for other reasons. We do not know whether workers who stay with their apprenticeship firm served in the military. The model implies the following:

1. Stayers earn more than laid-off workers and voluntary quitters; that is, $w^*(\tau) > v(\tau)$. It is easy to see that this has to be true. If $w^*(\tau) < v(\tau)$, stayers would quit and be better off in the secondhand market.
2. Military quitters leaving for largely exogenous reasons earn more than other quitters; that is, $v_m(\tau) > v(\tau)$. The reason they earn more than quitters is that they are of higher average ability, and both groups are paid their average marginal product in the secondhand market. This is true even if there is some selection among military quitters, as long as training firms rehire only very high ability workers after they complete their service. If there is some cost to doing so (for instance, in the case that firms need to keep certain positions open for these workers), training firms will be more selective in rehiring workers who have done military service than in retaining other workers. This more stringent selection implies higher wages for military quitters than for other quitters.
3. Military quitters *may* earn more than stayers; i.e., $v_m(\tau) \gtreqless w^*(\tau)$. The reason for this is that military quitters are freed from adverse selection (at least to some degree). For example, with no rehiring of military quitters by the original firm, they are paid the average marginal product of the group. Stayers, on the other hand, are paid below their marginal product due to the informational monopsony of the firm. If the monopsony power of the training firm over stayers is large enough compared with the quality of the pool of military quitters, the equilibrium wage for stayers may be below the wage of workers separated for an exogenous reason (see Acemoglu and Pischke [1996a] for a numerical example). However, cru-

cially, in all cases the average ability of stayers is above the average quality of military quitters.

Our empirical strategy to test these implications is to run a wage regression controlling for worker characteristics known at the time the apprenticeship starts (e.g., information on schooling and on the type of apprenticeship firm). Dummy variables for stayers and military quitters will let us discern the wage differentials for these two groups relative to other quitters.

In our data only a small fraction of workers leaves the training firm because of military service. Some workers serve in the military right after their apprenticeship ended, while others are drafted only when they already started working in a permanent position. We treat these differences in the timing of the draft as random (after controlling for the previous level of schooling completed). Some apprentices who get drafted right after their apprenticeship will return to their original employer. The fact that these workers will be classified as stayers in our data will tend to weaken, but not invalidate, the implications described above. Military quitters will still earn more than other quitters. But it will be less likely that we observe military quitters earning more than stayers. This is because the group of stayers will include some workers who have served in the military but are rehired by the training firm. They command a higher wage on the outside market and consequently from the incumbent firm.

This test provides only indirect evidence that firms might finance apprenticeship training in Germany. It is possible that adverse selection exists in the market for trained apprentices but firms do not actually pay for training because training and ability are not complementary. Before proceeding to the analysis of wages as just described, we therefore provide some evidence on three important issues: that firms actually pay for apprenticeship training; that the training is general; and that training and ability are complementary.

Three studies have been carried out by surveying training firms about their accounting costs and apprentice productivity to assess the net cost of training in Germany. The most careful of these was conducted in 1991 by the Federal Institute for Vocational Training (Bundesinstitut für Berufsbildung) and is described in von Bardeleben, Beicht, and Fehér [1995]. Similar studies exist for other countries. For example, Ryan [1980] examined welder apprentices at one particular U. S. shipyard, and Jones [1986] looked at apprentices in British manufacturing.

All these studies have used basically similar methodologies and find substantial net costs for training apprentices.

The first step is to calculate gross costs as the sum of apprentices' payroll costs; the payroll costs of training personnel; costs of material, equipment, and structures used in the training; and direct costs of any external training that the firm pays for. In addition, the studies assess apprentice output. This is done by surveying supervisors about the jobs done by apprentices, and their productivity relative to skilled workers. A money measure of the output contribution is constructed by multiplying the time spent in productive activity with the payroll costs of a skilled worker and the relative apprentice productivity. This calculation assumes implicitly that the wages of skilled workers are set competitively and therefore reflect marginal products. If the firm has market power over these workers, marginal product may exceed wages, so that this would yield an underestimate of apprentices' output contribution.

A second problem arises from the fact that in many, especially smaller establishments, most trainers are not engaged in training full-time but mostly work in productive activities. The German study for 1991 takes two approaches to this problem. The first is to prorate the time spent on training by part-time personnel. However, this is an average and not a marginal cost of training. The marginal cost of training may be much lower if training takes place during times when the trainers' time is valued less (e.g., during slack times). As an alternative, the study excludes the costs of part-time trainers completely from the cost calculation (they refer to this as variable cost).

Table I illustrates the role of these assumptions using data from von Bardeleben, Beicht, and Fehér [1995] for Germany. Average total gross costs per apprentice per year amounted to almost DM 30,000. Excluding part-time trainers yielded a variable cost of only DM 18,000. Apprentices' productivity, valued at skilled worker wages amounted to about DM 12,000. Under the perfect market assumptions, and using full costs, this yields a net cost of training of around DM 18,000. Using variable costs, the net costs are only around DM 6000. Instead, suppose that skilled worker wages are not set in perfect markets but (arbitrarily) assume that marginal products are twice a skilled workers' wage. Net training costs would then be about DM 6000 using full costs but DM −5000 using only variable costs. The latter number makes very conservative assumptions, and we regard it therefore

TABLE I
COSTS OF APPRENTICESHIP TRAINING IN GERMANY 1991 (GERMAN MARKS PER YEAR)

	All firms	By firm size (number of employees)			
		0–9	10–49	50–499	500+
A) Total gross costs	29,573	27,473	28,176	30,344	35,692
B) Variable gross costs	18,051	13,867	15,074	20,283	28,197
C) Aprentice productivity	11,711	12,221	11,465	12,099	10,311
Perfect markets					
Total net costs (A − C)	17,862	15,252	16,711	18,245	25,381
Variable net cost (B − C)	6,340	1,646	3,609	8,184	17,886
Imperfect markets (50% markdown)					
Total net costs (A − 2 * C)	6,151	3,031	5,246	6,146	15,070
Variable net costs (B − 2 * C)	−5,371	−10,575	−7,856	−3,915	7,575

Source. von Bardeleben, Beicht, and Fehér [1995], Chart 27 and Table 12.

as a lower bound for the net costs of training. However, even this very conservative estimate implies that the largest firms in Germany (those with more than 500 employees) still have positive training costs of around DM 7500. On the other hand, even with the perfect market assumption, the costs for the smallest firms (less than ten employees) using the variable cost concept are close to zero (DM 1500). This has also been observed by Soskice [1994] and Harhoff and Kane [1997]. We will therefore focus on apprentices trained in larger firms in our empirical analysis.

Apprenticeship training is largely general. Firms that train apprentices have to follow a prescribed curriculum, and apprentices take a rigorous outside exam in their trade at the end of the apprenticeship. The industry or crafts chambers certify which firms fulfill the requirements to train apprentices adequately, and firms will not be allowed to train without this approval. Works councils in the firms themselves monitor the training and resolve grievances. At least in certain technical and business occupations, the training curricula limit the firms' choices over the training content fairly severely. For example, a trainer in a large bank told us that apprentices typically do not find the time to learn about the more company-specific products during the apprenticeship and therefore attend additional courses on these topics afterwards.[6] Overall, the institutional arrangements severely limit the firm's ability to structure the apprenticeship training so that it involves mostly firm-specific skills.

The data we analyze provide some additional insight on this. Respondents were asked how much of the content of their apprenticeship they use in their current job. The results (shown at the bottom of Table II) indicate that 64 percent of all workers in our sample use "a lot" or "much" of their apprenticeship skills. By comparing workers who have stayed with their apprenticeship firm with those who work elsewhere, it is possible to assess how much of the apprenticeship skill is firm specific. This comparison is complicated by the fact that firm changes are often associated with changes in the occupation a worker works in as well. If we condition on working in the same occupation, we find that 80 percent of those still with their apprenticeship firm say that they use "a lot" or "much" of their apprenticeship skills, compared with 72 percent of those switching firms. For workers changing occupations the fractions are 46 percent versus 31 percent. These numbers indicates that apprenticeship skills are more useful inside the training firm, but the difference is very small. Most of the skills learned during the apprenticeship constitute general training.

A final important issue for our test is that ability and training have to be complements for adverse selection to explain training investments of the firm. Evidence for adverse selection in the market for trained workers alone is not sufficient to establish that firms pay for the training, as a referee suggested. If training and worker ability do not interact, there could be adverse selection, and firms obtain rents, but workers still pay for all the costs of general training themselves through a reduction in the first period wage. We present some evidence that higher skilled workers receive more workplace training, although these skills are not always unobserved by employers. Altonji and Spletzer [1991] find a strong relationship between training and previous education in the National Longitudinal Survey of the High School Class of 1972. They also find a positive relationship between SAT scores in math and the incidence of training. Bartel and Sicher-

6. Interestingly, the occupations where the training curriculum is more extensive, like industrial mechanic or bank clerk, are also the ones for which von Bardeleben, Beicht, and Fehér [1995] have found higher net costs.

man [1995] also report a positive relationship between training incidence and schooling as well as AFQT scores using the National Longitudinal Survey of Youth. We find similar results in the data for Germany we analyze below. Our data set for 1979 contains grades in math and German in secondary school. In the sample of apprentices described in Section VI, math grades are strongly related to the incidence of job related postapprenticeship training, holding the same covariates constant as in Table III, while grades in German are not. In a more representative sample, both math and German grades as well as previous education are positively associated with training. This consistent finding of an association between higher academic achievement or aptitude on the one hand and training on the other is highly suggestive that ability and training are indeed complements.

V. The Data

We use data from three cross sections of the German "Qualification and Career Survey" (QaC) conducted in 1979, 1985–1986 and in 1991–1992. Each cross section samples approximately 25,000 employed workers of German nationality in the age group 15 to 65. Besides standard questions on demographics and the current job, the surveys contain rich information on job attributes and job content; the qualifications needed for the job and how they were obtained; and retrospective questions on a worker's career path, education, and training history. In particular, the survey contains information on the apprenticeship firm of a worker, and whether and why a respondent separated from that firm.

We restrict the samples we use to workers who finished secondary school after grade 9 or 10, went through an apprenticeship, and did not obtain any higher school-based education. Students who complete twelfth or thirteenth grade in secondary school are supposed to be college bound, and even if they select to do an apprenticeship (which is relatively rare), they are likely to differ from other apprentices and often they attend college after the apprenticeship. Also, to make sure that we look at a population which is likely to have completed both their apprenticeship and their military service, we limit the sample to those between age 23 and 59. We only consider men because women have very different career patterns, and they often train for very different occupations than men. About 99 percent of men with these characteristics work full-time in Germany; we therefore also limit the sample to those working full-time.

We eliminate workers whose apprenticeship was in the public sector or who work currently in the public sector or in the construction industry. Retention rules after the apprenticeship and pay determination in the public sector are likely to be governed by rules different from the private sector. The construction industry is the only sector in Germany where training is not financed voluntarily by firms. Instead, a levy is imposed on those firms that do not train apprentices and redistributed to those who do. This indicates that turnover in the construction industry may be sufficiently high that the mechanism we discuss in this paper does not support training by firms. We would prefer to exclude only workers whose apprenticeship, rather than the current job, was in the construction industry. Unfortunately, the 1985–1986 survey does not contain this information. We further limit the sample to those workers whose apprenticeship was in a firm with 50 employees or more because, as noted above, only large German firms seem to finance apprenticeship training. Furthermore, we restrict the sample to those who left secondary school in 1948 or later. Most military quitters in the earlier period served in World War II, which may introduce survivorship or other biases. Because the samples we select are rather small, we pool the data from all three surveys for the analysis below. The final sample has 5355 observations.

The earnings variable on the surveys is the gross monthly wage. Respondents in the 1979 survey were asked to report their earnings in 13 brackets, in the 1985–1986 survey in 22 brackets, and in 1991–1992 in 15 brackets. We assign each individual earnings equal to the bracket midpoint.[7] We then convert the variable to an hourly wage by dividing by the number of weekly hours. We construct the standard variable for potential experience as age − schooling − 5. The number of years of schooling is derived from the survey information on the types of schools attended and degrees obtained, following Krueger and Pischke [1995].

7. Because of the large number of brackets, this is unlikely to introduce much more measurement error than is done by respondents' rounding continuous amounts. The top bracket in 1979 was DM 5,000 or more to which we assigned a value of DM 7,500, in 1985–1986 it was DM 15,000 or more to which we assigned a value of DM 17,500 and in 1991–1992 it was DM 8,000 or more to which we assigned a value of 12,500. Only 1.6 percent of sample observations are in the top income bracket.

Our key variables of interest are related to the separation from the training firm after the apprenticeship. The survey asks when respondents left their training firm, and the reason for the separation. We only focus on separations immediately after the apprenticeship. Apprenticeship contracts in Germany are generally fixed term, and a new contract has to be written for further employment. Typically, a training firm will make employment offers to most of its apprentices, but it is under no legal obligation to do so. As reasons for a separation, the 1979 survey allowed six responses: left for military service, laid off, could obtain a higher income elsewhere, better working conditions elsewhere, obtained additional training or education, and other. Multiple responses were allowed. In 1985–1986 only three mutually exclusive answers were possible: military conscription, laid off by the firm, or quit voluntarily. In 1991–1992 six mutually exclusive responses were possible to the question: what happened to you immediately after you left your training firm? The responses are immediately obtained a job commensurate with my training, immediately obtained a job but not commensurate with my training, unemployed, completed additional training, military service or voluntary social year, or other. We create two dummy variables from these questions. The variable *stayers* is used for those who stayed with their apprenticeship firm after the end of the apprenticeship, although these workers may have left the firm by the time of the survey. The variable *military quitters* is one for those who left their apprenticeship firm immediately and mention military service as the reason (even if other categories are also mentioned in the 1979 survey).

In order to assess whether our results may be biased due to selection into or experience from military service we also use data from the first wave of the German Socioeconomic Panel (SOEP). This data set has retrospective questions on the employment history of all workers from age sixteen. In particular, we know whether any respondent served in the military at any particular year of his life. Using these retrospective data, we construct a dummy variable for everybody who ever served. We select a sample analogous to the QaC Surveys. However, unlike the samples from the QaC, the SOEP sample includes respondents who might have returned to higher education after the apprenticeship and those who did their apprenticeship in small firms or in the public sector but work in the private sector at the time of the

1984 survey. These exclusions are not possible in the SOEP. The sample includes 513 observations.

VI. Empirical Results

A. Summary Statistics

Table II gives summary statistics for the full sample from the QaC Survey as well as for stayers, military quitters, and other quitters separately. Most apprentices leave secondary school after ninth grade. Military quitters are more likely to have a tenth grade education and are on average three years younger than stayers or other quitters. This is mainly a reflection of a cohort effect because military service has been mandatory for able-bodies German males only since 1957, which means that cohorts born in 1938 or earlier did not serve. Also, the fraction of a cohort drafted has varied over time, and draft rates were lower in the late 1950s and early 1960s. Apart from cohort effects, the higher schooling of military quitters also results from the fact that many teenagers leaving school after grade 9 do not reach draft age (eighteen) when they complete their apprenticeship. 63 percent of grade 9 school leavers are eighteen years or younger when they finish their apprenticeship compared with only 18 percent of those who leave school after grade 10. It is therefore important that we control for this difference in schooling in the regressions below.

Eighty-four percent of apprentices stay in their apprenticeship firm initially, although only 37 percent remain there by the time of the survey. Military quitters make up only 4 percent of the sample—much fewer than the fraction of young men in Germany who actually served in the military during the period covered by the sample. This reflects the fact that many workers stay with their firm after their apprenticeship and only then interrupt their career for military service, or simply return to their previous employer even after military service. Neither of these are counted as military quitters. As discussed in Section IV, this does not invalidate our test. On the contrary, it makes the test more stringent. It is sufficient for our test that the rehiring decision of workers who serve in the military right after their apprenticeship is more stringent than the retention decision for other stayers. In this case, the group of military quitters will be of higher average quality than other quitters and therefore earn higher wages. As explained in the previous section, military quitters may also earn

TABLE II
SUMMARY STATISTICS

Variable	All	Stayers	Other quitters	Military quitters
Schooling:				
9 years	0.68	0.68	0.70	0.61
10 years	0.32	0.32	0.30	0.39
Training as master craftsman	0.07	0.08	0.07	0.04
Uses computer on the job	0.24	0.24	0.24	0.25
Age	37.1	37.1	37.4	34.3
Years of potential experience	20.5	20.6	20.8	17.7
Age at the end of apprenticeship	18.8	18.8	18.9	19.2
Stayers	0.84	1.00	0.00	0.00
Military quitters	0.04	0.00	0.00	1.00
Still in apprenticeship firm	0.37	0.44	0.00	0.00
Number of employers:				
1	0.39	0.44	0.11	0.23
2	0.26	0.25	0.31	0.30
3	0.18	0.16	0.26	0.24
4	0.12	0.10	0.23	0.16
5 or more	0.06	0.05	0.09	0.07
Characteristics of apprenticeship firm				
Sector:				
manufacturing	0.68	0.69	0.63	0.60
crafts	0.16	0.15	0.23	0.20
trade	0.08	0.08	0.08	0.12
services and other	0.08	0.08	0.07	0.08
Firm size:				
50–99	0.24	0.22	0.32	0.36
100–499	0.35	0.35	0.37	0.36
500–999	0.12	0.12	0.12	0.10
1000+	0.29	0.31	0.19	0.18
Characteristics of current job				
Position:				
semiskilled blue-collar	0.06	0.05	0.11	0.05
skilled blue-collar	0.41	0.41	0.37	0.42
master craftsman or supervisor	0.10	0.10	0.10	0.07
low skilled white-collar	0.06	0.06	0.06	0.08
qualified white-collar	0.19	0.19	0.16	0.17
highly qualified/manager	0.12	0.12	0.11	0.15
executive	0.04	0.04	0.04	0.04
self-employed	0.03	0.03	0.06	0.03
Sector:				
manufacturing	0.67	0.68	0.60	0.57
crafts	0.11	0.10	0.14	0.13
trade	0.11	0.10	0.12	0.15
services and other	0.09	0.09	0.09	0.11
Firm size:				
1–9	0.07	0.06	0.12	0.13
10–99	0.24	0.23	0.29	0.36
100–999	0.40	0.41	0.35	0.28
1000+	0.30	0.31	0.25	0.23
Use apprenticeship skill: no	0.10	0.09	0.17	0.10
little	0.09	0.08	0.12	0.07
some	0.18	0.18	0.19	0.23
much	0.24	0.24	0.21	0.24
a lot	0.40	0.41	0.31	0.36
Number of observations	5355	4515	631	209

Data are pooled from the 1979, 1985/86, and 1991/92 German Qualification and Career Surveys. Sample consists of German males, age 23–59, with nine or ten years of schooling, who left secondary school in 1948 or later, completed private sector apprenticeship training without returning to school after the apprenticeship, who did an apprenticeship in a firm with 50 employees or more, were employed in the private sector outside construction, and were working full-time.

higher wages than stayers because they are freed from adverse selection.

What is important for our empirical strategy is that some sample members were separated from their apprenticeship firm because of the draft but would have continued to work there otherwise. A back-of-the-envelope calculation suggests that we would expect to see many fewer than 4 percent military quitters if they were all just regular quitters or laid-off workers. First of all, only 84 percent of the sample were born after 1938, and are therefore draft eligible. Many apprentices who eventually serve do not get drafted immediately when their apprenticeship ends. Among those reporting that they ever leave their apprenticeship firm for military service, only 29 percent leave immediately after the apprenticeship (this information is available in the 1979 and 1985 surveys). If we take this fraction as representative for all workers drafted, then 24 percent ($=0.29 * 0.84$) are at risk of being drafted at the time when service would separate them from their employer at the end of the apprenticeship contract. Since the 1960s about 60 to 75 percent of a cohort have actually served [Ullrich 1984]. Some males may not serve because they are not physically fit. Even among those eligible for conscription, not everybody actually gets drafted. Assuming that on average 70 percent of a cohort actually serve, and if the group at risk to be drafted after the apprenticeship had the same proportion of workers returning to their original employer as the entire sample (84 percent), then we should see only 2.7 percent of military quitters in the sample ($=0.24 * 0.70 * (1 - 0.84)$). The number of actual military quitters is about 50 percent higher than this, indicating that a substantial number of military quitters would have stayed at their training firm if it were not for the draft.

Military quitters, while a small group overall, are about a quarter of all initial separations from the training firm. This discussion suggests that the variation in who becomes a military quitter results from year of birth, the age of the worker at the end of the apprenticeship, and aggregate draft conditions at that time. Figure I shows that the number of military quitters varies over time (i.e., by graduation year from the apprenticeship) in a fashion expected from aggregate draft rates. The fraction of military quitters increases from the mid-1950s and stabilizes around 1970.

A few other characteristics of the sample in Table II are worth noting. A large fraction of apprenticeships is in manufacturing.

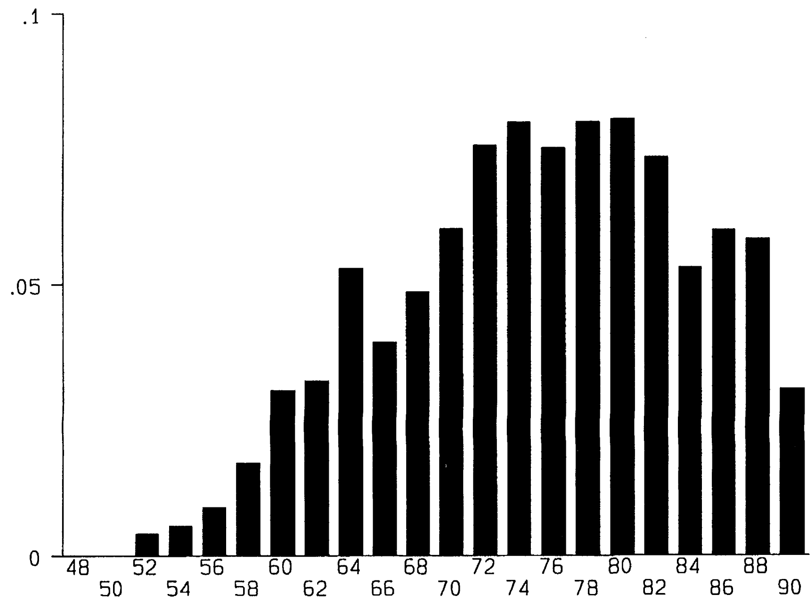

FIGURE I.
Fraction of Military Quitters by Year

This is a reflection of the fact that we restricted the sample to large apprenticeship firms. Military and other quitters are more likely to come out of the crafts sector and from smaller firms because retention rates after the apprenticeship are lower in those firms. The distribution of current job types is rather similar across groups. About 12 percent of the sample work in relatively unskilled blue- or white-collar positions, while 14 percent reach a supervisory blue-collar or executive white-collar position.

B. Basic Results

Stayers and military quitters differ in many observable dimensions, including some that were clearly discernible at the beginning of the apprenticeship. Since there is likely to be a good deal of selection into better and worse apprenticeships among school leavers (Soskice [1994] argues that this is the case), it is important to control for these observable attributes in our wage regressions. On the other hand, we do not want to control for any attributes of the jobs workers obtained after their apprenticeship since these will be related to worker ability and selection, which is part of the outcome we are interested in. We therefore present wage regressions in Table III which include a dummy for leaving school after tenth grade, a quartic in potential experience, dummies for the sector of the apprenticeship firm, and for the size of the apprenticeship firm.

Our main result is that both stayers and military quitters earn higher wages than those who left their apprenticeship firm for other reasons, but the coefficients are not estimated very precisely. The stayer coefficient is not significantly different from zero, and the two coefficients are not significantly different from each other. However, the point estimates imply that military quitters actually earn more than stayers, which is consistent with our model. Recall that according to our model military quitters have higher ability than other quitters but *lower ability* than stayers. However, once they separate, military quitters can obtain their marginal product while the incumbent firm extracts rents from the stayers. If these rents are large enough, they may outweigh the quality advantage of the stayer group. The test of whether military quitters earn wages higher than or wages as high as stayers is therefore a challenging test of our model and receives some support.

An alternative interpretation of our results would be that the military positively selects draftees or that military service imparts skills or attitudes which are valued in the civilian labor market. We find this unlikely to be the case of skilled workers in Germany. The military conducts a physical exam but no other tests to determine draft eligibility. This limits the ability of the military to select draftees on the basis of skill. We are not aware of any empirical studies for Germany on the impact of military service on civilian earnings. Imbens and van der Klaauw [1995] find that mandatory military service in the Netherlands, which has a very similar draft system, lowers earnings by about 5 percent for those who served, which implies that veterans lose approximately the returns to experience during the time of service. This suggests that there is little evidence of military service raising earnings in civilian jobs, for example through training.

To provide more direct evidence on this question, we ran a wage regression for a sample of men from the 1984 wave of SOEP, which we constructed to be as similar as possible to the QaC sample. The SOEP allows us to include a dummy variable for

TABLE III
BASIC WAGE REGRESSIONS (DEPENDENT VARIABLE: LOG AVERAGE HOURLY EARNINGS)

Independent variable	Qualification and career survey (1)	(2)	1984 SOEP (3)
Attended 10th grade	0.160	0.162	0.153
	(0.011)	(0.007)	(0.028)
Experience	0.123	0.108	0.028
	(0.024)	(0.017)	(0.098)
Experience^2/100	−0.694	−0.532	−0.086
	(0.182)	(0.126)	(0.848)
Experience^3/10,000	1.840	1.229	0.161
	(0.558)	(0.379)	(3.007)
Experience^4/1,000,000	−1.806	−1.063	−0.123
	(0.597)	(0.399)	(3.740)
Apprenticeship in manufacturing	0.024	—	—
	(0.013)		
Apprenticeship in trade	0.036	—	—
	(0.022)		
Apprenticeship in other sector	0.041	—	—
	(0.021)		
Apprenticeship firm had 100–499 employees	0.045	—	—
	(0.013)		
Apprenticeship firm had 500–999 employees	0.072	—	—
	(0.016)		
Apprenticeship firm had 1000+ employees	0.095	—	—
	(0.014)		
Stayer	0.012	0.027	—
	(0.015)	(0.008)	
Military quitter	0.045	0.011	—
	(0.025)	(0.014)	
Ever did military service	—	—	−0.022
			(0.024)
R^2	0.384	0.337	0.126

White standard errors are in parentheses. Samples in the first two columns are pooled from the 1979, 1985/86, and 1991/92 German Qualification and Career Surveys and consist of German males, age 23–59, with nine or ten years of schooling, who left secondary school in 1948 or later, completed private sector apprenticeship training without returning to school after the apprenticeship, were employed in the private sector outside construction, and were working full-time. Column (1) includes workers who did an apprenticeship in a firm with 50 employees or more; column (2) uses apprentices from firms of all sizes. Number of observations is 5,355 in column (1) and 13,051 in column (2). "Stayers" are those workers who continued in their apprenticeship firm after training; "military quitters" are those who left their training firm for military service. Sample in the last column is from the Socioeconomic Panel and consists of German males, age 23–59, with nine or ten years of schooling, who left secondary school in 1948 or later, were employed in the private sector outside construction, and were working full-time. Number of observations is 513. All regressions also include a constant, and the regressions in columns (1) and (2) include two additional dummies for the survey year.

everybody who ever served in the military.[8] If draftees are positively selected or if military experience is valuable in the civilian labor market, then military service should be associated with a positive wage differential. Instead, we find in column (3) that military service is actually associated with slightly lower earnings, although this effect is also not significant. This indicates that our findings for military quitters are not due to selection or training in the military, thus leaving our explanation as a more likely candidate. In order to make these results more closely comparable to the QaC results, we reran the regression for that data set with a specification and sample that matches the SOEP more closely. The stayer coefficient is slightly higher now because controls for the size of the apprenticeship firm are excluded. Apprentices in larger firms are more likely to stay with their training firm and earn higher wages. The military quitter coefficient is lower in column (2), but the point estimate is still positive. This is because the sample in column (2) includes workers who did their apprenticeship in smaller firms, a sector where adverse selection is less likely to be important (see below).[9]

C. Robustness Checks

We investigate the robustness of our results next and extend our findings in some other directions. Table IV reports the coefficients on the stayer and military quitter dummies for a number of different samples and specifications, making modifications to the baseline one at a time. First, we add a dummy variable indicating whether a worker is still with his apprenticeship firm. If a large enough part of apprenticeship skills is firm-specific and general and specific training are complements, the training firm may be able to recoup its investment even in the general human capital component of the apprenticeship [Acemoglu and Pischke 1996b]. We include a dummy for everybody still with their apprenticeship firm in the wage regression. In a specific human capital model, the coefficient on this variable should be positive. On the other hand, the model says nothing about the earnings of

8. This data set does not allow us to construct some of the other controls about the type of apprenticeship because such data are not collected in the survey.

9. This test does not necessarily show that training and selection are unimportant for the group of military quitters. There might be a small group of draftees who learn valuable skills in the military but this hardly affects average returns from service. These skills might differ from the individual's previous skill set so that the draftees who have learned a lot are most likely to become "military quitters" in order to capitalize on these skills in a new job.

TABLE IV
SPECIFICATION CHECKS (DEPENDENT VARIABLE: LOG AVERAGE HOURLY EARNINGS)

Specification or sample	Stayers	Military quitters	Sample size
Basic regression	0.012	0.045	5355
	(0.015)	(0.025)	
Add control for still working at the apprenticeship firm	0.026	0.045	5355
	(0.016)	(0.025)	
Limit the sample to workers 35 years or younger	0.033	0.051	2478
	(0.019)	(0.032)	
Allow separate coefficients on all other covariates by survey year	0.011	0.042	5355
	(0.015)	(0.025)	
Exclude self-employed	−0.003	0.035	5198
	(0.014)	(0.025)	
Include public sector and construction industry	0.020	0.025	6343
	(0.013)	(0.025)	
Include those with 12 or 13 years of school	0.007	0.028	5792
	(0.014)	(0.024)	
Include those who left school before 1948	0.008	0.029	5946
	(0.015)	(0.024)	
Only include those born after 1938 and include full set of experience dummies	0.013	0.042	4499
	(0.016)	(0.026)	
Respondents with apprenticeship in firms with less than 50 employees	−0.003	−0.017	7696
	(0.010)	(0.017)	
Respondents with apprenticeship in firms with more than 100 employees	0.010	0.029	4085
	(0.018)	(0.031)	

White standard errors are in parentheses. Data are pooled from the 1979, 1985/86, and 1991/92 German Qualification and Career Surveys. Samples for the base regression consist of German males, age 23–59, with nine or ten years of schooling, who left secondary school in 1948 or later, completed private sector apprenticeship training without returning to school after the apprenticeship, who did an apprenticeship in a firm with 50 employees or more, employed in the private sector outside construction, and were working full-time; changes to the base sample are as specified in the table. "Stayers" are those workers who continued in their apprenticeship firm after training; "military quitters" are those who left their training firm for military service. Except where otherwise noted, additional controls are a dummy for tenth grade education, a quartic in potential experience, three dummies for the sector of the apprenticeship firm, three dummies for the size of the apprenticeship firm, a constant, and two dummies for the survey year.

leavers depending on the timing of the separation. Thus, we would expect zero coefficients on the stayer and military quitter dummies. The empirical results show exactly the opposite pattern. The positive wage effect for stayers actually becomes stronger now. The coefficient on the variable for still being in the apprenticeship firm at the time of the survey is negative (and significant). It may therefore reflect matching considerations; only workers with better outside offers leave their jobs.

The market power of the incumbent employer may wear off with the labor market experience of the worker, or relevant abilities of workers may change with age (see subsection III.D). The next row in Table IV limits the sample to worker 35 years of age or younger. We might expect that the effects of adverse selection are stronger in this group, and this is in fact the case in our sample. Next, we find that the coefficients for both stayers and military quitters change little if we allow separate coefficients on all other regressors by survey year. If we eliminate the self-employed from the sample, the coefficients on both stayers and military quitters are lower now. The base sample excludes workers in the public sector and in construction. There are good reasons to exclude workers who do their apprenticeships in these sectors, but not necessarily workers who choose employment in these sectors later. Including these industries lowers the military quitter coefficient but does not change the basic pattern. We also investigate the effect of including those with twelve or thirteen years of education which again lowers both the stayer and military quitter effects. We find similar results in the next row when we look at respondents who finished school before 1948. This adds workers who completed their apprenticeship during the 1930s and 1940s, increasing the diversity of military quitters by including World War II veterans.

Our baseline sample uses both variation in the military draft rates during the draft and the fact that the draft only started in 1957 to help identify the military quitter dummy. To ensure that our military quitter dummy does not pick up any nonlinearities in age or cohort not already controlled for, we limit the sample to those who were born after 1938. These are the cohorts at risk of being drafted for military service. In addition, instead of controlling for a quartic in experience, we put in separate dummies for each year of experience of the respondents in the sample. The coefficients are hardly changed from the baseline specification.

The most dramatic change occurs when we limit the sample to apprenticeships in small firms. This is the sector where training may be paid for by apprentices directly and adverse selection plays less of a role because more apprentices leave their training firm. Both coefficients are negative now. This suggests that, as it could have been expected, adverse selection may not be a major concern in the market for apprentices who leave small firms. These firms, which train only few apprentices, may be much more constrained by business needs on when they can keep a worker after the apprenticeship. The difference in the results for those getting their training in small firms also makes us more confident that our earlier results on military quitters did not simply reflect selection or training in the military. However, the relationship between the wage differentials for stayers and military quitters and firm size is not simple. When we restrict the sample to large firms with more than 100 employees, both coefficients are attenuated compared with the baseline model.

In summary, the results in Tables III and IV are consistent with the predictions of our model. Workers who stay with their firm after the apprenticeship earn higher wages than those who quit or are laid off. Military quitters also earn higher wages than other quitters because they are separated from their firm for an exogenous reason and should therefore be of higher quality than other quitters. In fact, their earnings appear to be higher than those of stayers in all of our regressions. This result is surprising at first, but consistent with our theory as explained above.

It is important to point out that these results are not easily reconciled with alternative explanations for training. One objection might be that there are ability differences between stayers, quitters, and military quitters because of selection, but these differences are observable to everyone. Yet, this explanation suggests that military quitters should earn strictly less than stayers who are the ablest group. This is inconsistent with our point estimates, although we cannot reject the hypothesis that military quitters earn less than stayers.

VII. Discussion and Conclusion

This paper has proposed a new theory of training and presented some empirical evidence from Germany in support of the theory. We believe that this model has some attractive features relative to the standard Beckerian theory as it predicts that firms should pay for the general training of their employees, which according to many is standard practice in the labor market. We also derived a multiplicity of equilibria from the interaction of quit decisions and the average quality of the secondhand worker pool.

Our theory is partly motivated by the institutional structure of the German labor market where firm-sponsored general training is very common. In Germany firms voluntarily offer apprenticeships to their workers. The skills provided in these programs are highly general, but firms bear a considerable fraction of the costs of training. This situation contrasts with the U. S. labor market where the incidence of general training is more limited. To conclude, we will draw some implications from our model regarding these different outcomes in the two countries.

1. For the U. S. labor market, Topel and Ward [1992] report that the median number of jobs that a worker with ten years of experience holds is six. In contrast, in our data we calculate this number to be one or two in Germany. Our model is capable of generating such different equilibrium quit rates in two different ways. First, in terms of our model, we can think of the United States at a low training and high quit equilibrium and Germany at a low quit, high training equilibrium. Second, there may be some small differences in the underlying "quit-preferences" (i.e., the function $G(\theta)$ in our model) which will be amplified in equilibrium. In either case, the informal claim made by Blinder and Krueger [1996] that turnover should be negatively related to the level of training is formalized.
2. If Germany and the United States can be thought of as in different equilibria, some other interesting implications also follow. Our discussion established that although the high training equilibrium is better because the investment margin is less distorted, it has a worse allocation of workers to jobs; that is, workers end up staying in jobs for which they have high disutility. Instead, if the disutility θ can also be thought of as a measure of "match-quality" between a firm and a worker, then the economy with high turnover should achieve a higher productivity due to better matches but lower productivity due to a lack of investment in general skills. Interestingly, Topel and Ward [1992] find that workers who change jobs in the U. S. labor market get a positive wage premium which is consistent with a better match in the firms to which they move. In contrast, we find that workers who change jobs right after the end of their apprenticeship in Germany get a wage reduction, unless they are drafted for the military.
3. Our model also implies that when raids are possible, the multiplicity of equilibria will disappear, and the economy will have a unique equilibrium with low training. The institutional differences, especially the role of works councils in Germany, make raids on workers in employment with other firms much harder (see, for example, Abraham and Houseman [1993]). This suggests that a different interpretation is that Germany is in the equilibrium with

high training because raids are not allowed, and the United States is in the unique low training equilibrium of our economy where firms can poach workers from each other.

4. Information regarding the ability of the worker may be more important in Germany than in the United States because of high firing costs at the later stages of a worker's career (i.e., after the apprenticeship). Since a firm can easily lay off its workers in the United States, the cost of hiring a low quality worker is not enormous, and thus adverse selection may be less of a problem. This observation also helps in thinking about the institutional differences between the labor markets of the United States and Germany in endogenous terms. High firing costs in Germany combined with the training system and the low quit equilibrium may not be too distortionary, while the same level of firing costs in the United States without any other changes in the labor market may be more damaging. This may also help explain the good performance of the German economy during the postwar period, despite the fact that the high degree of labor market regulation might have hampered the optimal allocation of workers across firms.
5. Given the lower ability of firms to exploit internal training opportunities, our model also suggests that school-based training financed by workers themselves should be more important in the United States which appears to be the case in practice.
6. Adverse selection in the labor market can also explain firm-sponsored training at all stages of a worker's career. Pischke [1996] reports evidence that further training in Germany does not lead to higher wages. This is true even though this training seems to be general: many workers reported receiving a written certificate that they would present when applying for a new job. Workers who received further training also report that it enables them to do their jobs better, so the training is productive. If similar levels of further training are given to workers of different abilities, adverse selection will prevent workers from capitalizing on their additional skills in the market. This may explain why the incidence of further firm-sponsored training should be higher in Germany than in the United States.

7. Any explanation for general training financed by firms will have to rely on some market imperfection. If labor markets are imperfect, comparisons of the U. S. and German training systems, which rely on wages to make inferences about productivity, might be misleading. For instance, Harhoff and Kane [1997] and, borrowing from them, Heckman [1993] point out that there may be no market failures preventing training in the United States since wages grow at the same rate over the life-cycle in the two countries, which they interpret as the same rate of human capital accumulation. In the model of this paper compensation is front loaded, and front loading is more extreme in equilibria with more monopsony power and thus with more training. Therefore, similar wage profiles in the United States and Germany are consistent with more human capital accumulation in Germany. In fact, our model implies that if two economies with different levels of quits have the same growth rate of wages, the one with higher quits must have slower accumulation of skills!

APPENDIX: PROOFS OF PROPOSITIONS AND LEMMATA

PROPOSITION 2. (i) See Acemoglu and Pischke [1996a]. (ii) Now we establish some properties of the equilibrium. Suppose that $\tau^* = 0$ is a solution. From (5) this implies that either (i) $q[w^*(\tau^*),v(\tau^*)] = 1$ or (ii) $\int_0^{\hat{\eta}^*(0)} (\alpha'(0)\eta - v'(0))\, dF(\eta) = 0$. But $q[w^*(\tau^*),v(\tau^*)] = 1$ is not consistent with firm optimization; i.e., (4). Thus, (i) cannot be the case. Next consider (ii). Recall that from Assumption 1, $\alpha'(0) = \infty$. Thus, from (7), $v'(0) = [(q\bar{\eta} + (1-q)\int_0^{\hat{\eta}^*(0)} \eta\, dF(\eta))/(q + (1-q)F(\hat{\eta}^*))]\alpha'(0)$ (i.e., all other terms in $v'(\tau)$ are finite, and as $\tau \to 0$, they can be ignored). This implies that for all $\hat{\eta}^*(0) > 0$, $v'(0) < \alpha'(0)$, and $\hat{\eta}^* > 0$ follows from $w^*(0) > 0$. Therefore, $\tau^* = 0$ cannot be a solution to (5)—a contradiction. Thus, $\tau^* > 0$.

Next suppose that $\tau^* = \tau^c$. Then $q[w^*(\tau^*),v(\tau^*)] = 0$, but $q = 0$ is only possible from (5) if $\hat{\eta}^* = 0$, which implies from (4) that $w^* = 0$. But then, all $\theta > 0$ would leave and obtain $v \geq 0$. Thus $q > 0$—a contradiction. Hence $\tau^* < \tau^c$.

QED

PROPOSITION 3. (i) The argument in the text establishes that the equilibrium quits and wages are given by the intersection of two upward-sloping schedules for a given τ. Both are continuous by Assumption 2 and Proposition 2. Since $g(.)$ is unrestricted, we can always obtain multiple equilibria. Once we have multiple equilibria for a given training level τ, then from (5) we can get the equilibrium levels of τ.

(ii) First, we show that $v^a > v^b$ is equivalent to $q^a > q^b$. Consider the first-order conditions of the maximization problem. Differentiating the definition of q, we can see that

$$\frac{dq}{dv} = \frac{\partial q}{\partial v} + \frac{\partial q}{\partial w^*}\frac{dw^*}{dv},$$

where ∂v denotes a small change in v at given τ. Next note that $\partial q/\partial v = g(w(\tau) - v(\tau))$ and $\partial q/\partial w^* = -g(w(\tau) - v(\tau))$. $\partial q/\partial w^*$ can be obtained by totally differentiating (4). This gives

$$\frac{dw^*}{dv} = \frac{\lambda g(.)(1 - F(\hat{\eta})) - \lambda g'(.)\int_{\hat{\eta}}[\alpha\eta - w]\,dF(\eta)}{2\lambda g(.)(1 - F(\hat{\eta})) - \lambda g'(.)\int_{\hat{\eta}}[\alpha\eta - w]\,dF(\eta) + (1/\alpha)[1 - \lambda(1 - G(.)) + \lambda g(.)]f(\hat{\eta})},$$

where $g(.)$ stands for $g(w - v)$ and similarly for $G(.)$ and $g'(.)$. Assumptions 2 and 3 ensure that both the numerator and the denominator are positive but the denominator is larger. Thus, $0 < dw^*/dv < 1$, which implies that $dq/dv > 0$ everywhere. Equivalence of $q^a > q^b$ and $\tau^a < \tau^b$ follows from (4).

PROPOSITION 4. Take a sequence of distribution functions G_n that converge to G_∞ in the sup norm, where G_∞ is dirac at 0. Let W_n be the sequence of training wages corresponding to the equilibrium with distribution function G_n. Then it is immediate that $W_\infty > 0$, since when G_∞ is degenerate $w^* = \hat{\eta} = v = q = 0$. Then $\exists n^* : \forall n > n^*, W_n > 0$.

QED

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

REFERENCES

Abe, Yukiko, "Specific Capital, Adverse Selection and Turnover: A Comparison of the United States and Japan," *Journal of the Japanese and International Economies,* VIII (1994), 272–294.

Abraham, Katharine G., and Susan N. Houseman, *Job Security in America: Lessons from Germany* (Washington, DC: Brookings Institution, 1993).
Acemoglu, Daron, "Training and Innovation in an Imperfect Labor Market," *Review of Economic Studies,* LXIV (1997), 445–464.
Acemoglu, Daron, and Jörn-Steffen Pischke, "Why Do Firms Train? Theory and Evidence," NBER Working Paper No. 5605, 1996a.
Acemoglu, Daron, and Jörn-Steffen Pischke, "The Structure of Wages and Investment in General Training," mimeo, MIT, 1996b.
Altonji, Joseph G., and James R. Spletzer, "Worker Characteristics, Job Characteristics, and the Receipt of on-the-job Training," *Industrial and Labor Relations Review,* XLV (1991), 58–79.
Bartel, Ann P., and Nachum Sicherman, "Technological Change and the Skill Acquisition of Young Workers," NBER Working Paper No. 5107, 1995.
Becker, Gary, *Human Capital* (Chicago: University of Chicago Press, 1964).
Blinder, Alan S., and Alan B. Krueger, "Labor Turnover in the USA and Japan: A Tale of Two Countries," *Pacific Economic Review,* I (1996), 27–57.
Burdett, Kenneth, "A Useful Restriction on the Offer Distribution in Job Search Models," in Gunnar Eliasson, Bertil Holmlund, and Frank P. Stafford, eds., *Studies in Labor Market Behavior: Sweden and the United States* (Stockholm: IUI Conference Report, 1979).
Chang, Chun, and Yijiang Wang, "A Framework for Understanding Differences in Labor Turnover and Human Capital Investments," *Journal of Economic Behavior and Organization,* XXVIII (1995), 91–105.
Chang, Chun, and Yijiang Wang, "Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revisited," *Journal of Labor Economics,* XVI (1996), 505–519.
Cho, In-Koo, and David M. Kreps, "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics,* CII (1987), 179–222.
Gibbons, Robert, and Lawrence F. Katz, "Layoffs and Lemons," *Journal of Labor Economics,* IX (1991), 351–380.
Glaeser, Edward, "Two Essays on Information and Labor Markets," University of Chicago, Ph.D. thesis, 1992.
Greenwald, Bruce C., "Adverse Selection in the Labor Market," *Review of Economic Studies,* LIII (1986), 325–347.
Harhoff, Dietmar, and Thomas J. Kane, "Is the German Apprenticeship System a Panacea for the U. S. Labor Market?" *Journal of Population Economics,* X (1997), 171–196.
Heckman, James, "Assessing Clinton's Program on Job Training, Workfare, and Education in the Workplace," NBER Working Paper No. 4428, 1993.
Imbens, Guido, and Wilbert van der Klaauw, "Evaluating the Cost of Conscription in the Netherlands," *Journal of Business and Economic Statistics,* XIII (1995), 207–215.
Jones, Ian S., "Apprentice Training Costs in British Manufacturing Establishments: Some New Evidence," *British Journal of Industrial Relations,* XXIV (1986), 333–362.
Katz, Eliakim, and Adrian Ziderman, "Investment in General Training: The Role of Information and Labour Mobility," *Economic Journal,* C (1990), 1147–1158.
Krueger, Alan B., and Jörn-Steffen Pischke, "A Comparative Analysis of East and West German Labor Markets: Before and After Unification," in Richard B. Freeman and Lawrence F. Katz, eds., *Differences and Changes in Wage Structures,* NBER Series in Comparative Labor Markets (Chicago: University of Chicago Press, 1995), pp. 405–445.
Lazear, Edward P., "Raids and Offer Matching," in Ronald Ehrenberg, ed., *Research in Labor Economics,* VIII (1986), 141–165.
Mincer, Jacob, *Schooling, Experience, and Earnings* (New York: Columbia University Press, 1974).
Prendergast, Canice, "Career Development and Specific Human Capital Collection," *Journal of the Japanese and International Economies,* VI (1992), 207–227.
Pischke, Jörn-Steffen, "Continuous Training in Germany," NBER Working Paper No. 5829, 1996.
Ryan, Paul, "The Cost of Job Training for a Transferable Skill," *British Journal of Industrial Relations,* XIII (1980), 334–352.

Soskice, David, "Reconciling Markets and Institutions: The German Apprenticeship System," in Lisa M. Lynch, ed., *Training and the Private Sector International Comparisons,* NBER Series in Comparative Labor Markets (Chicago: University of Chicago Press, 1994), pp. 25–60.
Topel, Robert H., and Michael P. Ward, "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics,* CVII (1992), 439–480.
Ullrich, A. F., *Die Bundeswehr. Eine Diskussionsgrundlage* (Dülmen, Germany: Laumann Verlag, 1984).
von Bardeleben, Richard, Ursula Beicht, and Kálmán Fehér, *Betriebliche Kosten und Nutzen der Ausbildung. Repräsentative Ergebnisse aus Indstrie, Handel und Handwerk,* Berichte zur beruflichen Bildung Heft 187 (Bielefeld, Germany: Bertelsmann, 1995).
Waldman, Michael, "Job Assignments, Signaling, and Efficiency," *Rand Journal of Economics,* XV (1984), 255–267.
——, "Up-or-Out Contracts: A Signaling Perspective," *Journal of Labor Economics,* VIII (1990), 230–250.