DISCUSSION PAPER SERIES

No. 2536

WHY DO FIRMS INVEST IN GENERAL TRAINING? 'GOOD' FIRMS AND 'BAD' FIRMS AS A SOURCE OF MONOPSONY POWER

Alison Booth and Gylfi Zoega

LABOUR ECONOMICS

Centre for Economic Policy Research

www.cepr.org

ISSN 0265-8003

WHY DO FIRMS INVEST IN GENERAL TRAINING? 'GOOD' FIRMS AND 'BAD' FIRMS AS A SOURCE OF MONOPSONY POWER

Alison Booth, University of Essex and CEPR
Gylfi Zoega, Birkbeck College, London and CEPR

Discussion Paper No. 2536
August 2000

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 2536

August 2000

ABSTRACT

Why Do Firms Invest in General Training? 'Good' Firms and 'Bad' Firms as a Source of Monopsony Power*

We develop a model demonstrating conditions under which firms will invest in the general training of their workers, and show that firms' incentives to invest in general training are increasing in task complexity. Workers' heterogeneous observable innate ability affects the variety of tasks that can be performed within a firm. This gives monopsony power to firms with 'better' workforces. As a result such firms are willing to expend resources to provide workers with general training. Since the degree of monopsony power is increasing with task complexity, firms whose workforces undertake more sophisticated tasks are more willing to finance general training. We conclude that training will take place in better-than-average firms, while bad firms will have underperforming but overpaid workers who are not likely to be trained by their current employer.

JEL Classification: J24, J31 and J42
Keywords: firm-financed general training, hierarchical assignment models and monopsony

Alison L Booth
ISER
University of Essex
Wivenhoe Park
Colchester CO4 3SQ
UK
Tel: (44 1206) 873789
Fax: (44 1206) 873151
Email: albooth@essex.ac.uk

Gylfi Zoega
Department of Economics
Birkbeck College
7-15 Gresse Street
London W1P 2PA
UK
Tel: (44 20) 7631 6406
Fax: (44 20) 7631 6416
Email: gzoega@econ.bbk.ac.uk

* We would like to thank, for their helpful comments on an earlier draft, Ron Smith, Hugh Davies, Steve Pischke and seminar participants at the European Summer Symposium in Labour Economics at Ammersee, Germany, in September 1999 and at the 1st EALE/SOLE World Conference in Milan in June 2000. Financial support from the European Commission TSER Award 'Labour Demand, Education and the Dynamics of Social Exclusion' (Grant No. SOE2-CT97-3052) is gratefully acknowledged.

Submitted 14 July 2000

NON-TECHNICAL SUMMARY

Firms exist not only because they are an efficient way of coordinating economic activity but also because they foster collaboration with colleagues who can supplement our own abilities and compensate for our weaknesses, teach us how to do our tasks better and allow us to specialize in what we do best. Moreover, colleagues can generate a stimulating and challenging working environment and they can impose sanctions on us when we fail to live up to prevailing norms. It follows that the quality of our colleagues matters and if some firms have better quality workers than others these provide for a better working environment. It is the objective of this paper to tie these commonplace observations from the personnel economics literature into the human-capital literature, in particular the literature on on-the-job training.

One of the tenets of labour economics is the distinction between general and firm-specific training due to Becker (1964). When training is general, workers must pay for it themselves. Since the market for their services is perfectly competitive, there is no divergence between their wages and marginal product. When the training is firm-specific, however, workers and firms share the costs and the benefits from training (Oi (1962)). The optimal sharing rule provides incentives for both parties to invest optimally, and reduces the possibility that human capital may be lost due to quits or layoffs (Hashimoto, (1981)).

There is considerable evidence, however, of firms providing purely general training to their workers (see for example Krueger (1993); Bishop (1996); Barron (1997); Autor (1998); Leuven and Oosterbeek (1999); Acemoglu and Pischke (1998, 1999a)). Reasons advanced in the recent literature as to why it may be in the employer's interest to pay for some general training include: asymmetric information about the extent of workers' training and inherent abilities leading to a wage compression; the fact that training frequently embodies both general and specific skills; stochastic quitting behaviour motivated by non-wage considerations; and trade union compression of wages.

In this paper we follow the recent literature in assuming that there are frictions in the labour market, but our point of departure is that we allow firms to differ in terms of their average quality of labour. In our model, there is a continuum of firms ranging from the best one – that is the firm where the average quality of the workforce is highest – to the worst one. If the productivity of a worker depends on the quality of colleagues, we find that this gives a good firm with a high quality workforce some monopsony power in the labour market. In particular, the best firm has complete monopsony in the market for the services of labour that can only be generated within its ranks. The next-best firm only competes with the best one, and so also has an element of monopsony power, while the worst firm faces a much more competitive labour market. We show that, as a result, the better firms would be willing to pay for the general training of their workers. In this we are supported by the empirical findings of Lynch and Black (1999) who, using US establishment data, show that the incidence of computer and teamwork training is positively related to the average education level within the firm and also positively related to the use of high-performance work practices such as TQ bench-marking and self-managed teams.

Our approach is also relevant to the findings of recent studies, using longitudinal and linked employer-employee surveys, which show that US firms in very narrowly-defined industries are characterized by considerable diversity in productivity and wages (see *inter alia* Haltiwanger, Lane and Spletzer (1999, 2000)). This evidence also suggests that such worker and firm heterogeneity can persist over time. This paper provides a model of how workers' innate ability – observable to the firm although not necessary the survey statistician – can affect the variety of tasks performed within a firm, and consequently the firm's productivity. Since work-related training is endogenous in our model, we also identify an additional avenue through which heterogeneity in the workforces of firms within a narrowly-defined industry can be translated into heterogeneity in firms' output – via differential investment in skills. Our approach is thus relevant not only to labour economics but also to industrial organization, since it provides a theoretical underpinning to the observed link between worker and firm heterogeneity in wages and output within a narrowly-defined industry.

The model developed in this paper shows that good firms may enjoy an element of monopsony power, which makes them willing to spend resources on the training of their workers. The monopsony power emanates from their ability to enhance the productivity of their workers beyond the level they could reach either on their own or by working for alternative employers. The more stimulating and supporting the working environment, the more tasks can be performed and the greater the employer's profits.

In the presence of monopsony power, firms are willing to pay for training that is specific to the task performed but general to the industry. Thus worker- and firm heterogeneity creates both pure economic profits as well as an incentive for firms to invest in the general training of their workers. Our model demonstrates conditions under which firms will invest in the general training of workers, and shows that firms' incentives to invest in general training are increasing in task complexity. Since the degree of monopsony power is increasing with task complexity, firms whose workforces undertake more sophisticated tasks are more willing to finance general training.

I want to say thank you to an astonishing young studio, DreamWorks, for having the courage to hire a bloke from English theatre to do a movie about American suburbia, and then trusting him. So thank you Glenn Williamson, Bob Cooper, Walter Parkes, Laurie MacDonald, Jeffrey Katzenberg, David Geffen, Terry Press, and especially thank you to Steven Spielberg, for handing me this script in the first place, for handing me this [referring to the Oscar] and for being so generous with your wisdom. Thank you to a fantastic, fantastic cast of actors. I will be forever in your debt. Thank you for giving so much of yourselves to this picture, particularly Kevin and Annette and Chris Cooper and the three younger actors – thank you Wes and Thora, Mena. Thank you so much to all the people, all my many, many collaborators on this movie. To Dan Jinks and Bruce Cohen, the producers. To Conrad Hall, you're an artist and I love you. To Tom Newman for your fantastic music. To Naomi Shohan and Julie Weiss, the designers. To my personal heroes, Beth Swofford, you're fabulous, and the equally fabulous Tara Cook. To my editor, Tariq Anwar.

Director Sam Mendes accepting the Oscar for *American Beauty*, 26 March 2000.

Firms exist not only because they are an efficient way of co-ordinating economic activity but also because, as the quotation suggests, they foster collaboration with colleagues, who can supplement our own abilities and compensate for our weaknesses, teach us how to do our tasks better and allow us to specialise in what we do best. Moreover, colleagues can generate a stimulating and challenging working environment and they can impose sanctions on us when we fail to live up to prevailing norms. It follows that the quality of our colleagues matters, and if some firms have better quality workers than others, these provide for a better working environment. It is the objective of this paper to tie these commonplace observations from the personnel economics literature into the human-capital literature, in particular the literature on on-the-job training.

One of the tenets of labour economics is the distinction between general and firm-specific training due to Becker (1964). When training is general, workers must pay for it themselves. Since the market for their services is perfectly competitive, there is no divergence between their wages and marginal product. When the training is firm-specific, however, workers and firms share the costs and the benefits from training (Oi, 1962). The optimal sharing rule provides incentives for both parties to invest optimally, and reduces the possibility that human capital may be lost due to quits or layoffs (Hashimoto, 1981).

However, there is considerable evidence of firms providing purely general training to their workers (see for example Krueger, 1993; Bishop, 1996; Barron, 1997; Autor, 1998; Leuven and Oosterbeek, 1999; Acemoglu and Pischke, 1998, 1999a). Reasons advanced in the recent literature as to why it may be in the employer's interest to pay for some general training include: asymmetric information about the extent of workers' training and inherent abilities leading to a wage compression; the fact that training frequently embodies both general and specific skills; stochastic quitting behaviour motivated by non-wage considerations; and trade union compression of wages.[1]

In this paper we follow the recent literature in assuming that there are frictions in the labour market, but our point of departure is that we allow firms to differ in terms of their average quality of labour and we allow for heterogeneous workers. In our model, there is a continuum of firms ranging from the best – that is the firm where the average quality of the workforce is highest – to the worst. If the productivity of a worker depends on the quality of his colleagues, we find that this gives a good firm monopsony power in the labour market. In particular, the best firm has complete monopsony in the market for the services of labour that can only be generated within its ranks. The next-best firm only competes with the best and so also has an element of monopsony power, while the worst firm faces a much more competitive labour market. We show that, as a result, the better firms would be willing to pay for the general training of their workers. In this we are supported by the empirical findings of Lynch and Black (1999) who, using U.S. establishment data, show that the incidence of computer- and teamwork training is positively related to the average education level within the firm and also positively related to the use of high-performance work practices such as TQ benchmarking and self-managed teams.

Recent studies, using longitudinal and linked employer-employee surveys, show that US firms in very narrowly defined industries are characterised not only by considerable diversity in productivity and wages, but also in organisational structure and size (see *inter alia* Haltiwanger, Lane and Spletzer, 1999, 2000). The evidence also suggests that such worker and firm heterogeneity can persist over time. In our paper, we provide a model of how workers' innate

---

[1] See for example Stevens (1996), Frank and Soskice (1995), Chang and Wang (1996) and Acemoglu and Pischke (1998, 1999a and 1999b), amongst others.

ability – observable to the firm although not necessarily the survey statistician – can affect the variety of tasks performed within a firm, and consequently the firm's productivity. Since work-related training is endogenous in our model, we also identify an additional avenue through which heterogeneity in the workforces of firms within a narrowly defined industry can be translated into heterogeneity in firms' output – via differential investment in skills. Our approach is thus relevant not only to labour economics but also to industrial organisation, since it provides a theoretical underpinning to the observed link between worker and firm heterogeneity in wages and output within a narrowly defined industry.

1. Why do firms pay for general training?

Before advancing the paper's main ideas it is useful to explain why firms are willing to pay for general training within our modelling framework. Four conditions have to be satisfied, building on Acemoglu and Piscke, 1999b. First, there are labour-market frictions that make the productivity of a worker exceed his or her outside option. This economic rent can in the simplest case be due to fixed costs of moving jobs but will be due to elements of monopsony stemming from worker's stochastic preferences – stochastic loyalty to present employer – in our model. Second, the propensity to quit is strictly less than one (in discrete time) since when a firm becomes certain that a worker will leave once trained, it is no longer willing to sponsor his training. Third, the firm's share of the joint surplus due to the labour-market frictions is nonzero – i.e. the workers' bargaining power is less than one – since if the firm cannot capture any of the surplus from a job match it is again not willing to pay for the training. Finally, the marginal effect of training on productivity has to exceed (in absolute terms) the marginal effect of training on wages.

Acemoglu and Pischke (1999a,b) refer to the last condition as “wage compression” but we will call it *absolute wage compression* to distinguish it from more commonly used definitions of the term (Booth and Zoega, 2000). What is implied is that by paying for increased training, a firm augments a worker’s output by more than his or her wage, that is, profits are increasing in training intensity at low levels of training. This condition simply states that firms only invest in general training if they benefit from doing so.

Before moving any further it is useful to define terms in a more precise manner. Assume that Y(ϕ) is output of a worker with training ϕ, W(ϕ) is his wage, P(ϕ) the difference between the two (Y-W) and, finally, p(ϕ) the ratio of the two (Y/W).

Definition 1
Absolute wage compression (decompression) occurs when P’(0) = Y’(0) – W’(0) > 0 (< 0). This implies that profits per worker in absolute terms are increasing (decreasing) in ϕ over some range.

Definition 2
Relative wage compression (decompression) occurs when p’(0) > 0 (< 0). This implies that the ratio of output to wages is increasing (decreasing) in ϕ. By taking logs we get:

$$\frac{d\log p(0)}{d\phi} = \frac{d\log Y(0)}{d\phi} - \frac{d\log W(0)}{d\phi} > 0 \ \ (<0)$$

the derivative of the log difference with respect to ϕ is positive (negative). In this case training increases output proportionately more (less) than wages.

To illustrate the difference between the two definitions, assume that as a result of increased training the productivity and wages of every worker doubles. In this case, we have no change in the ratio of output to wages for any worker – there is neither relative wage compression nor decompression – and there is no change in relative wages or relative productivity levels. However, there is absolute wage compression since the difference between output and wages is now higher for those who have received training compared to those who have not.

Denote a worker’s inherent productivity by $\bar{y}$. Then, following Acemoglu and Pischke (1999a,b), assume that training adds to a worker’s productivity (*Y*) in an additive fashion;

$$Y(\phi) = \bar{y} + f(\phi) \tag{1}$$

where *f(ϕ)* is a strictly concave function. Similarly assume that wages can be denoted by $\bar{w}$ in the absence of training

$$W(\phi) = \bar{w} + w(\phi) \tag{2}$$

where *w(ϕ)* is also a strictly concave function and $\bar{y}$ and $\bar{w}$ can take any value. Assume that the probability that a worker stays on after training is equal to (1-*q*), where *q* is the propensity to quit. This is taken to be a constant and independent of relative wages. Expected profits from a worker receiving training ϕ can then be written as

$$P(\phi) = (1-q)[Y(\phi) - W(\phi)] - c(\phi) = (1-q)[\bar{y} - \bar{w} + f(\phi) - w(\phi)] - c(\phi). \tag{3}$$

Here, absolute wage compression occurs if P'(0) > 0. The equality of the expected marginal profit from training – $(1-q)[Y'(\phi) - W'(\phi)]$ – and the marginal training costs c'(ϕ) – where c(ϕ) is a strictly convex training-cost function and c(0) = 0 – gives the optimal level of training ϕ*:

$$(1-q)[Y'(\phi) - W'(\phi)] = (1-q)[f'(\phi^*) - w'(\phi^*)] = c'(\phi^*) \tag{4}$$

It follows that ϕ* > 0 iff $P'(0) > 0$ which implies $f'(0) > w'(0)$.[2]

Now, instead of assuming that training adds to both productivity and wages in an additive fashion, suppose that it adds in a multiplicative or log-additive way. We change equations (1) and (2) so they become

$$Y(\phi) = \bar{y}\ f(\phi) \tag{5}$$

$$W(\phi) = \bar{w}\ f(\phi) \tag{6}$$

where we have set $f(\phi) = w(\phi)$ to emphasise that $f'(\phi) = w'(\phi)$ at all levels of ϕ. Thus the ratio of output to wages p(ϕ) is a constant and equal to $\bar{y}/\bar{w}$. However, we will see that absolute wage compression is quite common and robust to different assumptions about the system of compensation.

---

[2] Note that in this setup, absolute wage compression implies relative wage compression p'(0) > 0 when $\bar{y} = \bar{w} = 0$ and f(0) = w(0).

Expected profits from employing the worker now become

$$P(\phi) = (1-q)[Y(\phi) - W(\phi)] - c(\phi) \; = \; (1-q)(\bar{y} - \bar{w}) f(\phi) - c(\phi) \tag{7}$$

and the first-order conditions with respect to training are now

$$(1-q)\,(\bar{y} - \bar{w})\, f'(\phi^*) = c'(\phi^*)\ . \tag{8}$$

It again follows that $\phi^* > 0$ iff $P'(0) > 0$ – there is absolute wage compression – but which now only implies $\bar{y} > \bar{w}$ that is labour-market frictions. The firm would benefit from increased training in the absence of relative wage compression, and would be willing to pay for it. It follows that absolute wage compression does not imply relative wage compression: firms may be willing to train in the absence of relative wage compression – relative wage decompression $p'(0) < 0$ not excluded.[3]

Equations (5) and (6) are consistent with a variety of compensation systems. One example is piece rates – or output-based pay where workers get paid $\bar{w}/\bar{y}$ for each unit of output produced. Since piece rates are often used to solve incentive – or moral-hazard – problems they entail a departure from perfect competition and $\bar{y} > \bar{w}$. But, independent of the difference between $\bar{y}$ and $\bar{w}$, if one worker is twice as productive as another, he is also paid twice as much. There is no relative wage compression but equation (8) shows that a compensation system such as piece rates yields absolute wage compression and a positive level of firm-sponsored training as long as $\bar{y} > \bar{w}$.[4]

We are left with the plausible conclusion that for firms to be willing to train – that is for there to be absolute wage compression – we only need an element of monopsony power in the labour market. If firms pay workers less than the value of marginal product for each unit produced, they are also willing to finance some training. Take a pure monopsonist deciding on

---

[3] The difference between (1) and (5) is simple. The first formulation implies that inherent abilities and trained productivity are perfect substitutes making the isoquants in the inherent ability-trained productivity ($\bar{y}$, f(τ)) space downward-sloping lines. The alternative multiplicative formulation implies that they are imperfect substitutes, so that the upper-contour set becomes strictly convex.

wages and employment for workers it has already trained. The post-training profits – this is expected profits once the cost of training has been incurred – can then we written as

$$\tilde{P} = (1-q)[f(\phi)\bar{y}L(w) - wL(w)] \tag{9}$$

where $L = w^a$, a >1, is the labour-supply function for the trained workers. The first-order condition for profit maximisation w.r.t. *w* is

$$af(\phi)\bar{y}w^{a-1} = (1+a)w^a \tag{10}$$

as first pointed out by Joan Robinson (1933). This gives the following solution for the wage;

$$w = \frac{a}{1+a} f(\phi)\bar{y} = f(\phi)\bar{w}, \tag{11}$$

where $\bar{w} = \frac{a}{1+a}\bar{y}$ which is independent of the level of ϕ. Expected profits before training costs are incurred can now be written as

$$P = (1-q)[f(\phi)\bar{y}L^* - f(\phi)\bar{w}L^*] - c(\phi)L^* \tag{12}$$

where $L^* = L(w)$ denotes the optimal level of employment and $\bar{y} > \bar{w}$. Dividing through by L^* and taking the derivative w.r.t. ϕ gives

$$(1-q)(\bar{y} - \bar{w})\, f'(\phi^*) = c'(\phi^*) \tag{13}$$

which is equivalent to equation (8) above. This goes to show that simple monopsony yields equation (8) and hence is all that is needed for firm-sponsored general training.

2. *Our general thesis*

While monopsony in its purest form may be a rare breed, dynamic models of labour demand demonstrate how monopsony power may be pervasive in the labour market. In labour markets where there are substantial costs of hiring and/or firing and asymmetric information and moral-

[4] The use of output-based pay is quite common: Salespeople on a straight commission receive output-based pay. Top executives receive compensation in the form of stocks or stock options. Agricultural workers often earn piece rates – paid for each fruit they pick.

hazard problems are present, employers may resort to their wage policy in order to attract, retain and motivate their workforce.

With employment adjustment costs, the question of whether firms have monopsony power is equivalent to the question of whether the own-wage elasticity of the hiring- (or poaching) function and the quit function is infinite. Empirical estimates suggest that the elasticity of quits and hires w.r.t own wages is a finite number. Campbell (1993) estimated the elasticity of the quit rate with respect to the wage and found that this was equal to –0.96 which is larger than that found in most other U.S. studies (Blau and Kahn, 1981; Viscusi, 1980; Shaw, 1985; Meitzen, 1986; and Light and Ureta, 1992). Estimates of the elasticity of hires tend to be larger, somewhere in the range 0.5 to 4.0 (Card and Krueger, 1995).

We will show how firm- and worker heterogeneity can enhance the level of monopsony defined in terms of the elasticity of quits with respect to own wages. Our central thesis involves demonstrating how the interaction of firm heterogeneity, employment-adjustment costs – in particular training costs – and workers' stochastic preferences can generate a substantial level of monopsony power. The implication is that the level of monopsony power varies across firms – and hence also the level of on-the-job training – in an empirically testable manner.

We assume that workers are heterogeneous in terms of their innate but observable ability, or human capital, which is measured in terms of their endowment of *efficiency units of labour* h_i. Some workers have more to give because of better education, prior work experience, personality traits or intelligence. Worker heterogeneity translates into firm heterogeneity when the average level of human capital is not identical across firms.[5] A high ability level makes a worker have an information set, intuition or drive that is useful to other workers within the firm.

---

[5] Hamilton et al (1999) consider worker heterogeneity in a model in which all workers have the same level of innate general human capital but differ in skill types, in a matching model. They model skill space as a circle around which workers' skills are uniformly distributed, with the density indicating the market size. Firms' job requirements are equally spaced around the circumference. Workers and firms are horizontally differentiated. This contrasts with our approach in which workers are heterogeneous with respect to innate general human capital and homogeneous with respect to skill types, as per the hierarchical job assignment literature surveyed in Sattinger (1993).

Firm heterogeneity in our model – in terms of the average human capital among its workers – affects the range of tasks that can be performed within the firm. To perform more complex tasks, firms need to hire workers of a higher observable innate talent. Only with a good workforce is the necessary synergy in place to accomplish the more difficult tasks. It is here that the employer's monopsony power lies. By definition, there is only one firm capable of performing the most advanced task, two firms capable of performing the next-advanced task, and so on, while all firms can do the simplest one. Thus there is only one buyer of labour trained to do the most difficult task while there may be very many firms employing workers to do the simplest tasks. As the number of firms doing a given task increases, we will show that the level of monopsony power is reduced.

With enhanced monopsony comes a greater willingness to invest in the general training of workers. The optimal level of training intensity is increasing in the difference between output and wages per efficiency unit of labour, and for this reason the determinants of wage setting also affect the level of training. We find the following factors all tend to raise a firm's optimal wage level relative to productivity and hence reduce the level of training: (i) having a low-ability workforce relative to the industry average (making it profitable to poach workers from elsewhere, and hence reduce training); (ii) being one of many employers of a particular type of labour – which also raises the expected benefits from poaching; and (iii) finding high wages effective at recruiting the better workers to raise the average quality of the workforce. We conclude that training will take place in better-than-average firms while bad firms will have underperforming but overpaid workers that are not likely to be trained by their current employer.

Surprisingly – and apparently paradoxically – workers performing the more complex tasks also benefit from their employer's monopsony position in an indirect way. Because of a bigger gap between productivity and wages per efficiency unit of their human capital, they receive more training and for that reason their human capital is augmented to a greater extent. It follows that their wages may be higher on that count in addition to the direct effect of higher productivity.

Thus a ‘good’ firm can offer higher wages to its well-trained workforce while enjoying monopsony profits due to the complexity of tasks performed within its ranks.

3. Model assumptions and the production technology

We rank firms in terms of the quality of their workforce, as will be shown below. The ranking is purely ordinal, and firms are ranked from the best quality workforce ($r = R$) to the worst ($r = 1$). It follows that firms differ only in their average level of human capital:

$$H_R > \ \ldots \ H_r > H_{r-1} \ldots\ldots \ H_1$$

The R-th firm hires workers with the highest average human capital and performs the most complex task, in addition to all simpler ones. Firm R-1 performs all the same tasks apart from the most advanced one, which is the one that distinguishes the best and the next-best firm. Thus, letting M_k denote the number of firms performing the k-th task, it is clear that M_k is decreasing in the complexity of the task performed. More precisely, we can write

$$M_k = R - k + 1 \tag{14}$$

where R is the total number of firms and k is the number of tasks performed at a particular firm. Equation (14) implies that there is one firm doing the most advanced task $k = R$, two firms doing the second most advanced, and R firms doing the most simple task.

Production is a linear function of labour (measured in efficiency units) such that an efficiency unit of labour devoted to a particular task produces $\bar{y}$ units of output. The firm can augment a worker’s efficiency units through on-the-job training. The cost of providing the worker with training ϕ is described by the strictly convex function c(ϕ);

$$c(\phi), \quad c'(\phi)>0, \quad c''(\phi) > 0, \quad c(0)= 0 .$$

The training function f(ϕ) is strictly concave;

$$f(\phi), \quad f'(\phi)>0, \quad f''(\phi) < 0, \quad f(0)= 1.$$

The profits of the best firm (performing *R* tasks) P_R are the sum of value added $\overline{y} - \overline{w}$ – the difference between output and wages per efficiency unit of labour – net of training costs coming from all efficiency units performing each of the *R* tasks, and can be written as

$$P_R = \sum_{r=1}^{R} \sum_{i=1}^{n_r} n_r f(\phi_r) h_i (\overline{y}_r - \overline{w}_r) - \sum_{r=1}^{R} n_r c(\phi_r) \tag{15}$$

where n_r is the number of workers doing task r and h_i is the human capital of worker i. The *i*-th unit of human capital would only produce $\overline{y}_1$ in the "worst" firm where the average quality of his colleagues is H_1 – performing the most mundane tasks and receiving wage $\overline{w}_1$ – but it would produce a higher level $\overline{y}_R$ in the best firm doing the most sophisticated task at wages $\overline{w}_R$ where the average quality is H_R.

Because of constant-returns-to-scale – one efficiency unit of labour devoted to task i produces $\overline{y}_i$ units of output – the firm's decisions about employment, on the one hand, and wages and training intensity, on the other hand, are separable. The number and quality of workers in each firm is initially calculated for given wages and training intensity. They receive on-the-job training in work-related skills that are general to all firms in the skilled sector but specific to the task they are being trained for. The degree of training intensity then determines the extent to which post-training productivity and wages are augmented.

Once trained, workers may choose either to stay with the firm that provided the training and produce, or quit to work in other firms in the skilled sector. Workers do not leave the skilled sector. The probability of retaining current workers – captured by the quit function q – and the probability of poaching workers trained by other firms – captured by the poaching function p – is a function of the wage differential between firms, amongst other factors. The two functions, and their first derivatives, determine the optimal level of wages and hence also the optimal level of training.

4. ***The allocation of labour across tasks***

We now consider both total employment (human capital) in each firm devoted to each task, and the optimal allocation of workers with different ability levels across tasks. We initially consider how employment and human capital are allocated for given optimally chosen levels of wages and training intensity. We assume that each firm is faced with a finite number of workers with each skill level. This captures the essence of a labour market in which workers show up randomly. Firms' task heterogeneity is determined by the ability of workers who show up. Not all workers who turn up are selected due to the cost of training them. The firm selects the best quality workers and then begins training and production.

First we consider *the number of workers doing each task,* n_k, where the subscript denotes the k-th task. Workers are given training on-the-job which is specific to their task. The training intensity is denoted by ϕ and the workers augment their inherent human capital h in a multiplicative way as in equation (5) above. Thus a worker trained for task k will have $f(\phi_k)h_i$ efficiency units of labour and he will be able to produce $f(\phi_k)h_i\bar{y}_k$ units of output.[6] Wages per efficiency unit of labour are denoted by $\bar{w}_k$ so the wages of a worker with human capital h_i working on task k are $f(\phi_k)h_i\bar{w}_k$ as in equation (6) above.

The r-th firm's chooses the number of workers doing the k-th task, n_k , by maximising

$$P_k = \sum_{i=1}^{n_k} \left\{ (1-q) f(\phi_k)(\bar{y}_k - \bar{w}_k)h_i - c(\phi_k) \right\} \tag{16}$$

subject to the number of available workers of each skill type. Note that remuneration and marginal productivity vary in accordance with a worker's innate productivity, h.

For the r-th firm, the number of workers trained to do task k has to be such that the following equation be satisfied for every worker trained:

$$(1-q) f(\phi_k^*)(\bar{y}_k - \bar{w}_k^*)h_i - c(\phi_k^*) \geq 0 \tag{17}$$

[6] Note that the relative productivity of all workers is not affected by their training because the training adds to a worker's human capital in a multiplicative way.

The left-hand side of the equation has the optimal values of ϕ and $\overline{w}$. The equation shows that the benefit from hiring worker *i* must be nonnegative. The firm starts with the most able person who has presented herself – assuming her human capital is high enough to make the benefit from training be positive. As it moves to less able workers, the benefit from training is falling in the number of workers trained since their ability level h_i is falling until the equality in equation (17) holds or the net benefit from training for task k becomes lower than the benefit from training the same worker for some other task.

Having considered the number of workers doing each task, we now take a look at *the allocation of workers across tasks.* Should the firm use a few of the best workers for the most difficult task – task *r* – or should we use many less able workers for that task? We will show that the best workers should be allocated to the most advanced tasks within the firm. Assume that there is a worker with ability level h_j doing task *r* – the most sophisticated within firm *r* – and another doing the least advanced task – task 1 – with a higher ability level h_i; $h_i > h_j$. If the firm could relocate them at zero cost so that the latter would be trained to do task r and the former to do task 1, the benefit to the firm would be the following since total training costs would be unchanged;

$$(1-q)\left[f(\phi_r)(\overline{y}_r-\overline{w}_r)-f(\phi_1)(\overline{y}_1-\overline{w}_1)\right]\left(h_i \overset{+}{-} h_j\right)>0 \tag{18}$$

which is positive iff

$$f(\phi_r)(\overline{y}_r-\overline{w}_r)\geq f(\phi_1)(\overline{y}_1-\overline{w}_1)$$

that is if the firm benefits more from the employment of a worker doing the sophisticated task. We will show this to be the case because of a greater degree of monopsony power for task *r* and a higher level of training, $\phi_r > \phi_1$. It will then become clear that the best workers would be allocated to the most advanced task and the worst to the least advanced one. At the margin between task *k* and task *k-1*, the following condition should hold:

$$(1-q)f(\phi_k)(\overline{y}_k-\overline{w}_k)h_i-c(\phi_k)=(1-q)f(\phi_{k-1})(\overline{y}_{k-1}-\overline{w}_{k-1})h_i-c(\phi_{k-1})\ \geq\ 0 \qquad (19)$$

The net benefit from allocating a new worker i – with human capital h_i – to task *k* and task *k-1* is the same.

Note that the inequality in equation (19) comes about when a firm has gone down the line and trained workers for the different tasks, starting with the most able worker being assigned to the most sophisticated task and finishing with the least able worker being trained for the most basic task. If workers who are at least as able as this marginal (minimum ability) worker – this is the pool of desired workers – are in short supply, the firm cannot hire as many workers for each task as it desires. (Recall our assumption that there is a finite number of each innate skill/ability type who present themselves for employment.) Faced with this constraint, the best the firm can do is to allocate workers efficiently across tasks, leaving unfilled vacancies throughout the firm.

5. *Monopsony and On-the-job Training*

We allow firms to augment each worker's human capital through on-the-job training. Each firm gives workers training that is general to one of the tasks performed within its ranks. Thus a worker trained for task *r* (the most sophisticated task in the *r*-th firm) can take this knowledge with her to work on that same task in any other firm. But the training is specific to task *r* and cannot be used to do any other task. Thus the training will be useless in firm *r*-1 and below since the *r*-th task is not performed there, but will be of value to all firms with task complexity greater than *r*.

We now return to equation (14). Firm *R* has is the only firm in the market for labour trained to perform its most sophisticated task and produce the last $\overline{y}_R$ units of output (per efficiency unit of labour). These units can only be produced through the most difficult task, which requires the co-operation of very able workers. Firm *R* then only competes with one other firm for the services of labour to produce the $\overline{y}_{R-1}$ units of output. The degree of labour-market competition is then rising until it reaches its maximum for the first $\bar{y}_1$ units. As noted above, we view these different lumps of output as resulting from increasingly sophisticated tasks performed within firms.

What determines labour mobility in this model? Following Salop (1979) we assume frictions or 'stickiness' in the movement of workers between firms, so that the retention probability is a function not only of the wage differential for each task performed but also contains a stochastic component. The latter reflects the fact that sometimes workers decide to leave their current employer and find another job for non-wage reasons. Workers' preferences are drawn from the same probability distribution but each worker has a different draw from the distribution, hence different preferences and may therefore like to work in a different workplace. This is the key labour-market friction in our model. Most importantly, we will show that the quitting/poaching probability is affected by the number of other firms performing the task for which a worker has been trained.

We can show these ideas more rigorously as follows. Suppose that an individual joins a new firm – is successfully poached – iff $\mu_i \geq (\overline{w} - \overline{w}^A)$, where μ denotes the non-pecuniary benefit from joining a given firm, $\overline{w}$ is his current wage and $\overline{w}^A$ is the average wage elsewhere. Assume cov(μ_i ,h_i) = 0 since there is no reason to suppose that innate ability and attitudes to location or the boss should be correlated. Suppose that μ_i is uniformly distributed as $[-\theta L_k,\ \theta]$ representing the fact that the non-pecuniary benefits from joining a given firm are decreasing in the total number of firms because there is more choice in terms of the non-pecuniary aspects of a job. The density $f(\mu)$ is then given by $1/[\theta(L_k+1)]$. The probability of a worker joining a firm is then given by the following equation

$$prob(\mu \geq \overline{w} - \overline{w}^A) = \frac{\theta - \left(\overline{w} - \overline{w}^A\right)}{\theta\left(L_k + 1\right)} = p(\overline{w} - \bar{\overline{w}}^A,\ \bar{L}_k) \tag{20}$$

which is equal to $1/M_k$ in symmetric equilibrium where $\overline{w} = \overline{w}^A$. The probability of successfully poaching a worker is decreasing in the number of firms and in own wages relative to wages elsewhere $\overline{w} - \overline{w}^A$.

The probability that a worker quits is equal to the probability that he be poached by any of the L_k alternative employers. It follows that the probability of quitting can be written as follows

$$q(\overline{w} - \overline{w}^A,\ \bar{L}_k) \ = \ p(\overline{w} - \overline{w}^A,\ \bar{L}_k)L_k \ = \frac{\theta L_k - \left(\overline{w} - \overline{w}^A\right)L_k}{\theta\left(L_k + 1\right)} \tag{20'}$$

where $\overline{w}$ is the worker's current wage as before and $\overline{w}^A$ is the wage being offered elsewhere. It follows that the quit rate is increasing in the number of firms L_k and decreasing in the worker's own wage.

We now consider the firm's profit function for task k in equation (16). Recall that P_k denotes expected profits from letting n_k able workers perform task k and produce $\sum_{i=1}^{n_k} f\left(\phi_k\right)h_i \overline{y}_k$ units of output, where $\overline{y}_k$ is output per efficiency unit of labour doing task k and wages per efficiency unit of labour are denoted by $\overline{w}_K$. The r-th firm's maximisation problem can now be written as follows;

$$\max_{w_k, \phi_k} P_k = [1 - q(x_k, L_k)]\, f(\phi_k)(\overline{y}_k - \overline{w}_k)\overline{h} n_k + p(x_k, L_k)\, f(\phi_k^A)(\overline{y}_k - \overline{w}_k)\overline{h} L_k n_k^A - n_k c\left(\phi_k\right) \tag{21}$$

where $x_k = \overline{w} - \overline{w}^A$, $\overline{h}$ is the average level of human capital doing task k and ϕ_k^A is the average level of training in other firms.[7] The first term on the left-hand side gives expected profits from employing retained workers trained in this firm. The second term on the left-hand side gives the expected profits from workers trained by all other firms performing this task but who have been poached by this firm. There are n_k^A such workers at each firm. Finally the right-hand side has the cost of training given by the strictly convex function c(ϕ). To make things simple, we assume that the 'poachees' are of the same ability level as worker *i*. However, we think that relaxing this

[7] Instead of writing the profit function using the average human capital of workers, we could have written one profit function for each worker employed doing task k. This would not affect our results as long as any prospective

assumption is sufficiently important to justify a separate section that comes at the end of the paper.

The first-order conditions for the level of training per efficiency unit for task *k* follows:

$$\left[1-q\left(x_k^*, L_k\right)\right] f'\left(\phi_k^*\right)\left(\bar{y}_k - \bar{w}_k^*\right)\bar{h} = c'\left(\phi_k^*\right) \text{ , } k = 1,2 \ldots R \tag{22}$$

The left-hand side has the expected marginal benefit from increased training while the right-hand side has the marginal cost. We now have our first proposition.[8]

> *Proposition 1*: The amount of training set by the firm is always positive, but is less than the efficient level given by ϕ^{**}.

Proof: According to equation (22) there will be under-investment in training because the firm discounts the marginal benefit from training $f'(\phi^*)$ by (1-q). In contrast, the first-best level ϕ^{**} is set to maximise the value of total output produced by *all* trained workers – both those retained by any firm, plus those who quit to work in other firms – less the costs to society of training, that is ϕ^{**} solves $f'\left(\phi_k^{**}\right)\left(\bar{y}_k - \bar{w}_k\right)\bar{h} = c'\left(\phi_k^{**}\right)$, and hence $\phi^* < \phi^{**}$.

Note that the level of training is an increasing function of a worker's inherent human capital. It follows that workers doing task k will not all receive the same level of on-the-job training as they will not all have the same level of inherent human capital.

The wage is determined by the following equation;

$$[p' L_k n_k^A f(\phi_k^A) - q' n_k f(\phi_k^*)](\bar{y}_k - \bar{w}_k^*) = pL_k n_k^A f(\phi_k^A) + (1-q) n_k f(\phi_k^*) \tag{23}$$

where the left-hand side has the marginal benefit of raising wages – in the form of fewer quits and more poachees – and the right-hand side has the marginal cost – in terms of a higher wage bill. The two equations (22) and (23) are separable in ϕ and $\bar{w}$. Equation (24) gives wages paid *per efficiency unit* for task *k*:

$$\bar{w}_k = \bar{y}_k - \frac{pf\left(\phi_k^A\right)L_k\, n_k^A / n_k + (1-q)\, f\left(\phi_k^*\right)}{p' f\left(\phi_k^A\right)L_k n_k^A / n_k - q' f\left(\phi_k^*\right)} = \bar{y}_k - \Psi_k \tag{24}$$

---

poachee is of equal quality to the relevant worker doing task k.

We should note at this stage that the second-order condition for maximum is satisfied, in that the Hessian matrix is negative definite.[9] Notice also that equation (24) does not give the actual wage received by a worker, as this is given by $f(\phi_k)h_i\overline{w}_k$.

We now come to the second proposition:

> *Proposition 2*: The smaller is the number of firms performing a given complexity of task, the larger is the gap between wages per efficiency unit and productivity. Hence the more complicated the task, the larger is this gap.

Proof: The poaching function (20) has the following properties

$$\lim_{L_k \to \infty} p(\cdot) = 0, \quad \lim_{L_k \to 0} p(\cdot) = 1, \qquad \lim_{L_k \to \infty} p'(\cdot) = 0, \quad \lim_{L_k \to 0} p'(\cdot) = \frac{1}{\theta}$$

in symmetric equilibrium. Similarly we find from (20') that

$$\lim_{L_k \to \infty} q(\cdot) = 1, \quad \lim_{L_k \to 0} q(\cdot) = 0, \qquad \lim_{L_k \to \infty} q'(\cdot) = -\frac{1}{\theta}, \quad \lim_{L_k \to 0} q'(\cdot) = 0$$

also in symmetric equilibrium. The term Ψ_k in equation (24) is a measure the degree of monopsony, defined as the difference between the level of productivity and wages per efficiency unit of labour. It follows that this term goes to zero when the number of alternative firms goes to infinity, and converges to a positive number when L_k goes to zero:

$$\lim_{L_k \to \infty} \Psi_k = 0 \qquad \text{and} \qquad \lim_{L_k \to 0} \Psi_k = \frac{f(\phi_k^A)L_k n_k^A / n_k + f(\phi_k^*)}{\frac{1}{\theta} f(\phi_k^A) L_k\, n_k^A / n_k} > 0$$

By construction, L_k is a function of how advanced is the relevant production task as shown in the figure below. At k = 1 (the most basic task), L_k has its maximum value (R-1) and the wage is close to productivity. At k = R the wage is at its lowest with only one firm in the market for trained labour doing the most sophisticated task (L_k =0).

[FIGURE 1]

---

[8] See also Stevens (1996) and Booth and Zoega (1999).

It follows that the more advanced is the relevant task, the higher are profits from performing the task within the firm: all firms can perform the simplest tasks, hence the degree of competition is highest for such tasks. But as firms become 'better', they gain elements of monopsony power in the market for labour trained to do tasks that are only feasible within their ranks – labour that can only produce output by enjoying the company of the high-quality workers currently employed.

We now consider the financing of this general training, where the firm's choice of training intensity is denoted by ϕ^*. This is our third proposition.

Proposition 3:

(i) In a frictional labour market in which the firm gets rent from the employment relation and there is some probability that the relation will continue, the firm will invest a positive amount in general training, that is $\phi^* > 0$.

(ii) The more advanced the task, the more firm-financed general training will be paid for by the firm.

Proof 3(i): Expected profits from a retained worker are $(1-q)(\bar{y}-\bar{w})f(\phi)\bar{h}$, and note that its derivative is $(1-q)(\bar{y}-w)f'(\phi)\bar{h}$. It follows that $P'(0)>0$ iff $\bar{y} > \bar{w}$. The proof of *3(ii)* follows from Propositions 2 and 3(i): The fewer the firms, the larger is $(\bar{y}-\bar{w})$, and hence the bigger is ϕ^*).

The firm is here willing to finance its chosen level of general training because it augments workers' productivity in a multiplicative way. As shown in Booth and Zoega (2000), we do not need to rely on the wage compression arguments adopted by Acemoglu and Pischke (1999a,b) to generate firm-financed general training. From equation (24) it follows that the larger is Ψ – the difference between output and wages – the greater is the level of ϕ. It follows that when we

[9] We consider a representative firm only, since the equilibrium is symmetric as shown in the appendix.

move to more sophisticated tasks, the firm finances more extensive training. Figure 2 shows the optimal level of training.

[FIGURE 2]

Now looking back at equation (22) we find that workers doing the most sophisticated tasks within the firm receive more training for two reasons. In addition to the monopsony element emphasised in Proposition 3 these workers also have greater inherent human capital *h* and receive more training for that reason.

5. Wages, training and complexity

Each worker performs one of the tasks carried out within the firm. In a symmetric equilibrium (when wages per efficiency unit of labour are equalised for each task across firms) the wage of a worker i performing task k w_{ik} is given by

$$w_{ik} = f(\phi_k)h_i\overline{w}_k = f(\phi_k)h_i(\overline{y}_k - \Psi_k). \tag{25}$$

While a worker's inherent human capital h_i is exogenous to our model, we have modelled the determination of both training and wages per efficiency unit. Wages received by worker i are a positive function of all three factors. The implied distribution of wages across tasks and workers depends on the interplay between these three factors.

Across tasks, we have shown that the more complex is the task, the greater the difference between output and wages per efficiency unit of labour, Ψ. This difference then makes firms choose a higher level of training. Taking the total differential of equation (22) gives,

$$\frac{d\phi}{d\Psi_k} = \frac{(1-q)\overline{h}\, f'(\phi)}{c''(\phi)-(1-q)f''(\phi)\overline{h}\Psi_k} > 0. \tag{26}$$

Not surprisingly, the effect of monopsony on training is decreasing in the convexity of training costs and the concavity of the training function f and increasing in the retention rate 1-*q* and the inherent human capital of workers. It follows that if both the cost function c and the training function f are close to linear, the effect on training can be very significant, especially for the better workers.

Finally, the effect on ϕ is decreasing in the level of Ψ_k as shown in Figure 2. The effect on human capital through training is then further decreasing in Ψ_k due to the concavity of the training function f. Depending on the value of the model's parameters, the stage is set for the hump-shaped relationship between wages (per worker) and task complexity shown in Figure 3.

[FIGURE 3]

Initially the indirect effect of monopsony through training on wages dominates its direct effect so that wages are increasing in task complexity. The indirect effect may then diminish more rapidly – although this depends on the parameters of the model – so that the direct effect starts dominating and wages start falling in the complexity of the task.[10] In summary, wages are increasing in task complexity because of increased training and also higher output per efficiency unit of labour $\bar{y}_k$. Wages are decreasing in complexity because of higher monopsony power Ψ_k.

Lucrative consulting opportunities among university professors may provide one example of the relevance of a hump-shaped relationship between complexity and wages. By performing less advanced tasks as consultants to both private companies and government agencies, professors often earn salaries that are significantly higher than those paid for the often more demanding tasks performed on university campuses. However, universities tend to be more generous when it comes to sabbaticals and other opportunities for self improvement, not surprisingly in light of

[10] The hump-shaped relationship depends crucially on productivity not increasing too rapidly in task complexity ($\bar{y}_k$ exceeding $\bar{y}_{k-1}$ by too much) and the shape of the training and cost functions being such that the second derivative of ϕ with respect to Ψ be sufficiently negative.

our model.[11]

Across workers performing the same task, workers with higher human capital also receive higher wages. The reason is twofold: First, they embody more units of human capital to start with. Second, because their endowment of human capital is greater, they also receive more training which raises their wage further. In a university department, the more productive professors would both be paid more due to their initially higher creativity but would also be given less teaching and more opportunities to increase their skills further.

6. *Poaching across tasks and the quest for a monopsony position*

So far, we have left out of our analysis the possibility that a firm may want to offer high wages in order to raid other firms of individuals who are of a higher ability than their current workforce. In particular, we wrote equation (21) as if poachees would be identical to existing workers. But now we want to relax this assumption and introduce asymmetries between current and prospective workers.

When a firm has trained all eligible workers from its own pool of applicants, it may pay to offer high wages in order to attract more able workers doing the same task at other firms – hence not in need of new induction skills – or more able workers doing different tasks (i.e. less sophisticated tasks at inferior firms) but needing some initial retraining.

We first describe poaching within tasks. The maximisation problem becomes

[11] The relationship between complexity – hence labour-market competition – and real wages is reminiscent of the findings of Calmfors and Driffill (1988) who found a hump-shaped relationship between real wages and the centralisation of wage bargaining. We can view centralisation among employers as enhanced monopsony power and thus directly compare our results to theirs. On the employer side, they find that "the potential for a price rise is larger, the more sectors that are simultaneously affected by wage increases and, hence, money profits are less adversely affected. However, this is accompanied by a stronger effect on the aggregate price level, hence less of an increase in *real* profits. The first effect tends to raise wages, the second to reduce them."[11] As a result, increased centralisation among employers leads initially to rising real wages and then eventually to falling real wages. We have found that as the level of monopsony in the labour market is increased, the real wage may initially rise due to the increased training of workers but eventually falls when training runs into diminishing returns and monopsony power is further increased.

$$\underset{w,\phi}{Max}\ P_k = (1-q(x_k))g(\phi_k)(\bar{y}_k - \bar{w}_k)\bar{h}n_K + p(x_k)L_k n_k^A g(\phi_k^A)[\bar{y}_k - \bar{w}_k]\bar{h}^A - c(\phi)n_k \qquad (27)$$

where $\bar{h}^A$ is the average ability level for workers doing task k at other firms. The FOC for ϕ is as before equation (22).[12] The FOC for wages now become

$$\bar{w}_k = \bar{y}_k - \frac{pf(\phi_k^A)L_k n_k^A \bar{h}^A / n_k \bar{h} + (1-q)f(\phi_k^*)}{p' f(\phi_k^A)L_k n_k^A \bar{h}^A / n_k \bar{h} - q' f(\phi_k^*)} = \bar{y}_k - \hat{\Psi}_k . \qquad (24')$$

Wages per efficiency unit of labour are a positive function of $\bar{h}^A / \bar{h}$;

$$\frac{d\hat{\Psi}}{d(\bar{h}^A/\bar{h})} = -[pq' - p'(1-q)] \frac{f(\phi_k^A)f(\phi_k)L_k \, n_k^A / n_k}{[p' f(\phi_k^A)L_k n_k^A \bar{h}^A / n_k \bar{h} - q' f(\phi_k^*)]^2} > 0 . \qquad (28)$$

The effect is increasing in the number of alternative firms L_k and in the relative size of these firms n_k^A / n_k . It follows that a worker – or a group of workers doing the same task – gets a higher wage (per efficiency unit of labour that he embodies) if he is of low ability relative to those doing the same task in other firms. Bad workers get more than they "deserve" because their employer is hoping to attract someone else. There are more efficiency units of labour to be poached and the marginal benefit of raising wages (at any given wage level) is higher.

It follows that a 'good' firm that can allocate high ability workers to a complicated task k enjoys higher monopsony profits for two reasons. First, there is a limited number of firms doing task k as discussed previously. Second, the firm is disproportionately large because of the superior workforce that it allocates to task k. For both reasons, the firm would pay lower wages per efficiency unit and have higher profits and, most importantly, it would train its employees more intensively according to proposition 3. The last effect benefits workers and they can receive higher wages despite the greater monopsony element.

---

[12] Note that we may no longer have a symmetric equilibrium if the firms differ in the abilities of workers doing task k. In particular ϕ_k does not have to equal ϕ_k^A which complicates the equations slightly.

We now turn to the possibility of poaching a worker who is currently doing a different task at another firm. In order to simplify the exposition, we assume that there is no poaching of workers doing the same task at other firms. The analysis has to be modified in two different ways. First, the key difference from the previous case lies in the need to give workers induction training for task k. If the firm is successful at poaching a worker from another firm, the worker will then only be trained if his ability level is sufficiently high to warrant giving him the training. The following condition has to be satisfied for a worker with ability level h_j to be profitably trained:

$$(1-q)\, f(\phi_k^*)(\bar{y}_k - \bar{w}_k)h_j - c\left(\phi_k^*\right) \geq 0 \tag{29}$$

It is clear that there is always a place for an 'excellent' worker. But if a worker falls below this threshold, he will be turned away. Now when firms set wages with the intention of both retaining their current workers and poaching workers from other firms who are currently doing other tasks, they have to take into account the probability γ that possible poachees will be above this threshold ability level. The second difference lies in the observation that workers who join a new firm to do a different task face both different wages and training levels with their new employer. The relevant wage differential becomes

$$x_k = f(\phi_k)\bar{w}_k - f(\phi_r)\bar{w}_r \tag{30}$$

when quitting task r in one firm to do task k in another.

While equation (22) is again unchanged, the FOC for wages will now look as follows**;**

$$\bar{w}_k = \bar{y}_k - \frac{p\gamma f\left(\phi_k^A\right)L_k n_k^A \bar{h}^{A'} / n_k \bar{h} + (1-q)\, f(\phi_k)}{p'\gamma f\left(\phi_k^A\right)L_k\, n_k^A \bar{h}^{A'} / n_k \bar{h} - q'\, f(\phi_k)} = \bar{y}_k - \hat{\hat{\Psi}}_k \tag{24''}$$

where $\bar{h}^{A'}$ is the average human capital of those workers falling above the threshold. Note that γ appears with L_k and $\bar{h}^{A'}$ – they all have the same effect of raising wages towards the level of productivity. If it is easy to get able workers, this makes poaching more lucrative – just as having more and better potential poachees – and the optimal wage is higher.

Finally, a firm may want to change its ranking by poaching high-ability workers who make it possible for a more sophisticated task to be performed. But note that this usually requires more than one worker. Assume that *m* high-ability workers are needed to change the ranking of firm *r* and that the probability that a given worker falls into this group is η. The probability that *m* such high-ability workers will be poached is equal to η^m, which can be a very low number since $\eta < 1$. If fewer than *m* workers are poached, the analysis above applies. But assume that the firm will only hire these new workers if their number is at least equal to *m* and that they will be used to do task r+1. Equation (21) will now have the following form;

$$\underset{w,\phi}{Max}\ P_{r+1} = \eta^m\ p(x_{r+1})^m L_{r+1} g\left(\phi_{r+1}^{A'}\right)\left(\overline{y}_{r+1} - \overline{w}_{r+1}\right)h^{A'} \tag{31}$$

where $x_{r+1} = \overline{w}_r - \overline{w}_r^A$ if poaching within tasks and given by equation (30) if poaching between tasks. The FOC with respect to wages will now read:

$$\overline{w}_{r+1} = \overline{y}_{r+1} - \frac{p(x_{r+1})}{mp'(x_{r+1})} = \overline{y}_{r+1} - \hat{\hat{\hat{\Psi}}}_{r+1} \tag{24'''}$$

An increase in *m* will raise $\overline{w}_{r+1}$. If we need many workers the sensible thing is to offer a high wage. The same applies if high wages are effective at attracting workers (p' is large). But if the workers come anyway (p is large), the optimal wage is low.

We can summarise the discussion in this section by stating the conditions that are conducive to high wages:

- There are many other firms with workers who have been trained to do the task in question.
- Workers at other firms are of a high ability and likely to be of sufficient ability for the firm to want to give them training.
- The firm needs a large number of new workers to create an environment for a new sophisticated task to be performed within its ranks, and high wages are an effective way of poaching workers.

It then follows from equation (22) that all of these conditions work against firm-sponsored general training. A firm which is one of many in the industry with a below-average workforce but that aspires to greatness through attractive wage offers is not going to expend resources training its underperforming and overpaid staff. Taking an example from the academic labour market, a below-average research department would maintain relatively competitive wages in the hope of attracting good people from other department, but would not be generous to its current staff in terms of the frequency or length of sabbaticals or in terms of research budgets.

In contrast a good and well established firm that offers its much-better-than-average workforce the opportunity of performing tasks rarely performed anywhere else can both pay less relative to productivity and then also offer more extensive training programmes. Workers may also benefit from their high average quality as they receive more training and for that reason possibly higher wages. An above-average research department would thus enjoy the best of all world: high profits for the university, high wages for staff and generous provisions for sabbaticals and other skill-enhancing opportunities.

7. Conclusions

We have found that good firms may enjoy an element of monopsony power which makes them willing to spend resources on the training of their workers. The monopsony power emanates from their ability to enhance the productivity of their workers beyond the level they could reach either on their own or by working for alternative employers. The more stimulating and supporting the working environment, the more tasks can be performed and the greater the employer’s profits.

In the presence of monopsony power, firms are willing to pay for training which is specific to the task performed but general to the industry. Thus worker- and firm heterogeneity creates both pure economic profits as well as an incentive for firms to invest in the general training of their workers. Our model demonstrates conditions under which firms will invest in the general training of workers, and shows that firms' incentives to invest in general training are increasing in task complexity. Since the degree of monopsony power is increasing with task complexity, firms whose workforces undertake more sophisticated tasks are more willing to finance general training.

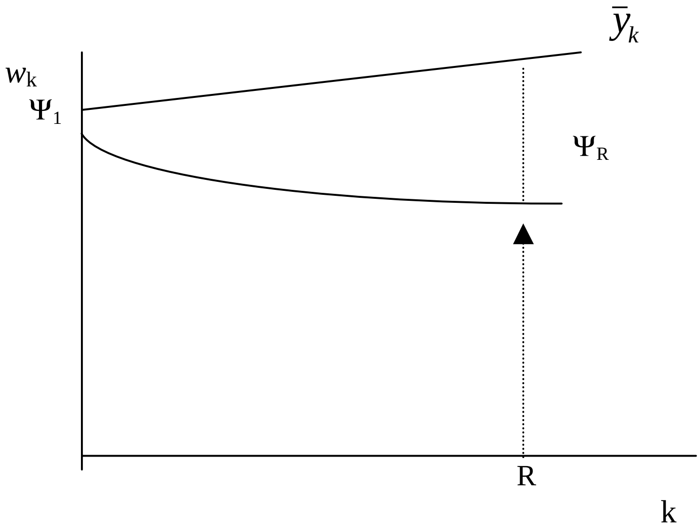

Figure 1 Wages, monopsony power and firm ranking

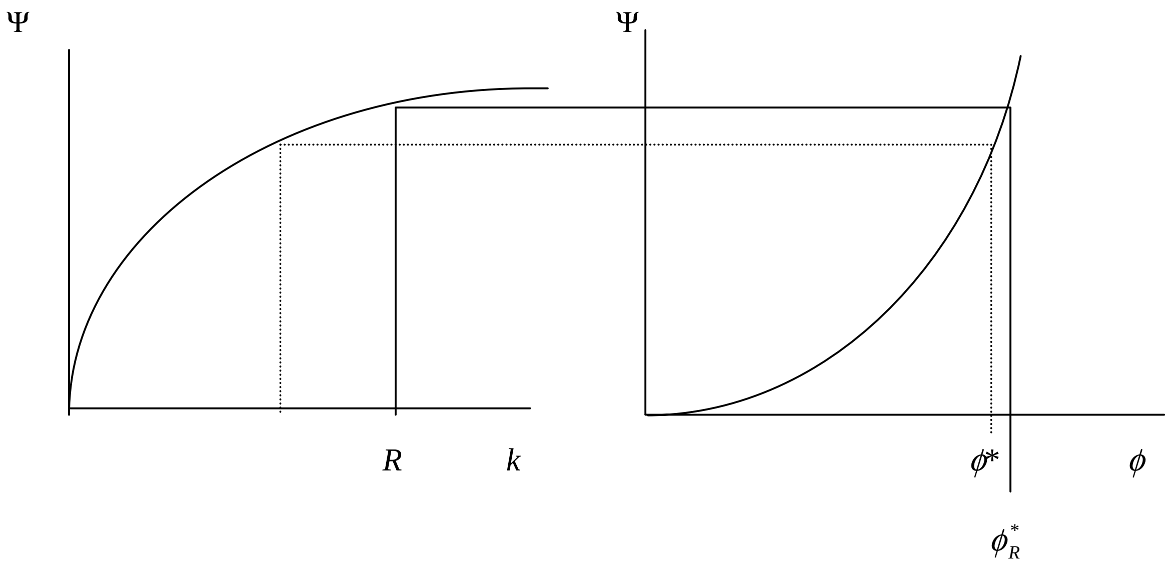

Figure 2 Training and the level of monopsony

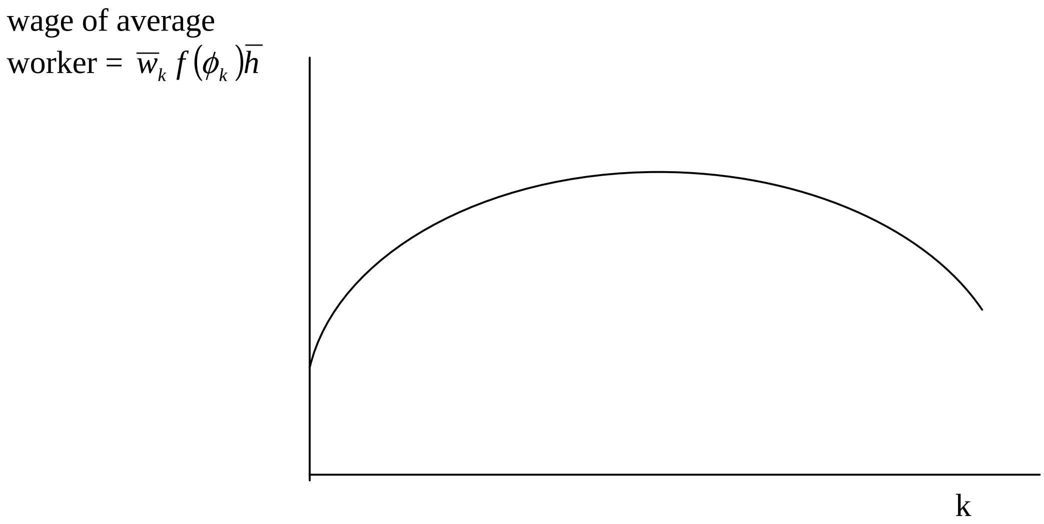

Figure 3 Wages and the level of monopsony

Appendix

Proof of Existence and Uniqueness of Symmetric Equilibrium for the Industry

For simplicity, the proof is done for the two-firm case. Profits for firm 1 can written as:

$$P_{1k} = (1-q_{1k})\, f(\phi_{1k})(\bar{y}_k - \bar{w}_{1k})\bar{h}\, n_{1k} + p_{1k}\, f(\phi_{2k})(\bar{y}_k - \bar{w}_{1k})\bar{h}\, n_{2k} - n_{1k} c(\phi_1) \qquad \text{(A1)}$$

Maximisation of (A1) w.r.t. training intensity ϕ_1 and wages w_1 respectively gives:

$$(1-q_{1k})\, f'(\phi_{1k})(\bar{y}_k - \bar{w}_{1k})\bar{h} = c'(\phi_1) \qquad \text{(A2)}$$

$$[p'_{1k}\, n_{2k}\, f(\phi_{2k}) - q'_{1k}\, n_{1k}\, f(\phi_{1k})](\bar{y}_k - \bar{w}_{1k}) = p_{1k} n_{2k}\, f(\phi_{2k}) + (1-q_{1k}) n_{1k}\, f(\phi_{1k}) \qquad \text{(A3)}$$

The analogous expected profit equation for firm 2 is:

$$P_{2k} = (1-q_{2k})\, f(\phi_{2k})(\bar{y}_k - \bar{w}_{2k})\bar{h}\, n_{2k} + p_{2k}\, f(\phi_{1k})(\bar{y}_k - \bar{w}_{2k})\bar{h}\, n_{1k} - n_2 c(\phi_2) \qquad \text{(A1b)}$$

The first order conditions (FOC) of firm 2 are:

$$(1-q_{2k})\, f'(\phi_{2k})(\bar{y}_k - \bar{w}_{2k})\bar{h} = c'(\phi_2) \qquad \text{(A2b)}$$

$$[p'_{2k}\, n_{1k}\, f(\phi_{1k}) - q'_{k2}\, n_{2k}\, f(\phi_{2k})](\bar{y}_k - \bar{w}_{2k}) = p_{2k} n_{1k}\, f(\phi_{1k}) + (1-q_{2k}) n_{2k}\, f(\phi_{2k}) \qquad \text{(A3b)}$$

Dropping the *k* subscripts for expositional ease, we now subtract (A3b) from (A3) to obtain after some manipulation the following function, assumed to be continuous from continuity of the underlying functions:

$$\Phi(x) = (w_1 - w_2) + \left[\frac{p_1 f(\phi_1)(n_1/n) + (1-q_2) f(\phi_2)}{p_1' f(\phi_1)(n_1/n) - q_2' f(\phi_2)} \right] - \left[\frac{p_2 f(\phi_2)(n_2/n) + (1-q_1) f(\phi_1)}{p_2' f(\phi_2)(n_2/n) - q_1' f(\phi_1)} \right] = 0$$

(A4)

We now show <u>existence</u> of an industry equilibrium by showing that equation (A4) only holds for

$x = (w_1 - w_2) = 0$, that is $w_1 = w_2$.

Case 1: $x \rightarrow \infty$

Here the first term in (A4) is ∞. Clearly $n_2 = 0$ when $x \rightarrow \infty$. Because there is consequently no quitting or poaching, $\Phi(w_1 - w_2) \rightarrow \infty$ as $x \rightarrow \infty$.

Case 2: $w_1 = w_2$

Here the first term in (A4) is zero. The second and third terms are only equal to zero when $n_1 = n_2$ and $\phi_1 = \phi_2$. This implies $\Phi(w_1 - w_2) = 0$ if $x = 0$.

Case 3: $x \rightarrow -\infty$

As $x \rightarrow -\infty$, the first term in (A4) $\rightarrow -\infty$. Because there is consequently no quitting or poaching, $\Phi(w_1 - w_2) \rightarrow -\infty$ as $x \rightarrow -\infty$.

To guarantee uniqueness of the symmetric equilibrium $w_1 = w_2$, note that

$$\Phi'(x) = 1 > 0. \tag{A5}$$

REFERENCES

Acemoglu, Daron and Jörn-Steffen Pischke (1998), "Why do Firms Train? Theory and Evidence," *Quarterly Journal of Economics*, February Vol. CXIII, No.1, pp.79-119.

Acemoglu, Daron and Jörn-Steffen Pischke (1999a), "Beyond Becker: Training in Imperfect Labour Markets," *The Economic Journal*, Vol. 109, No.453, pp. F112-F142, February.

Acemoglu, Daron and Jörn-Steffen Pischke (1999b), "The Structure of Wages and Investment in General Training," *Journal of Political Economy*, 107(3), pp. 539-572.

Autor, David (1998) "Why do Temporary Help Firms Provide Free General Skills Training? Theory and Evidence," mimeo, MIT.

Barron, John M., Mark C. Berger and Dan A. Black (1997), "On-the-job Training," Kalamazoo: W.E. Upjohn Institute for Employment Research.

Becker, Gary (1964), *Human Capital*, Chicago: The University of Chicago Press.

Bishop, John H. (1996), "What We Know about Employer-provided Training: a Review of the Literature," Cornell University, Center for Advanced Human Resources Studies Working Paper, 96-09.

Blau, Francine, and Lawrence M. Kahn (1981), “Race and Sex Differences in Quits by Young Workers,” *Industrial and Labor Relations Review*, 34, pp. 563-77.

Booth, Alison L. and Gylfi Zoega (1999a), “Do Quits Cause Under-training?,” *Oxford Economic Papers*, 51, pp.374-386.

Booth, Alison L., Marco Francesconi and Gylfi Zoega (1999b), “Training, Rent-sharing and Unions,” discussion paper # 16/99 Birkbeck College.

Booth, Alison L. and Gylfi Zoega (2000), “Is Wage Compression a Necessary Condition for Firm-financed General Training? A Comment on Acemoglu and Pischke,” discussion paper # 2/00 Birkbeck College.

Calmfors, Lars and John Driffill (1988), “Bargaining Structure, Corporatism and Macroeconomic Performance,” *Economic Policy*, pp. 14-47.

Campbell, III, Carl M. (1993), “Do Firms Pay Efficiency Wages? Evidence with Data at the Firm Level,” *Journal of Labor Economics*, 11, pp. 442-70.

Card, David and Alan B. Krueger (1995), *Myth and Measurement: The New Economics of the Minimum Wage*, Princeton University Press.

Chang, Chun and Yijiang Wang (1996), “Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revisited,” *Journal of Labor Economics*, 14, pp. 505-519.

Frank, Wolfgang and David Soskice (1995), “The German Apprenticeship System” , in Institutional Frameworks and Labor Market Performance. Comparative Views on the German and United States Economies, eds. F. Buttler et al., London and New York: Routledge, pp. 208-34.

Gaynor, Martin, and Mark V. Pauly (1990), “Compensation and Productive Efficiency in Partnerships: Evidence from Medical Group Practice,” *Journal of Political Economy*, 98, pp. 544-573.

Haltiwanger, John C., Julia I. Lane and James R. Spletzer (1999), “Productivity Differences across Employers: The Roles of Employer Size, Age and Human Capital,” *American Economic Review Papers and Proceedings*, 89, pp. 94-98.

Haltiwanger, John C., Julia I. Lane and James R. Spletzer (2000), “Wages, Productivity and the Dynamic Interaction of Businesses and Workers.” Paper presented at the Education and Employment Economics Group (EEEG) Workshop, University of Southampton, July 2000.

Hamilton Jonathan, Jacques-Francois Thisse and Yves Zenou (1999) “Wage Competition with Heterogeneous Workers and Firms” CEPR Discussion Paper No. 2141, forthcoming in the *Journal of Labor Economics.*

Hashimoto, M (1981) “Firm-specific Human Capital as a Shared Investment’, *American Economic Review*, 71(3), pp. 475-82.

Krueger, Alan B. (1993) “How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989,” *Quarterly Journal of Economics*, Vol. 108, pp. 33-60.

Lazear, Edward P. (1998), *Personnel Economics for Managers*, John Wiley & Sons, Inc.

Leuven Edwin and Hessel Oosterbeek (1999) “Demand and Supply of Work-related Training: Evidence from Four Countries,” *Research in Labor Economics*, Fall 1999.

Light, Audrey, and Manuelita Ureta (1992), “Panel Estimates of Male and Female Job Turnover Behavior: Can Female Nonquitters Be Identified?” *Journal of Labor Economics*, 10, pp. 156-81.

Lynch, Lisa M. and Sandra E. Black (1999), “Beyond the Incidence of Employer-Provided Training,” Industrial and Labor Relations Review, vol. 52, pp. 64-81.

Meitzen, Mark E. (1986), “Differences in Male and Female Job-Quitting Behavior,” *Journal of Labor Economics*, 4, pp. 151-67.

Oi, Walter Y (1962) “Labor as a Quasi-fixed Factor,” *Journal of Political Economy*, 70, pp. 538-55.

Pischke, Jörn-Steffen (1998), “Continuous Training in Germany,” mimeo MIT.

Robinson, Joan (1933), *The Economics of Imperfect Competition.* St. Martin’s Press, London.

Salop, Steven (1979), “Monopolistic Competition with Outside Goods,” *Bell Journal of Economics*, vol. 10, pp. 141-56.

Sattinger, M. (1993) “Assignment Models of the Distribution of Earnings,” *Journal of Economic Literature* 31, pp. 831-880.

Shaw, Kathryn L. (1985), “The Quit Decision of Married Men,” *Journal of Labor Economics*, 5, pp. 533-60.

Stevens, Margaret (1994), “A Theoretical Model of On-the-job Training with Imperfect Competition,” *Oxford Economic Papers*, Vol.46, pp.537-62.

Stevens, Margaret (1996), “Transferable Training and Poaching Externalities,” Chapter 2 in A.L. Booth and D.J. Snower (Eds.) *Acquiring Skills*, Cambridge: Cambridge University Press.

Viscusi, W.K. (1979), “Job Hazards and Worker Quit Rates: An Analysis of Adaptive Worker Behavior,” *International Economic Review*, 20, pp. 29-58.