Wage Differentiation via Subsidised General Training *

V. Bhaskar
University of Essex†

Steinar Holden
University of Oslo‡

December 2002
(First version: May 2002)

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 for firm subsidised general training rely on wage compression *across* skill categories.

JEL Categories: J31, D82.

1

1 Introduction

In his seminal work, Becker [7] 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 the that general training induces a larger increase in productivity than in the wage. Following Pigou, a number of authors such as Stevens [28] and Acemoglu and Pischke [1] argue that this is possible with an imperfectly competitive labor market.[1] As noted by Acemoglu and Pischke, the key issue is whether the distribution of wages is compressed relative to the distribution of marginal products arising from general training.[2] If this is the case, i.e. *across* skill wage differentials are compressed, the worker's return to training is less than the rise in her marginal product. Then the worker has insufficient incentives to engage in general training: On the other hand, the firm does have an incentive to provide general training, as the wage increases less than the productivity (although poaching externalities may imply that market provision of general training will be insufficient). This explanation is widely used in order to explain the difference in training systems between Germany and the United States or UK. It is argued that differentials between skilled and unskilled wages are lower in Germany, which provides incentives for employers to finance general training.

In this paper we provide a different explanation for the employer provision of general training, in the context of a *competitive* labor market. We shall 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: for example, the ability of the worker may not be observable at the time of hiring, and may not be verifiable subsequently,

---

[1]Other explanations include that general and specific skills are complementary (see discussion and references in Stevens ([29])), asymmetric information (Katz and Ziderman ([20]) and Acemoglu and Pischke ([1])) and matching frictions (Burdett and Smith, [14])

[2]Bhaskar and To [9] 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 as excessive incentives to acquire general training. Dustmann and Schönberg [15] 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.

so that wages cannot be credibly conditioned on the employer's subsequent observation. The inability to wage differentiate may also depend upon institutional factors; trade unions may oppose employer discretion in wage gradations within skill levels and enforce uniformity within skill categories. 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 selection of more able workers. In a competitive labor market, 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 wage premium which is equal to the increase in productivity caused by training. If fact, if training works as a signal of higher productivity, the wage premium 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.

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. It also shows that a minimum wage increases the subsidy to training. Section 3 extends this analysis by considering competition in the market for trained workers and shows that our basic results continue to be true. The final section discusses the empirical evidence and concludes.

2   The Model

We assume that there is a continuum of workers. A worker's productivity (or ability) equals θ, where θ is distributed on 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 also assume that there are at least two firms, so that the labor market is competitive.

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. In this case, clearly the wage cannot be ability dependent. More realistically, we may assume that ability cannot be observed by the firm at the time of hiring, but that the firm gradually gets to observe θ over time, as the worker engages in productive activity. Nevertheless, if θ is not verifiable, the firm cannot credibly commit, at the time of hiring, to raise the worker's wage in line with her productivity. If the firm were to promise to raise w in line with its observation of θ, it would have an incentive to claim that θ was low, in order to reduce its wage payment. That is, the firm suffers from a moral hazard problem which makes its promise to raise w in line with θ not credible.[3] In contrast, the firm can easily contractually commit to a training program, which is available, in principle, to all workers.

2. 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, [18], Loewenstein, Thompson, and Bazerman [23] or Bolton and Ockenfels, [11]. Bewley's [7] survey shows that employers are wary of adopting internal wage policies which have deleterious effects on worker morale. Marsden, French and Katsuyuki [26] analyze evidence from a survey on performance-related pay in the British public sector services. They find widespread de-motivating effects arising

---

[3]It may be argued that the firm's reputational concerns can alleviate this moral hazard problem, but this possibility maybe limited if outside workers cannot observe the productivity of inside workers, and hence cannot verify whether the firm has carried out its promise.

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. A training program which induces self-selection results in a *fair allocation*, in the sense of Foley [17], so that the worker who does not train does not envy the allocation of the worker who chooses to train. In contrast, wage differentiation according to ability results in envy. Thus even if ability was observable, employers may prefer to induce self-selection rather than tailor wages directly to ability, since this will not generate resentment on the part of less able workers.

3. Considerations of fairness may interact with moral hazard considerations and make wage differentiation impossible. The literature on employer moral hazard in wage setting has pointed out that if the number of workers is sufficiently large, such moral hazard may be overcome by tournament schemes where the total wage bill is fixed - see, for example, Malcomson [24]. 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.

4. Finally, there may be institutional constraints on differentiating wages among workers of equal skill levels. Unions are often opposed to discretion 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 and centralized wage setting and wage compression, (see Blau and Kahn [10] and Wallerstein [30]).

The key point of this paper is that the firm provides general training as a way of credibly differentiating effective compensation in order to reflect differences in ability. To sharpen the focus on this motivation for providing

---

[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, we assume that training does not affect a worker's productivity in the current period. However, before we turn to this, consider 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. The monetary cost of training a worker is k. In addition, training entails a non-monetary cost or disutility to the worker, which depends on her ability, and is denoted by $c(\theta)$, where 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

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

Let θ^{**} be the value of θ such that $\pi(\theta) = k$, i.e. the payoff to the worker is equal to the monetary cost of training. We shall assume in this section that R coincides with the social return to training, so that θ^{**} is the socially optimal level of training. Fig. 1 illustrates this situation – when the worker pays for his training in full, only workers with productivity to the right of θ^{**} will choose to train.

We shall assume that each firm offers a contract of the type (w, λ), where w denotes the wage which paid, and $\lambda \in [0, 1]$ is the fraction of the training cost borne by the worker. We shall assume also that the firm cannot condition the wage payment on the training decision, i.e. it cannot pay distinct wages to those workers who choose to train and those workers who choose not to. The firm might conceivably want to do this; however, we shall assume that the workers do not have to commit to train at the time that they choose to work at the firm. Furthermore, 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.
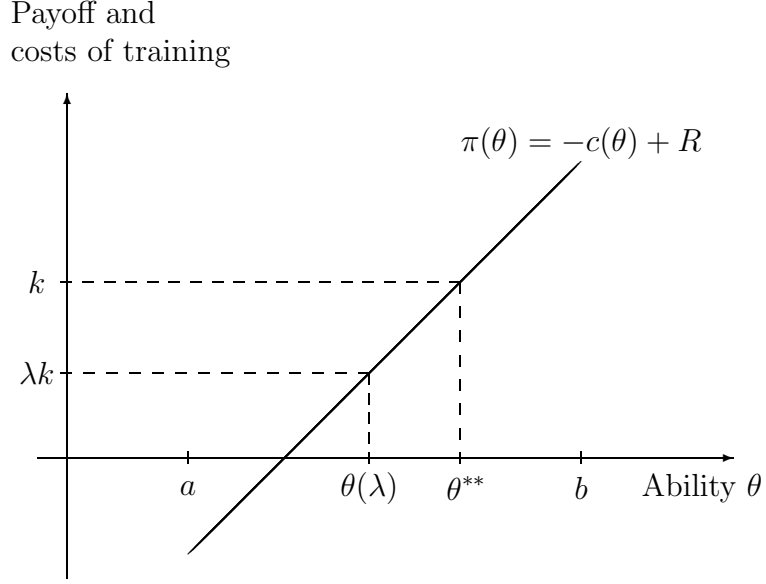
Payoff and
costs of training

Figure 1: Ability and the Costs/Benefits of 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 the $k < -c(b) + R$, 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). Clearly, the critical value $\theta(\lambda)$ is decreasing in R - a higher return to training makes less productive workers also willing to train - but to simply notation we suppress the effect of R. The socially optimally level of training, θ^{**}, prevails when workers pay the entire cost of training, i.e. $\theta(1)$, such that $\pi(\theta) = k$.

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

$$U(w, \lambda, \theta) = \begin{cases} w \text{ if } \theta \leq \theta(\lambda) \\ w + R - c(\theta) - \lambda k \text{ if } \theta \in (\theta(\lambda), \theta^{**}) \\ w + (1 - \lambda)k \text{ if } \theta \geq \theta^{**} \end{cases} \tag{3}$$

We note therefore that the utility of a worker is a continuous increasing function her ability as long as $\lambda < 1$. Observe also that the 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 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 all active firms do not pay for training. Since the market is competitive, they must offer a wage equal to $\bar{\theta}$. Now let one firm, firm A, offer (w, λ) where $w < \bar{\theta} < w + (1 - \lambda)k$ (implying that $\lambda < 1$). 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 $\bar{\theta} < w + (1 - \lambda)k$. It must also be greater than a since $c(a) > R$. 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). A firm's profits from the workers who choose training is given by

$$\pi_T = \frac{1}{2} \int_{\theta(\lambda)}^{b} (\theta - w - (1 - \lambda)k) f(\theta) d\theta \tag{4}$$

8

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

$$\pi_U = \frac{1}{2} \int_a^{\theta(\lambda)} (\theta - w) f(\theta) d\theta \tag{5}$$

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, $\mathbf{E}(\theta|\theta \leq \theta(\lambda)) = a + \frac{\theta(\lambda)}{2}$, while $\mathbf{E}(\theta|\theta > \theta(\lambda)) = a + \frac{1+\theta(\lambda)}{2}$. 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 $k < 1/2$. 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.

$$\hat{w} = a + \frac{1}{2} - k(1 - \theta^{**}) \tag{6}$$

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 $k \geq 1/2$. There exists an equilibrium where each firm pays training costs equal to the expected productivity difference $1/2$, i.e. that firm sets $1 - \hat{\lambda} = 1/(2k)$. The wage is set equal to the average productivity of the workers who do not train, i.e. $\hat{w} = a + \frac{\theta(\lambda)}{2}$, 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 $\lambda < 1$.

2.1 The Effects of a minimum wage

There is a considerable literature on the effects of minimum wages of training. In competitive labor markets, it is argued that firm financed general training is one form of worker compensation. If a minimum wage is imposed, firms will respond by reducing non-wage benefits such as training. We now show that this presumption is false once one allows for heterogeneity in worker ability. Indeed, a minimum wage may increase the equilibrium level of training in the labor market. The intuition is the following: in a competitive labor market, minimum wage must be paid for, in the form of higher worker productivity. Typically, this is achieved by a reduction in aggregate employment, which raises the marginal revenue product of labor, say by raising the

product price. Thus the minimum wage raises the productivity of all workers. This also increases the *difference* in productivity between able and less able workers. Since able workers are more attractive, firms respond by increasing the training subsidy.

To see this argument formally, let us consider a situation where firms initially pays only a part of the training costs, i.e. that λ is at an interior value. (Clearly, if the firm pays in full initially, there is no scope for increasing training in our setup, and a small rise in the minimum wage will have no impact.) Our focus is on the effect of a minimum wage upon the equilibrium value of λ.

A minimum wage will cause rationing in the labor market so that only a fraction β of the labor force may be employed. As shown above, when λ is interior, each firm makes zero profits on skilled and unskilled workers separately. Recall at the time of hiring, the firm cannot observe the worker's productivity. Thus the probability that the firm allocates the job to one worker rather than another cannot depend upon the worker's productivity. Therefore rationing will be uniform, in the sense that the probability of any worker getting a job does not depend on her type. The total level of output, y, will be given by

$$y = \beta \int_a^b \theta f(\theta) d\theta \qquad (7)$$

To capture the effect of lower employment induced by a minimum wage, we extend the model slightly by assuming that the product price $p(y)$ is a decreasing function of total output. Equilibrium in an unconstrained labor market corresponds to the case where $\beta = 1$, with the corresponding total output y^*. The revenue productivity of a worker is thus $p(y^*)\theta$. Our previous analysis has thus implicitly normalized $p(y^*)$ to 1.

A minimum wage will induce a lower level of employment in the labor market so that output falls and the product price rises. We show that a small minimum wage (i.e. one which is just above the free market level) leads to a larger subsidy to training, and therefore a larger fraction of employed workers undertake training.

Proposition 2 *A minimum wage strictly increases the firm subsidy to training, unless firms are fully paying for training.*

Proof.   Consider the case where the new equilibrium is also characterized

by a zero profit condition on skilled and unskilled workers. (The minimum wage may increase the training subsidy up till 100 percent, $\lambda^m = 0$, in which case the Proposition clearly holds.) For unskilled workers, the zero profit condition implies that

$$p^m \mathbf{E}(\theta | \theta < \theta(\lambda^m)) = w^m \tag{8}$$

where w^m denotes the minimum wage, p^m the equilibrium price post minimum wage, and λ^m the training component of the post minimum wage contract. Similarly, for skilled workers, the zero profit condition is

$$p^m \mathbf{E}(\theta | \theta \geq \theta(\lambda^m)) = w^m + (1 - \lambda^m)k \tag{9}$$

To prove the Proposition, assume the opposite, that the minimum wage does not affect firms' subsidy, i.e. $\lambda^m = \lambda$. In this case workers' choice of whether to train will not be affected, implying that $\theta(\lambda^m) = \theta(\lambda)$. Thus the average productivity of unskilled workers is unchanged with the minimum wage, and by the zero profit condition for unskilled workers, p^m must rise proportionally with w^m. However, this implies that the profits on skilled workers is strictly positive, given by

$$p^m \left\{ \mathbf{E}(\theta | \theta \geq \theta(\lambda^m)) - \frac{w^m}{p^m} - \frac{(1 - \lambda^m)k}{p^m} \right\} > 0 \tag{10}$$

where this inequality follows from the initial zero profit condition that $\mathbf{E}(\theta | \theta \geq \theta(\lambda)) - \frac{w^*}{p^*} - \frac{(1-\lambda^*)k}{p^*} = 0$.(the only difference between the two conditions being $p^* < p^m$.) Intuitively, the minimum wage raises the revenue productivity of all workers, including skilled workers, by raising the market price. If the amount spent on training does not rise, then profits on skilled workers must become positive. Thus, in equilibrium, the subsidy to training must also rise in order restore profits to zero. In the same way, it is straightforward to show that if $\lambda^m > \lambda$, then the average productivity of unskilled workers would rise, making it even more profitable to hire skilled workers, thus making firms reduce λ^m. ∎

Our results are consistent with the findings of Arulampalam. Booth and Bryan [4], who find no evidence that minimum wages reduced firm provision of training. They also find some evidence that minimum wages increase the level of training provided by firms.

3 Training, Signaling and Labor Market Competition

We have so far assumed that the worker's return to ability (R) is exogenous, and unrelated to her ability, and that this coincides with the social return. This implies that the worker is entirely able to capture the increase in productivity due to general training. This assumption has been in order to focus attention on the main point of the paper, since other work has already argued that the training firm will have an incentive to pay for general training if it is able to capture part of the return. We now consider a dynamic model, with two periods. The basic model described in the previous section 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

$$w^U = E\left[\theta | \theta < \theta(\lambda)\right] \tag{11}$$

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

$$w^S = E\left[\gamma(\theta) | \theta \geq \theta(\lambda)\right] \tag{12}$$

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:

14

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

$$R = \{E\left[\theta | \theta \geq \theta(\lambda)\right] + Z\} - E\left[\theta | \theta < \theta(\lambda)\right] > Z \qquad (13)$$

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 $w^U = \mathbf{E}(\theta | \theta \leq \theta(\lambda)) = a + \frac{\theta(\lambda)}{2}$, while the wage of skilled workers is $w^S = \mathbf{E}(\theta | \theta > \theta(\lambda)) + Z = a + \frac{1 + \theta(\lambda)}{2} + Z$. The wage premium, $w^S - w^U = 1/2 + Z$.

Thus, in this case there are two mechanisms that causes 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 exists unless the 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 both firms 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:

$$\pi_{2T}(\lambda) = E\left[\gamma(\theta)|\theta \geq \theta(\lambda)\right] - \gamma(\theta(\lambda)) \tag{14}$$

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

$$\pi_{2U}(\lambda) = E\left[\theta|\theta < \theta(\lambda)\right] - a \tag{15}$$

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

$$\pi_T = \frac{1}{2} \int_{\theta(\lambda)}^b (\theta - w - (1-\lambda)k)f(\theta)d\theta + \pi_{LT}(\lambda) \qquad (16)$$

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

$$\pi_U = \frac{1}{2} \int_a^{\theta(\lambda)} (\theta - w)f(\theta)d\theta + \pi_{UT}(\lambda) \qquad (17)$$

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

$$R = \gamma(\theta(\lambda)) - a \qquad (18)$$

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. However, while there is strong evidence that wages are generally more compressed in Germany than in the US, the evidence in Freeman and Schettkat [9] 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 the claim that wages are more compressed across skills in Germany than in the US. However, the residual in a standard wage equation, is greater in the US than in Germany. This suggests 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 [12] 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 [27], 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 ([19]) and Wallerstein ([30]).

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. However, 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. A key point in Autor's story is that firms performing training will observe worker ability, whereas non-training firms will not. Autor justifies this assumption by the empirical feature that the training provided by firms began and ended with assessment. However, against this justification one can argue that other firms at fairly low cost can perform the same assessment that the training firms performed. 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 [13], who asks the question, why do employers pay for the college education of their employees.[5] 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.

Our explanation is consistent with survey evidence based on in-depth interviews of 11 Norwegian firms, Larsen et al. ([22], 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.

To conclude, our basic argument is if employers are unable to differentiate

---

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

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. Furthermore, the introduction of a minimum wage will increase firms' subsidization of 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.

5   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

$$\phi(\lambda) = \mathbf{E}[\theta | \theta \geq \theta(\lambda)] - \mathbf{E}[\theta | \theta < \theta(\lambda)] - (1 - \lambda)k \qquad (19)$$

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 conditional expectations 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 $\phi(\lambda) = 0$.

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 $\phi(\lambda) = 0$. (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] Akerlof, G., and J. Yellen (1990). The Fair Wage-Effort Hypothesis and Unemployment, *Quarterly Journal of Economics* 105, 255-283.

[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* 115,

[6] Bac, M. (2000),On-the-Job Specific Training and Efficient Screening *Journal of Labor Economics* 18(4): 681-701.

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

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

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

[10] 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.

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

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

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

[14] 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.

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

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

[17] Foley, D. (1967). Resouce Allocation and the Public Sector, *Yale Economic Essays* 7, 45-98.

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

[19] 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.

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

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

[22] 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.

[23] 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.

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

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

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

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

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

[29] 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.

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