Wage Differentiation via Subsidised General Training *

V. Bhaskar
University of Essex†

Steinar Holden
University of Oslo‡

June 2004

Abstract

We provide a new explanation for why firms pay for general training in a competitive labor market. If firms are unable to tailor individual wages to ability, for informational or institutional reasons, they will pay for general training in order to attract better quality workers. The market provision of training may well exceed the first best level. Our explanation relies on wage compression *within* skill categories, while imperfect competition based explanations rely on wage compression *across* skill categories.

JEL Categories: J31, D82.

*Thanks to Alison Booth, Espen Moen and seminar participants at Essex for useful comments. V. Bhaskar thanks the Economic and Social Research Council, UK for its support under research grant L138251029.

†Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK. Email:vbhas@essex.ac.uk.

‡Department of Economics, University of Oslo, Box 1095 Blindern, 0317 Oslo, Norway. Email: steinar.holden@econ.uio.no.

1 Introduction

In his seminal work, Becker [6] argued that with a competitive labor market, individual workers would be paid their marginal product, and would therefore be able to capture the entire returns to general training in the form of additional wages. There is however a range of empirical evidence which finds that firms do seem to provide and pay for general training. The leading explanation for this phenomenon is that general training induces a larger increase in productivity than in the wage. Following Pigou, a number of authors such as Stevens [27] and Acemoglu and Pischke [1] argue that this is possible with an imperfectly competitive labor market.[1] As noted by Acemoglu and Pischke, if the distribution of wages is compressed relative to the distribution of marginal products arising from general training[2], then the worker's return to training is less than the rise in her marginal product. In this case the worker has insufficient incentives to engage in general training, while the firm does have an incentive to provide general training, as the wage increases less than the productivity.

In this paper we provide a novel explanation for the employer provision of general training. We argue that employers are sometimes unable to differentiate wages *within* skill categories according to the ability of the worker. This inability can arise due to a variety of reasons, as imperfect information about ability, institutional factors like trade unions, or practical concerns as to the costs of individual wage bargaining. Given undifferentiated wages, employers clearly have an incentive to use other means of attracting the more able workers. If able workers have lower subjective costs of undertaking training (as in Spence's analysis of job market signaling), the employer provision of general training is a way for employers to partially tailor wages according to ability. By providing free training, employers are able to induce a positive

[1] Other explanations include that general and specific skills are complementary (see discussion and references in Stevens ([28]), and Kessler and Lulfesmann, [19]), asymmetric information (Katz and Ziderman ([18]) and Acemoglu and Pischke ([1])) and matching frictions (Burdett and Smith, [13])

[2] Bhaskar and To [8] show that under plausible models of imperfect competition, wage differentials can be compressed or magnified relative to skill differentials, so that workers may have either insufficient or excessive incentives to acquire general training. Dustmann and Schönberg [14] argue that employer market power is insufficient to explain training incentives in Germany. They argue that union mandated wage floors interact with employer market power in providing training incentives.

selection of more able workers.

We show that this sort of employer financed general training may arise even in a competitive labor market. In a symmetric model, all employers will be forced to offer subsidized general training, since otherwise they will be left only with the pool of lower ability workers. If the costs of training are not too high relative to the productivity differential, training will be fully subsidized by the firms. The subsidized training will induce excessive training relative to the social optimum, even if training involves a subsequent wage premium which is equal to the increase in productivity caused by training. A further point is that when firms offer subsidized training to select the more able workers, training will work as a signal of higher productivity. In this case the wage differential for trained workers is likely to exceed the increase in productivity caused by training. Yet our earlier results prevail: in equilibrium firms will offer subsidized training, in sharp contrast to the requirement of Acemoglu and Pischke [2], where a crucial condition is that the wage differential is smaller than the productivity increase caused by training.

We argue that in several noteworthy cases, our model provides a better explanation of existing evidence than do previous theories. Acemoglu and Pischke [2] use German firms' provision of firm-sponsored general training in the form of apprenticeships as a leading example of their theory, based on the claim that wages are more compressed across skills in Germany than in the US.. However, the evidence in Freeman and Schettkat [8] show that the wage increase from unskilled to skill level II (some college/ hauptschule + apprenticeship/meister) is in fact greater in Germany (Table 2), casting some doubt on this claim. In contrast, the residual variance in a standard wage equation is greater in the US than in Germany, suggesting that there is more wage compression among workers of the same observational skill level in Germany than in the US. Within our model this would suggests more firm-provided general training in Germany than the US.

Autor[5] explains the finding that temporary help supply (THS) establishments provide free training in portable computer skills by the idea that free training induces self-selection of high ability workers, as in our model. In Autor's model, firms profit from doing this by obtaining an informational advantage about the ability of the workers, which then can be exploited by offering wages below marginal productivity. However, this justification begs the question of why non-training firms cannot obtain the same information by undertaking a similar assessment of computer skills as the training firms do, at low cost. Our model offers an alternative interpretation of why THS

establishments provide free general training: Because THS firms initially are unable to observe worker ability, and thus unable to differentiate the wage accordingly, free training is offered to attract the high ability workers (assuming that training is more valuable or less costly to high ability workers). As shown in our model, this may constitute a sufficient motivation for firms to provide general training, and no ex-post informational advantage of the incumbent firm is necessary.

The remaining of the paper is organized as follows. Section 2 sets out the basic model of labor market competition when worker ability is unobserved by firms, and shows that firms will pay for training in equilibrium. To sharpen the focus on the key idea, this model is made simple and unrealistic. In section 3, we extend the analysis in several ways, by allowing for a second time period where the wage depends on the training in period 1, and various assumptions on employers' information as to workers' ability. We show that our basic results continue to be true. Section 4 discusses the empirical evidence and section 5 concludes.

2 The Model

We assume that there is a continuum of workers. A worker's productivity (or ability) equals θ, where the support of θ is the interval $[a, b]$, with a strictly positive density f. We assume that the worker knows her own ability, but that this is not observable by the firm at the time of the hiring decision. The worker has a reservation wage r; to abstract from problems of adverse selection, we assume that r is fixed and independent of θ. We also assume $r < b$, since otherwise there can never be any employment.

A firm has constant returns production technology, and the value of its output is the sum of productivities of the individual workers. We assume that the labor market is perfectly competitive. We model this by assuming that there are at least two firms, who are engaged in Bertrand style competition for workers.

Our key assumption is that the worker's wage cannot be conditioned on her ability. In contrast, a training program which is universally offered but induces self-selection does not require such conditioning. The inability to tailor wages can arise due to a number of reasons:

1. The simplest case is when ability cannot be observed by the firm, as we have set out above.

2. Individual wage setting may involve considerable costs both with negotiating wages and with possible costs arising from a dispute in the wage setting.

3. Considerations of fairness may also constrain inequality in wage setting. If θ is a subjective measure of performance, a worker may not agree with the firm's evaluation of her productivity, and may resent being paid lower wages than other workers whom she perceives as equally or less able. The importance of such considerations are borne out by equity theory (Adams, [3]), and also by the evidence on inequity aversion (Fehr and Schmidt,[15], Loewenstein, Thompson, and Bazerman [21] or Bolton and Ockenfels, [10]. Bewley's [6] survey shows that employers are wary of adopting internal wage policies which have deleterious effects on worker morale. Marsden, French and Kubo [24] analyze evidence from a survey on performance-related pay in the British public sector services. They find widespread de-motivating effects arising from difficulties of measuring and evaluating performance fairly. In contrast, differential benefits from a training program are due to self-selection by workers, and are consequently unlikely to have negative morale effects.

4. Considerations of fairness may interact with moral hazard considerations and make wage differentiation difficult. While the firm gradually

   will observe productivity over time, productivity may not be verifiable, and thus difficult to contract upon as the firm then would have a moral hazard problem in the form of an incentive to claim that productivity was low, in order to reduce its wage payment. As has pointed out by Malcomson [22] and others, such moral hazard may be overcome by tournament schemes where the total wage bill is fixed. However, these schemes require that any worker should be able to verify the wages paid to other workers, and thus a worker who is paid less will suffer a loss of status which makes such low pay doubly unpalatable. While moral hazard problems may exist also in the provision of training[3], it seems less likely to be important here than in the setting of individual pay.

5. Finally, there may be institutional constraints on differentiating wages among workers of equal skill levels. Unions are often opposed to discre-

[3]Malcomson et. al. [23] provide a model of apprentice system where such moral hazard plays an important role.

tion in individual wage setting, and this may prevent the firm tailoring wages to reflect differences in ability.[4] One reason for union opposition is the subjectivity of performance evaluation. This makes unions fear that employers can use discretion as a way of punishing workers who are actively involved in the union. International comparisons show a clear positive link between high unionization, centralized wage setting and wage compression, (see Blau and Kahn [9] and Wallerstein [29]).

Note that we do not argue that these reasons can never be circumvented; obviously, we do observe a lot of wage differentiation. Our claim is more modest, that wage differentiation in many cases and for various reasons will be smaller than ability differences. We then consider the extreme case, where there is no wage differentiation.

As a benchmark, consider first the case where firms cannot offer training. Let $\bar{\theta}$ denote the mean of θ. In market equilibrium, each firm offers a wage equal to $\bar{\theta}$. If $\bar{\theta} \geq r$, all workers will be employed, even those whose productivity is lower than r. If $\bar{\theta} < r$, no worker will be employed. In either event, firms make zero profits, reflecting the competitiveness of the market.

Now suppose that firms can offer training. To sharpen the focus on training as a means of attracting high-ability workers, we assume in this section that training does not affect a worker's productivity. The monetary cost of training a worker is k. In addition, training entails a non-monetary cost or disutility to the worker, and is denoted by $c(\theta)$. Following Spence [26], we assume that the disutility of training is lower for high ability workers, so that c is strictly decreasing and differentiable. If the worker trains, then she gets a return R which is exogenous for the moment. That is, R does not depend upon the ability of the individual worker or on the average ability of the trained worker. Then the payoff for the worker from choosing to train is

$$\pi(\theta) = -c(\theta) + R \tag{1}$$

If workers pay the full costs of training, only high ability workers for whom $\pi(\theta) \geq 0$ will choose to train, implying a cut off value θ^*given by $\pi(\theta^*) = k$, i.e. the payoff to the worker is equal to the monetary cost of

[4]In Norway, the teachers' union has several times publically opposed individual wage setting. Interviews with personnel managers of unionized firms in India confirm that unions are opposed to managerial discretion in paying bonuses.

training (see Figure 1). If R coincides with the social return to training, as we shall assume in this section, θ^* is the socially optimal level of training.

We shall assume that each firm offers a contract of the type (w, λ), where w denotes the wage being paid, and $\lambda \in [0, 1]$ is the fraction of the training costs borne by the worker. For example, the firm may pay the fees in a college education, but let the worker cover all the other costs. For the time being, we assume that the firm cannot condition the wage payment on the training decision. A possible justification for this assumption is that if the firm were to directly offer higher wages to workers who train, then it is possible that workers might sign up for the training program but not put in any effort, thereby not incurring the subjective cost of training. We shall also assume that the firm has no special monopoly in the provision of training — the worker can always acquire training by paying for it outside the firm. So the value of training provision is no greater than $(1 - \lambda)k$, and can indeed be less, for example if the worker chooses not to undertake training.

Let us now analyze the training decision of a worker, given that she has chosen employment in the firm. Clearly, a worker will choose to train if and only if

$$\lambda k \leq -c(\theta) + R \tag{2}$$

Let us assume that $k < -c(b)+R$, so that the most productive worker will always choose to train even if she pays the entire cost of training. Assume also that $0 > -c(a)+R$, so that the least productive type of worker will never train, even if training is for free. Under these conditions, for any λ, there will be a type of worker $\theta(\lambda) \in (a, b)$ who is indifferent between training and not training (see Figure 1), while workers with lower ability, where $\theta < \theta(\lambda)$, will not train. The critical value $\theta(\lambda)$ is decreasing in R - a higher return to training makes less productive workers also willing to train - but to simplify notation we suppress the effect of R.

The utility of a worker from a package (w, λ) is given by

$$U(w, \lambda, \theta) = \begin{cases} w \text{ if } \theta < \theta(\lambda) \\ w + R - c(\theta) - \lambda k \text{ if } \theta \geq \theta(\lambda) \end{cases} \tag{3}$$

We note therefore that the utility of a worker is a continuous increasing function of her ability.Observe also that paying for training is an inefficient means of paying workers. If the firm incurs a cost of $(1 - \lambda)k$ on training a worker, the benefit to the worker equals $(1 - \lambda)k$ only if the worker's ability is greater than θ^*. If $\theta \in (\theta(\lambda), \theta^*)$, then the benefit to the worker is strictly

less than $(1-\lambda)$, as the worker in this case wouldn't have chosen training if she had to pay in full.

We assume that there are n firms in the labor market, where $n \geq 2$. Each firm chooses a pair (w, λ), and every worker responds by choosing employment at a firm which offers an optimal compensation package, corresponding to the worker's type θ. If a worker is indifferent between two or more firms, she is equally likely to work at any one of these firms. The labor market is in equilibrium if for any firm i, its profits from the chosen compensation package are greater than its profits from any alternative, given the offers made by the other firms, and given the optimizing response of workers.

We first show that there cannot be an equilibrium with positive employment where firms do not pay for training. If no firm pays at all for training, one firm can offer a selective wage increase by paying for training, thus attracting the more able workers. If it combines this with a reduction in the wage, it discourages the less able workers, who will not avail of training. This raises total productivity and more than pays for the small increase in average compensation.

To see this argument more formally, suppose that we have an equilibrium where no active firms pays for training. Since the market is competitive, firms must offer a wage equal to $\bar{\theta}$. Now let one firm, firm A, offer (w, λ) where $0 < \lambda < 1$ and $w < \bar{\theta} < w + (1-\lambda)k$. Let $\hat{\theta}$ denote the type of worker who is indifferent between the offers $(\bar{\theta}, 1)$ and (w, λ). $\hat{\theta}$ must be interior: it must be lower than θ^* since $\bar{\theta} < w + (1-\lambda)k$ (subsidization of training makes more workers choose to train). It must also be greater than a since $c(a) > R$ (by assumption, the least able workers do not want training even if it the firm pays all costs). Therefore the productivity of the workers who join firm A is given by $\mathbf{E}(\theta|\theta \geq \hat{\theta}) > \bar{\theta}$. Now by selecting (w, λ) such that $w + (1-\lambda)k$ is arbitrarily close to but greater than $\bar{\theta}$, firm A can make positive profits from this deviation. Thus there cannot be an equilibrium where firms do not pay at all for training.

Since firms will pay at least partly for general training in any equilibrium with positive employment, it follows immediately that there is excessive training relative to the social optimum: since training is subsidized, more workers will avail of it than when they pay for it themselves.

Let us now turn to a characterization of equilibrium. Let us consider an equilibrium where all active firms offer the same package (w, λ). Let us suppose that there are n such active firms, so that a fraction $1/n$ of each type (θ) of worker is employed by any firm. The firm's profits from the workers

who choose training is given by

$$\pi_T = \frac{1}{n}\int_{\theta(\lambda)}^{b} (\theta - w - (1-\lambda)k) f(\theta) d\theta \tag{4}$$

while its profits from workers who do not train is given by

$$\pi_U = \frac{1}{n}\int_{a}^{\theta(\lambda)} (\theta - w) f(\theta) d\theta \tag{5}$$

Now we must have that $\pi_T + \pi_U = 0$, under Bertrand competition — if profits were strictly positive, one firm could raise w slightly and attract all the workers.

Now it is also the case that either (i) the profits from training workers are zero at λ or (ii) λ is at its boundary value, 0. Otherwise, the firm could reduce λ slightly, and get all the productive workers from the other firms. In the case where $\lambda = 0$ and $\pi_T > \pi_U$, all firms strictly prefer to subsidize training, and thus all active firms must offer the same package. In the case where $\pi_T = 0$ at the interior value of λ, then some firms could offer training while other firms do not, since firms are indifferent between the two categories of worker. In any event, this possible asymmetry does not affect the level of wages and training subsidy available in the market.

Our main results can be summarized in the following theorem, which is proved in the appendix.

Theorem 1 *Suppose that $\bar{\theta} > r$. There always exists a labor market equilibrium, which must have positive training subsidies. In any equilibrium firms make zero profits, and either training is fully subsidized ($\lambda = 0$) or the firm makes zero profits from trained workers ($\pi_T = 0$). The extent of training is greater than the socially optimal level.*

The theorem focuses on the case where $\bar{\theta} > r$, so that the labor market is active in the absence of training. However, even if $\bar{\theta} < r$, so that in any equilibrium without training, there is no employment, training may allow an active labor market. A firm can offer $w < r$, $w + (1-\lambda)k = r$. As long as $\mathbf{E}(\theta|\theta \geq \theta^*) > r$ the firm can make positive profits by offering training.

To illustrate the results in the theorem, let us specialize to an example where θ is uniformly distributed on $[a, a+1], a > 0$. Let us also assume that $r = 0$. We also maintain the assumptions on the cost of training that

ensure that $\theta(\lambda)$ is interior. Under the uniform distribution, if $\theta(\lambda)$ is interior, $\mathbf{E}(\theta|\theta \leq \theta(\lambda)) = \frac{a+\theta(\lambda)}{2}$, while $\mathbf{E}(\theta|\theta > \theta(\lambda)) = \frac{a+1+\theta(\lambda)}{2}$. Hence the difference in productivity between the workers who train and those who do not is 1/2, independent of $\theta(\lambda)$.

Let us consider first the case where training costs are low; $k < 1/2$. We show first that there exists a symmetric equilibrium where each firm sets $\hat{\lambda} = 0$, so that it pays the entire training cost. The wage that is offered solves the zero profit condition, i.e.

$$\hat{w} = a + \frac{1}{2} - k[a + 1 - \theta(0)] \tag{6}$$

To see that this is an equilibrium, note that if a firm reduces the wage slightly, it will lose all its workers. Also, if the firm raises λ, it will also lose all its more able workers. Note also that this is the only possible equilibrium, since if a firm chooses $\lambda > 0$, the other firm will attract all the able workers by choosing a lower λ. We find in this equilibrium that firms pay for all the training, and furthermore, the average productivity of the training workers exceeds their total compensation, while the average productivity of those workers who do not train is below their compensation. We also find that there is excessive training relative to the social optimum , since $\theta(0) < \theta^*$.

Now let us consider the case where training costs are high; $k \geq 1/2$. There exists an equilibrium where each firm pays training costs equal to the expected productivity difference 1/2, i.e. that firm sets $(1 - \hat{\lambda})k = 1/2$. The wage is set equal to the average productivity of the workers who do not train, i.e. $\hat{w} = \frac{a+\theta(\lambda)}{2}$, implying that firms make zero profit for both types of workers. Thus, in this case there is no cross subsidization between those workers who train and those who do not. Once again there is excessive training since $\theta(\lambda) < \theta^*$, for $\lambda < 1$.

Let us consider briefly the welfare and policy implications of our explanation for employer subsidised training. Imperfect competition based explanations (e.g. Stevens or Acemoglu-Pischke) suggest that the pair consisting of the worker and her current employer will not be able to fully appropriate the return to training, due to poaching externalities. Thus market provision of training will be insufficient, which provides a rationale for government subsidies to training programs. In contrast, our explanation implies that training is excessive relative to the first best. This suggests that a tax on training might be an optimal response. Indeed, if the equilibrium value of λ is interior, then it is easy to see that a tax on training will reduce the extent of

the training subsidy, thus reducing the amount of excess training. However, one must exercise caution about this policy implication, since our model is simple, and leaves out important considerations. For instance, the screening role of training may result in a direct welfare gain, by promoting a superior allocation of workers to tasks within the organization.[5] Additionally, workers may be credit constrained, so that firm provision of training may play an important role in offsetting this capital market imperfection.

We have assumed that firms cannot condition the wage upon the training decision, i.e. both workers who train and those who do not are paid the same wage. The effect of this assumption is to constrain the subsidy to those workers who train so that it does not exceed the financial cost of training, k. This assumption is justified, in our view. If a firm fully paid for training and also pays a worker who trains a higher wage than a worker who does not, then a worker could simply shirk on training effort, thus not incurring the non-monetary cost of training. Nevertheless, let us briefly consider the implications of allowing firms to condition the base wage on the worker's training decision. For parameter values where the market equilibrium is such that firms make zero profits from trained workers, this modification would have no effect — in equilibrium the firm chooses an interior value of λ, and as such the constraint upon wages implied by our existing assumption is not binding. However, in the case where the market equilibrium is such that firms make positive profits from trained workers when training is fully subsidized, i.e. $\pi_T > 0$ in equilibrium, firms would have an incentive to raise wages of workers who choose to train and reduce the wages of workers who choose not to. In equilibrium, profits would be zero on both categories of workers, those who train and those who do not. As compared to the current model, even more workers would choose to train, thereby reducing the average productivity of both trained and untrained workers. Thus the provision of training would be even more excessive, as compared to the social optimum.

[5]Note however that it is possible to construct both a model where tax on training improves this screening function, and one where it retards the screening function.

3 Training, Signaling and Labor Market Competition

In this section we extend the model by adding a second period to the basic model described in the previous section, which constitutes the first period. In the second period, firms are assumed not to offer any training. Workers who did not receive any training in period one (which we shall refer to as unskilled), have unchanged productivity θ, while workers who did receive training in period one (skilled workers), have a higher productivity level $\gamma(\theta) > \theta$. Thus the social return to training for worker of type θ is $\gamma(\theta) - \theta$. We assume that the return to training is non-decreasing in θ. If it is strictly increasing, then we have complementarity between skills and ability, as is assumed by Acemoglu and Pischke [1] — indeed, in their model, this assumption is essential if firms are to provide training. However, we also allow for the possibility that $\gamma(\theta) - \theta$ is constant and independent of θ. That is, we allow for the possibility that $\gamma(\theta)$ is separable in ability and training, and can be written as $\theta + Z$, where Z is constant and independent of θ.

We shall consider three alternative scenarios for period two, based on different assumptions concerning the observability of individual ability, by the firm which provides training and outside firms. This will determine the nature of competition in the labor market.

3.1 Ability observed by all firms

Let us first consider the case where individual ability is common knowledge between employers in period two, and where firms can differentiate wages fully between individuals, without any constraints. These assumptions, especially with regard to information, are somewhat unrealistic, but the case is nevertheless useful as a benchmark. It is then straightforward to see that a standard competitive equilibrium prevails in period two. All workers with productivity above their reservation wage are employed. All employed workers are paid equal to their productivity, i.e. for unskilled workers $w^U = \theta$, while for skilled workers, $w^S = \gamma(\theta)$. The return to training of an individual worker is equal to $R(\theta) = \gamma(\theta) - \theta$. Although the return to the worker is type dependent, it is clear that the analysis of period one given above is unaffected by adding period two; the worker appropriates the entire return to training, R, in period two, while the period two profit of the firm is not

changed. Thus, the period one results still applies– firms will offer subsidized training in order to attract better workers in period one, even though they are unable to exploit these workers in period two.

3.2 Ability unobserved by outside firms, unexploitable by the inside firm

Now consider the case where individual productivity is not observable by the outside firms and where also the inside firm is unable to pay different wages to workers of the same skill level. This could arise because the internal firm also does not observe worker ability, or because of its inability to differentiate wages, for the reasons that we have set out earlier. Here we must also rule out the possibility that firms combine a higher wage offer with the use of selective dismissals of workers of lower quality. However, all firms can observe whether a worker is skilled or not, i.e. whether the worker received any training in period one, and can also differentiate wages according to skill level. In equilibrium, firms will offer both groups a wage equal to their expected productivity. Thus, the wage for unskilled workers, w^U, will be given by

$$w^U = E\left[\theta | \theta < \theta(\lambda)\right] \tag{7}$$

while the wage for skilled workers, w^S, will be given by

$$w^S = E\left[\gamma(\theta) | \theta \geq \theta(\lambda)\right] \tag{8}$$

Thus the wage premium for skilled workers equals the difference in average productivity between the two types of worker, with the firm getting no share of the increase in productivity induced by training. As in case 1 above, this implies that the period one analysis given in section 2 above is not affected: firms have an incentive to offer subsidized training arising from the period one effect on selection of workers.

However, a novel point is that the individual return to training for a worker is, on average, greater than the increase in productivity caused by training. To see this as simply as possible, assume that there is no complementarity between ability and training so that $\gamma(\theta) = \theta + Z$. In this case, the return to training of an individual worker is

$$R = w^S - w^U = \{E[\theta|\theta \geq \theta(\lambda)] + Z\} - E[\theta|\theta < \theta(\lambda)] > Z \qquad (9)$$

The reason for this is straightforward – as in Spence, training acts as a signal for higher ability, and thus workers have a greater incentive to undertake it. A numerical example can illustrate this effect. Assume that θ is uniformly distributed on $[a, a+1], a > 0$, and where $r = 0$. The wage of unskilled workers is $w^U = \mathbf{E}(\theta|\theta \leq \theta(\lambda)) = \frac{a+\theta(\lambda)}{2}$, while the wage of skilled workers is $w^S = \mathbf{E}(\theta|\theta > \theta(\lambda)) + Z = \frac{a+1+\theta(\lambda)}{2} + Z$. The wage premium, $w^S - w^U = 1/2 + Z$.

Thus, in this case there are two mechanisms that cause excessive training relative to the socially optimal level. Firms subsidize training to attract the more productive workers, and workers choose to train as a signal of higher productivity. Note also that the signalling effect of training follows directly from the feature that firms use training as a screening role in period one. Put differently, whenever firms use training to screen workers in period one, the screening will work as a signal in period two, and workers have an additional incentive to choose training. This effect will exist unless there is no informational effect in period two of the screening in period 1 (as in the first case in this subsection, where we assumed that individual productivity was observable in period two), or if there is no gain from the additional information in period two (e.g. if the wages for skilled and unskilled for some exogenous reason were assumed equal).

The results of our analysis so far are in sharp contrast to the results of Acemoglu and Pischke [1], where firms pay for general training because it yields an increase in productivity that exceeds the increase in wages. Here we see that the wage differential between skilled and unskilled exceeds the increase in productivity that is caused by training, and nevertheless the firms profit from paying for general training. This has clear implications for analyzing the empirical evidence on differentials between skilled and unskilled workers.

3.3 Ability observed and exploitable only by the inside firm

The final case that we analyze is where the inside firm observes the ability of the worker while the outside firm does not do so. Our model of the second period is rather similar to Acemoglu and Pischke [1]. We assume also that

the inside firm may condition its wage offer upon the ability of the worker. Competition for each worker is given by an English auction or bidding war, where each firm may raise its offer in response to the offer of its rival. We may model this competition equivalently as a sealed bid second price auction, where all n firms in the market submit wage offers, and the worker goes to the firm submitting the highest offer at the wage offered by the next highest bidder. Outside firms clearly suffer from the winner's curse in this situation. (The only reason the inside firm would let the outside firm win the bidding would be because the outside firm offers more than the worker's productivity.) Thus the equilibrium outside offer for trained workers will be $\gamma(\theta(\lambda))$, the productivity level of the lowest type of trained worker. Hence all trained workers will be retained by the inside firm, at a wage equal to $\gamma(\theta(\lambda))$. The profits made by the firm on trained workers in period two is given by:

$$\pi_{2T}(\lambda) = \frac{1}{n}\left(E\left[\gamma(\theta)|\theta \geq \theta(\lambda)\right] - \gamma(\theta(\lambda))\right) \tag{10}$$

Then consider wage determination for unskilled workers. If we assume that untrained worker's ability is also similarly observed by the inside firm, they will also be retained by the inside firm at a wage equal to the productivity of the lowest type, a. Thus the profits on unskilled workers in period two is given by

$$\pi_{2U}(\lambda) = \frac{1}{n}\left(E\left[\theta|\theta < \theta(\lambda)\right] - a\right) \tag{11}$$

Thus we see that period two profits may be greater for skilled or unskilled workers, depending on whether the dispersion of abilities above the lower bound is larger in the former group or the later. Note also that in contrast to cases 1 and 2 above, the inside firm does make excess profit in period two by exploiting its informational advantage relative to outside firms about workers' ability. This informational advantage will clearly affect firms' wage offers in period one. A firm's profits over the two periods (neglecting discounting) from the workers who choose training is given by

$$\pi_T = \frac{1}{n}\left(\int_{\theta(\lambda)}^{b} (\theta - w - (1-\lambda)k) f(\theta) d\theta + \pi_{LT}(\lambda)\right) \tag{12}$$

while its profits from workers who do not train is given by

$$\pi_U = \frac{1}{n}\left(\int_a^{\theta(\lambda)} (\theta - w) f(\theta) d\theta + \pi_{UT}(\lambda)\right) \tag{13}$$

Workers' return from training in period one is a higher wage in period two, given by

$$R = \gamma(\theta(\lambda)) - a \tag{14}$$

Taking into consideration period two does not affect the qualitative analysis of period one given above. The screening motivation for providing subsidized general training still applies, so firms will offer subsidized training even if there are no additional benefits to be reaped in period two. However, if period two profits are greater for skilled workers, $\pi_{2T}(\lambda) > \pi_{2U}(\lambda)$, firms' incentive to subsidize training in period one is strengthened. This additional effect is similar to Acemoglu and Pischke's [1] argument of training yielding an informational advantage.

There are many reasons to suppose that productivity dispersion is greater for skilled workers. In addition, it may be harder for the inside firm to exploit its superior information in the case of untrained workers than for trained workers. In the case of unskilled workers, such exploitation may require a wage reduction in period two, from the level $E[\theta|\theta < \theta(\lambda)]$ to a. On the other hand, for skilled workers, such exploitation can be achieved even with a wage increase which reflects their higher productivity. This may be an additional reason why period two profits are greater from trained workers.

Let us conclude our analysis of this case by considering the case where training and ability are not complementary in production, so that $\gamma(\theta) = \theta + Z$. In this case we note that all increases in productivity due to training will accrue to the worker. Nevertheless, the firm may make higher profits on trained workers than on untrained if productivity dispersion is greater for skilled workers. In any event, firms still have an incentive to train in order to induce a positive selection. This is in contrast with the results of Acemoglu and Pischke [1], where in the separable case, firms have no incentive to train.

To summarize, we have explored several different scenarios of second period labor market competition for trained workers. For reasons of signalling, workers have greater incentives to train in a dynamic context. Although firms may be able to make greater profits from trained workers, this is not necessary for them to offer training in the first place. Indeed, we find that

firms may offer training even though the wages of trained workers rise by more than the increase in productivity induced by such training.

4 Empirical Implications and Evidence

At the more general level, both our idea here and the idea of Acemoglu and Pischke [2] explain firm-provision of general training by wage compression. However, there are also crucial differences. In Acemoglu and Pischke [2], wage compression is across skill levels, so that training leads to a smaller increase in the wage than in productivity. In other words, firms pay for training in order to exploit future differences in productivity. In our model, there is wage compression within skill levels, between workers of different ability, and firms offer cheap or free training so as to attract the more productive workers. In other words, training works as a screening device, paid for by current productivity differences. There is no need for the wage increase induced by training to be smaller than the increase in productivity induced by training; indeed, under plausible assumptions the wage premium can exceed the productivity increase associated with training as training may involve a signal of higher innate productivity.

In more practical terms, the theory of Acemoglu and Pischke would suggest that firm-provided general training will take place in environments where the worker's return to training is small, for example in countries where skill premia are small. In contrast, our theory suggest that firm-provided training will take place in environments where firms for some reason are constrained in differentiating pay between individuals of different ability not directly related to training. For example, one could think of countries or environments where the wage distribution is compressed also when one compare workers of the same observational characteristics, including education and training levels.

Acemoglu and Pischke [2] use German firms' provision of firm-sponsored general training in the form of apprenticeships as a leading example of their theory, based on the claim that wages are more compressed across skills in Germany than in the US. However, the evidence in Freeman and Schettkat [8] show that the wage increase from unskilled to skill level II (some college/ hauptschule + apprenticeship/meister) is in fact greater in Germany (Table 2), casting some doubt on this claim. In contrast, the residual variance in

a standard wage equation is greater in the US than in Germany, suggesting that there is more wage compression among workers of the same observational skill level in Germany than in the US. Within our model this would suggests more firm-provided general training in Germany than the US.

A recent empirical study by Brunello [11] is also relevant in this context. Brunello calculates measures of wage compression within groups that are sorted according to country, occupation, sector and age. He then finds a significant positive relationship between wage compression and firm-provided general training, over a sample with the associated groups. It is however not clear whether this wage compression should be interpreted as relating to individuals of different training levels (the Acemoglu and Pischke's story) or individuals of different productivity levels not directly related to training (our story). Thus, Brunello's evidence is consistent with either of the hypotheses, and can not be used to distinguish between them.

There is also evidence showing that firms are more likely to provide training in the public sector, and in unionized sectors, cf. OECD [25], see e.g. tables 3.11 and 3.12. This is also consistent with either of the stories, as wages generally are more compressed both in the public sector and in unionized sectors, cf. Gregory and Borland ([17]) and Wallerstein ([29]).

Autor[5] discusses why temporary help supply (THS) establishments provide free training in portable computer skills. A basic assumption in Autor's model is that training is more productive and therefore more valuable for high productive workers. As in our model, firms offer training to induce self-selection so as to attract high ability workers. In contrast to our model, firms profit from doing this by obtaining an informational advantage about the ability of the workers, which then can be exploited by offering wages below marginal productivity.However, this justification begs the question of why non-training firms cannot obtain the same information by undertaking a similar assessment of computer skills as the training firms do, at low cost. Furthermore, one would expect that also non-training firms would obtain information about the ability of their employees on the basis of possible reports and reactions from the clients of the THS firms, where the employees actually work. Our model offers an alternative interpretation of why THS establishments provide free general training: Because THS firms initially are unable to observe worker ability, and thus unable to differentiate the wage accordingly, free training is offered to attract the high ability workers (assuming that training is more valuable or less costly to high ability workers). As shown in our model, this may constitute a sufficient motivation for firms

to provide general training, and no ex-post informational advantage of the incumbent firm is necessary.

Our paper is closely related to Cappelli [12], who asks the question, why do employers pay for the college education of their employees.[6] It documents substantial payment for college education, and argues that employers do this in order to improve employee quality and reduce turnover. The empirical evidence shows that provision of tuition assistance improves the quality of hires, much as in the model of this paper. Cappelli's paper is indeed complementary to ours – while we have focused on a theoretical analysis of a labor market where training helps select better workers, his paper shows that such a mechanism is empirically important. The model presented in this paper may help better understand this empirical evidence. For example, Cappelli suggests that tuition assistance is so common that the employers offering it may not be able to induce positive self-selection. In terms of our model, this is not a puzzle, since in symmetric equilibrium all employers offer assistance, but the average quality of hires is the same as in the situation where no one offers assistance. Nevertheless, it is optimal for an individual employer to offer assistance since otherwise better quality workers would go to the other employers.

5 Concluding remarks

We argue that if employers are unable to differentiate wages according to worker ability, for informational or institutional reasons, this will provide employers with an incentive to provide subsidized general training to attract the more able workers. Under fairly weak conditions, labor market equilibrium will always involve some subsidization of general training. In contrast to existing popular explanations, firms may benefit from subsidizing training even if they do not capture any of the return from training in the form of a productivity increase exceeding the associated wage increase. Our explanation for the firm provision of training appears consistent with available evidence, and can be viewed as being complementary to existing theories.Since training is paid for by contemporaneous productivity differences between workers, not reflected in their wage, our theory is more applicable in situations where training occupies a relatively small fraction of the employee's

[6] We came across his paper after having written the first versions of ours.

work time. Thus it is a plausible explanation of situations where employers pay for evening classes, or where the trainee goes to college for one day a week.

Our explanation is consistent with survey evidence based on in-depth interviews of 11 Norwegian firms, Larsen et al. ([20], chapter 6). A few of the firms had a recruitment strategy with emphasis on a good apprentice system and an explicit aim of being viewed as a firm oriented towards competence among its employees. Almost all firms expressed a desire to have employees who themselves took a responsibility for his/her own task and development, and individual initiative to training was viewed as a signal that the employee was of the type that the firm was interested in. Finally, in all firms the salary was basically associated with the job, either through collective agreements or job descriptions, while any individual variation was negligible.

6 Appendix

We now prove theorem 1. We have already proved the characterization results in the text of the paper, so it remains to prove existence of equilibrium. Consider the function $\phi(\lambda)$ defined on $[0,1]$ by

$$\phi(\lambda) = \mathbf{E}[\theta|\theta \geq \theta(\lambda)] - \mathbf{E}[\theta|\theta < \theta(\lambda)] - (1-\lambda)k \tag{15}$$

$\phi(\lambda)$ is the difference in profits between workers who train and those who do not at λ. Since the probability distribution over productivities is atomless, (being given by the density f), the conditional expectations are continuous functions of λ. Thus ϕ is a continuous function. We have assumed that when $\lambda = 1$, $\theta(1) = \theta^{**}$ is in the interior of $[a,b]$. Thus $\phi(\lambda) > 0$ if λ is sufficiently close to 1, since the difference in productivities between workers who train and those who do not is strictly positive.

From our characterization results, a necessary condition for the pair $(w,\lambda), \lambda > 0$, to be an equilibrium profile is that $\phi(\lambda) = 0$. Furthermore, if $\phi(0) \geq 0$, then workers who train are more profitable than those who do not, even when the firm pays for all training.

Let us define $\hat{\lambda}$ as follows. If $\phi(0) \geq 0$, let $\hat{\lambda} = 0$. Otherwise, let $\hat{\lambda}$ denote the smallest value of λ such that $\phi(\lambda) = 0$. (Although our assumptions do not ensure uniqueness of the zeroes of the ϕ function, continuity of ϕ ensures that $\hat{\lambda}$ is well defined, and lies in the unit interval).

We now show that there is an equilibrium provided that two or more firms offer the pair $(\hat{w}, \hat{\lambda})$, where $\hat{w}$ is the wage which ensures zero profits, i.e. $\hat{w} = \bar{\theta} - [1 - F(\theta(\hat{\lambda}))](1 - \hat{\lambda})k$.

Clearly, this pair satisfies the sufficient conditions in the characterization results. If any firm offers a lower subsidy to training, then this is not profitable since at $\hat{\lambda}$, workers who train are at least as profitable as those who do not. If $\hat{\lambda} > 0$, then a firm which offers a higher subsidy will attract all the better workers, but this is also not profitable. ■

References

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

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

[3] Adams, J.S. (1963). Towards an Understanding of Inequity, *Journal of Abnormal and Social Psychology,* 67, 422-436.

[4] Arulampalam, W, A.L. Booth, and M.L.Bryan (2002). Work-Related Training and the New National Minimum Wage in Britain. IZA Discussion Paper 595, Bonn.

[5] Autor, D.H. (2000). Why do Temporary Help Firms Provide Free General Training?, *Quarterly Journal of Economics* 116, 1409-1448.

[6] Becker, G. (1964), *Human Capital*, Chicago: University of Chicago Press.

[7] Bewley, T. (1999), *Why Wages don't Fall in Recessions,* Cambridge: Harvard University Press.

[8] Bhaskar, V. and T. To (2002), Insufficient versus Excessive Training in Oligopsonistic Labor Markets, mimeo.

[9] Blau, F.D. and L.M. Kahn (2000). International Differences in male Wage Inequality: Institutions versus Market Forces. *Journal of Political Economy 104*, 791-837.

[10] Bolton, G., and A. Ockenfels (2000). Erc: A Theory of Equity, Reciprocity and Competition, *American Economic Review* 90, 166-193.

[11] Brunello, G. (2002). Is Training More Frequent when Wage Compression is Higher? Evidence from 11 European Countries. CESifo Working paper 637.

[12] Cappelli, P. (2002). Why Employers Pay for College, NBER Working paper 9225.

[13] Burdett, K. and E. Smith (1996). Education and matching externalities. In A. Booth and D.J. Snower (eds). *Acquiring Skills. Market failures, their symptoms and policy and policy responses,* CEPR, Cambridge University Press.

[14] Dustmann, C., and U. Schönberg (2002). Training and Unions, mimeo, University College London.

[15] Fehr, E. and K. Schmidt (1999), A Theory of Fairness, Competition and Cooperation, *Quarterly Journal of Economics* 114, 817-868.

[16] Freeman, R and R. Schettkat (2001). Skill Compression, Wage Differentials, and Employment: Germany vs the US. *Oxford Economic Papers 3*, 582-603.

[17] Gregory, R.G. and J. Borland (1999). Recent Developments in Public Sector Labor Markets. In O. Ashenfelter and D.Card (eds). *Handbook of Labor Economics*, Vol 3, North-Holland.

[18] Katz, E. and A. Ziderman (1990). Investment in General Training: The Role of Information and Labour Mobility. *Economic Journal*, 100, 1147-1158..

[19] Kessler, A. and C. Lulfesmann (2001), The Theory of Human Capital Revisited: On the Interaction of General and Specific Investments, mimeo, Bonn.

[20] Larsen, K.A., F. Longva, A. Pape, and A. N. Reichborn (1997). Bedriften som lærested (The firm as place of training). Fafo report 212, Oslo, Norway.

[21] Loewenstein, G.F., L. Thompson, and M.H.Bazerman (1989). Social Utility and Decision Making in Interpersonal Contexts. *Journal of Personality and Social Psychology LVII,* 426-441.

[22] Malcomson, J. (1984). Work Incentives, Hierarchy and Internal Labor Markets, *Journal of Political Economy* 92, 486-507.

[23] Malcomson, J., J. Maw and B. McCormick. (2002). General Training by Firms, Apprentice Contracts and Public Policy, CESifo Working Paper No. 696 (4).

[24] Marsden, D., S. French, and K. Kubo. (2001). Does Performance Pay De-Motivate, and Does It Matter?, Centre for Economic Performance Discussion Paper

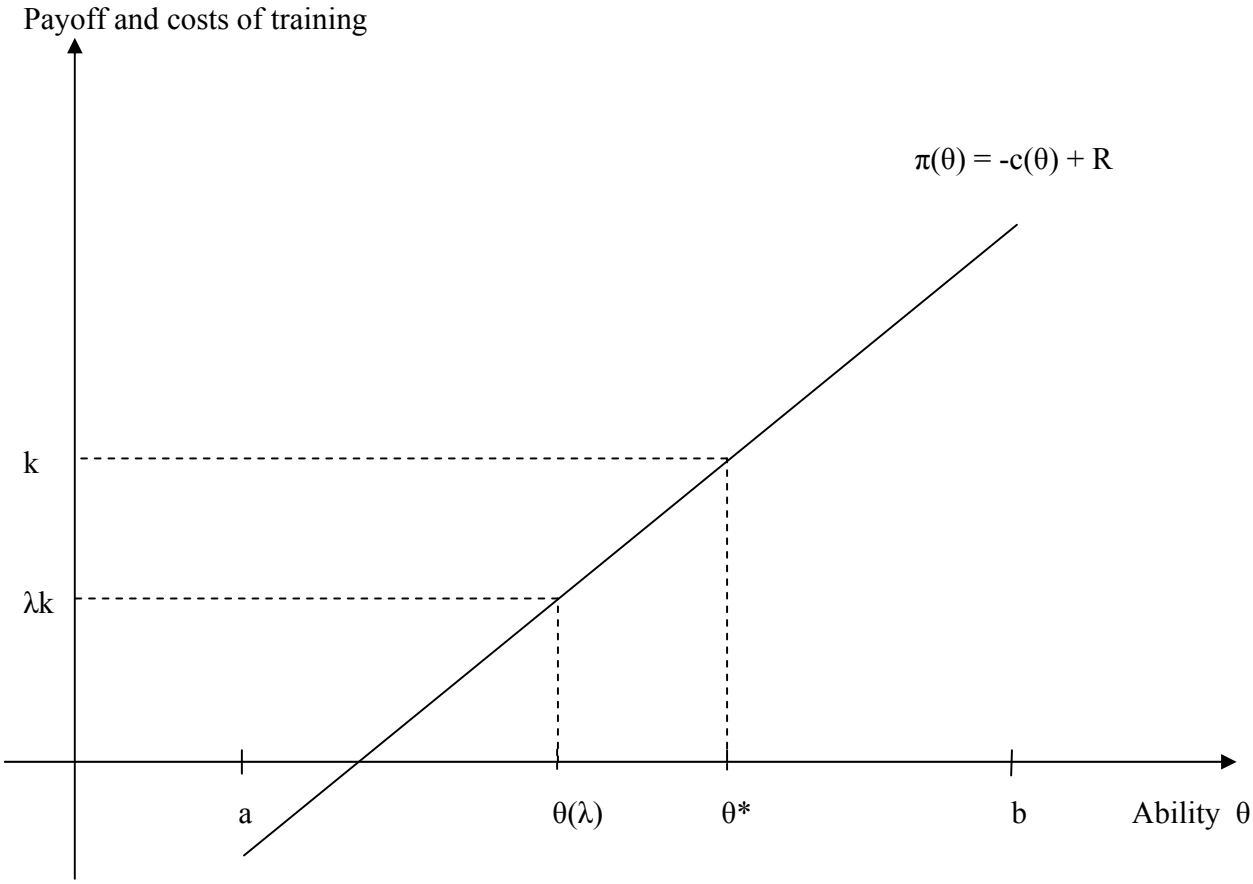
[25] OECD (1999). Employment Outlook. OECD, Paris.

[26] Spence, M. (1973), Job Market Signaling, *Quarterly Journal of Economics 87* .

[27] Stevens, M. (1994), A Theoretical Model of On-the job Training with Imperfect Competition, *Oxford Economic Papers* 46, 537-562.

[28] Stevens, M. (1996), Transferable training and poaching externalities. In A. Booth and D.J. Snower (eds). *Acquiring Skills. Market failures, their symptoms and policy and policy responses,* CEPR, Cambridge University Press.

[29] Wallerstein, M. (1999). Wage-Setting Institutions and Pay Inequality in Advanced Industrialized Societies. *American Journal of Political Science 43*, 649-680.


Payoff and costs of training
π(θ) = -c(θ) + R
k
λk
a
θ(λ)
θ*
b
Ability θ