Worker heterogeneity, new monopsony, and training

Working Paper

http://eprints.bbk.ac.uk/3049

Version: Published Working Paper

Citation:


© 2007

Publisher version

All articles available through Birkbeck ePrints are protected by intellectual property law, including copyright law. Any use made of the contents should comply with the relevant law.
Worker heterogeneity, new monopsony and training

Alison L. Boogh and Gylfi Zoega

Address: Alison L. Booth
University of Essex
Wivenhoe Park
Colchester CO4 3SQ

Email: albooth@essex.ac.uk

Address: Gylfi Zoega
Faculty of Economics and Business Administration
University of Iceland
Oddi, at Sturlugata, 101 Reykjavik
Iceland

Email: gz@hi.is
ABSTRACT

A worker’s output depends not only on his or her own ability but also on that of colleagues, who can facilitate the performance of tasks that each individual cannot accomplish on his own. We show that this common-sense observation generates monopsony power and is sufficient to explain why employers might expend resources on training employees even when the training is of use to other firms. We show that training will take place in better-than-average or ‘good’ firms enjoying greater monopsony power, while ‘bad’ firms will have low-ability workers unlikely to receive much training.

Keywords: Firm-financed general training, worker heterogeneity, hierarchical assignment models, monopsony.

JEL Classification: J24, J31, J42

Address for correspondence:
Gylfi Zoega
Department of Economics, University of Iceland
101 Reykjavik, Iceland
E-mail: gz@hi.is

* We should like to thank, for their helpful comments on an earlier draft, Mark Bryan, the late Hugh Davies, Steve Pischke, Ron Smith, and seminar participants at the European Summer Symposium in Labour Economics at Ammersee Germany.. Financial support from the European Commission TSER Award “Labour Demand, Education and the Dynamics of Social Exclusion” is gratefully acknowledged.
1. Introduction

Able colleagues often facilitate the performance of tasks that individuals cannot accomplish on their own. New art movements emanate from groups of similarly minded artists; a concentration of world-class physicists developed the atomic bomb at Los Alamos; the attractiveness of university departments depends on the collective ability of members of staff; a violinist excels as a part of a symphony orchestra, to name just a few examples that come to mind. This simple observation can explain why firms may want to pay for the general training of their workers. In this paper we show how worker and firm heterogeneity can enhance the level of “new monopsony” (see Manning, 2003) in labour markets, and thus create conditions whereby firms are willing to sponsor the general training of their workers. Our central thesis involves demonstrating how the interaction of firm and worker heterogeneity and employment-adjustment costs – in particular training costs – can exacerbate monopsony power in a frictional labour market.

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 specific to firms, 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, 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 and financing purely general training (see for example Krueger, 1993; Bishop, 2000; Barron, Berger and Black, 1997; Autor, 1998; Leuven and Oosterbeek, 1999; Acemoglu and Pischke, 1998, 1999; Booth and Bryan, 2003). Reasons advanced to explain 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; minimum wages; and trade union compression of wages.\(^1\)

In this paper we follow the recent literature in assuming there are frictions in the labour
market, but our point of departure is to allow for heterogeneous workers. Our model has a
continuum of firms ranging from the best to the worst – the best characterised by a workforce of
the highest (total) innate ability. The better firms enable individual workers to perform better
than they would do elsewhere. We model performance by dividing work into increasingly
sophisticated tasks. Workers employed in a ‘good’ firm can perform more sophisticated tasks
than they can perform elsewhere. In this framework, the collective ability of workers at a firm
then determines the range of tasks that can be performed within that firm. A high ability worker
confers on a firm an externality, since her ability raises talent within that firm and increases the
range of tasks that can be performed. Since firms characterised by workforces of higher ability
are able to perform a greater range of more complex tasks, it follows that sophisticated tasks can
be performed in a smaller number of firms than the more simple tasks.

In this framework, we show conditions under which a ‘good’ firm may obtain monopsony
power in the labour market. In particular, the best firm has complete monopsony in the market
for labour doing the most complex tasks that can only be undertaken within that firm. 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, better firms
would be more 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 find that the incidence of training is
positively related to the average education level within a firm.\(^2\)

---

\(^1\) See for example Stevens (1996), Frank and Soskice (1995), Chang and Wang (1996) and Acemoglu and Pischke
(1998, 1999), amongst others.

\(^2\) Lynch and Black use U.S. establishment data to 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.
Kremer (1992) and Sobel (1992) also investigate task complexity, although in a different context. In their sequential production-task model, mistakes made doing the early tasks can be easily corrected – without destroying the output of any subsequent worker. However, mistakes made performing the finishing touches can ruin the work performed doing all earlier tasks. Hence they show that it is of paramount importance to put the more skilled and reliable workers in charge of the final tasks. Although our model is also directed at task complexity, it is in a different context, since our production process is not sequential. While we find, like Kremer (1992) and Sobel (1992), that more able workers should be delegated to the more advanced tasks, it is for a different reason. In our model, a firm assigns the best workers to the most complex tasks because it then reaps maximum profits from its monopsony position – it is more profitable to be in a monopsony position vis-à-vis the more productive workers. Interestingly, the workers may also gain, because they benefit from producing more valuable output as well as from more training – and this augmentation of human capital may result in higher pay.

Our findings also relate to recent studies, using longitudinal and linked employer-employee surveys, showing 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). 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 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.
The remainder of this paper is set out as follows. In Section 2, we set out the model motivation and assumptions about the initial endowment of agents in our economy. In Section 3, we investigate the allocation of heterogeneous workers across tasks. In Section 4 we derive equilibrium wages and training intensity, and in Section 5 show how these differ across tasks. Section 6 discusses some labour-market implications of our model. The final section concludes.

2. The Setup of the Model and its Assumptions

Our model is based on the suppositions that there is an exogenously given distribution of innate ability across workers and that the labour market is imperfectly competitive. This first assumption relates to worker heterogeneity. It seems plausible to suppose that some workers have more to give because of better education, prior work experience, personality traits or intelligence. We formalise this by assuming 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 $h_i$. Our second major assumption – of imperfect competition in the labour market – is captured simply by assuming idiosyncratic match values. Workers have independent and randomly distributed tastes affecting the probability they will stay with a firm. We discuss each of these assumptions – and their ramifications - in more detail below.

2.1 Exogenously given distribution of innate ability across workers

Firms are described by the type and number of workers hired, and the range of tasks any firm can perform is determined by the collective ability of its entire workforce. Each employer advertises its vacancies and receives a finite number of applications from the (finite) pool of potential applicants. Some employers then have a pool of workers of higher ability than others,
and ability within each firm will differ across firms. Thus worker heterogeneity translates into firm heterogeneity when collective abilities within firms are not identical.\(^3\)

Firm heterogeneity in our model – in terms of the collective ability or innate human capital of its workforce – affects the range of tasks that can be performed within the firm. This can be viewed as an intra-firm externality. Our approach is analogous to that of Lucas (1988), who postulates – in a slightly different context – that each producer in an economy benefits from working amongst people of high ability (or what might be termed an intra-economy externality).

### 2.2 How Task Complexity Varies Across Firms

We model each firm as comprising workers doing increasingly sophisticated tasks. There is a market for the output produced from each task. Some firms are better, in the sense that some of their workers do more sophisticated tasks. Examples illustrative of our model are not difficult to come by. Consider for instance the concentration of world-class physicists that developed the atomic bomb at Los Alamos; the continuous flow of new products by Microsoft and other concentrations of skilled programmers; new art movements emanating from groups of similarly minded artists such as the impressionists or the dot-art developed by one group of aborigines in Central Australia. An example closer to home is university departments, which help create and foster synergy in research. Departments with higher concentrations of more talented individuals produce more sophisticated output. Collaborations on research projects can generate results that are often not within the individual reach of those involved.

### 2.3 Task Complexity and Monopsony

Task complexity generates monopsony power. By definition in our model 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 task. As the number of firms doing a given task increases, the level of monopsony power is reduced. Imagine a physicist working at the Los Alamos laboratories in the 1940s. There is only one potential employer at which the task of building an atomic bomb can be accomplished. Clearly, the individuals involved in that project were not able to extract the benefits from the project, which stemmed from the discovery of nuclear energy for peaceful as well as non-peaceful purposes! At the other end of the scale we have very many high schools and colleges where physics is taught. This is the more competitive end of the task scale. When we move from the college to the university – doing research and higher level teaching – the number of potential employers falls. Similarly, as we move towards better and better university departments – doing increasingly sophisticated research and teaching more demanding and receptive students – the level of monopsony is further increased. An outstanding university might be able to perform most of the tasks while the poor inner-city college can only provide teaching in the most basic elements of physics. Finally, when we move to Los Alamos, we are faced with a single potential employer who realises that the very brightest of physicists – who aspire to accomplish the most challenging of tasks – do not have any alternative employer.

Note that the more able workers do not have any offsetting monopoly power. The reason for this is that, on their own, they cannot enable a firm to do a new task. Only a collection of able workers can create the conditions necessary for a new complicated task to be performed. A high-ability worker confers on a firm an externality since her presence raises the level of talent within that firm, and in so doing increases the range of tasks that can be performed. However the high ability worker cannot individually exploit this because it is not her ability that determines task complexity but the intra-firm total.\(^4\) Think of a bright violinist offering her talents to the London Philharmonic or a computer whiz approaching Microsoft for a job. Once the collection of

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).\(^4\) Only if all high-ability employees at a firm co-ordinated their activities – for example through a union – could this
workers is in place, the employer has monopsony power in its relations to any new employee. The employer can give a worker the opportunity to perform a given task – for example only in a good symphony orchestra can a violinist realise the dream of playing the most-challenging of symphonies – but a single worker cannot enable a firm to have a new task performed within its ranks.\footnote{Our approach is not intended to model the behaviour of individual super-stars, such as Pavarotti or Vengerov. Instead we wish to model team players, such as those making up university departments, or who enable Microsoft to}

2.4 The Production Technology

History matters in our model. When a firm starts up, the complexity of its production process is determined by the collective ability of the workforce it selects from workers showing up at its doors. We should emphasize that we intend this only to be a metaphor. We postulate that some firms are better than others because their workers have higher collective ability. This is consistent with our perception of the real world. We do not investigate in this paper the reasons for this being so. We use this construct not only on the grounds that it generates important new insights but also that it offers a more realistic characterization of the labour market than the usual assumption in the literature of a homogeneous workforce. We show that firm heterogeneity can result in firms being willing to finance the general training of their employees.

Continuing the metaphor, each firm can choose not to hire all the workers who show up at its gates, as we show in Section 3. We rank firms in terms of the ability of their workforce. The ranking is purely ordinal, and firms are ranked from the best quality workforce \((r = R)\) to the worst \((r = 1)\). Thus firms differ only in their collective ability or innate human capital

\[
H_R > \ldots > H_i > H_{r-1} \ldots \ldots \ldots \ldots \ldots H_1
\]

We assume increasing returns to ability at the firm level, which takes a particular form shown in Figure 1: the higher is collective ability in the firm, the more advanced are the tasks that can be performed within it.
We do not model the within-firm interactions between workers. Instead, we simply assume that innate ability differs across firms depending on workers’ endowments, and that more complex tasks can only be performed where ability is higher.\(^6\)

The firm ranked \(R\) can potentially employ workers who have the highest ability and therefore perform the most complex task, in addition to all simpler ones. Whether it actually does so will depend on profitability, whether the profits from so doing exceed the costs of training the workers to do task \(R\). Firm \(R-1\) can perform 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 able to perform task \(k\), it is clear that \(M_k\) is decreasing in the complexity of the task performed. More precisely, we can write

\[
M_k = R - k + 1
\]

where \(R\) is the total number of firms and \(k\) is the number of tasks (potentially) performed at a particular firm, \(k \in [1,R]\). Equation (1) implies that there is (potentially) one firm doing the most advanced task \(k = R\), two firms doing the second most advanced, and \(R\) firms doing the most simple task.

### 2.5 The Training Technology and Firms’ Profits

We assume that production or value added arising from each task is a linear function of labour (measured in efficiency units). We measure the value of output from each task in terms of units of output emanating from the simplest task. Thus \(y_k\) is the value of output per efficiency unit of labour devoted to task \(k\) measured in units of output coming from task \(1\). It follows that \(y_1 = 1\).

---

\(^6\) In our model no individual perceives that, on her own, she might have the ability level to switch a firm from one level of task complexity to another. This follows from our assumption that people individually but simultaneously make their decisions about which firm to apply to.
Workers at a firm receive on-the-job training that is specific to their task but general to all firms undertaking that task. (If there is only one firm carrying out a task, then the skill for that task is pure specific.) The amount of training a worker receives – or training intensity - is denoted by $\phi$. Training intensity then determines the extent to which post-training productivity and wages are augmented, as described below. The cost of providing the worker with training $\phi$ is described by the strictly convex function $c_k(\phi)$:

$$c_k(\phi) = k + c(\phi), \quad k \in [1, R], \quad c'(\phi) > 0, \quad c''(\phi) > 0$$

(2)

We assume, quite plausibly, that the cost of training is rising in task complexity as well as training intensity. By assumption the former is a linear relationship while the latter is strictly convex. However, the assumption of a linear relationship between training costs and task complexity can be easily relaxed.

A worker’s innate ability is augmented multiplicatively (or log-additively) through training to become $h_i f(\phi_i)$, which denotes post-training human capital, both innate and acquired at the workplace. Thus a worker trained for task $k$ has $f(\phi_{ik})h_i$ efficiency units of labour and produces output valued at $f(\phi_{ik})h_iy_k$. The function $f(\phi)$ is strictly concave:

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

Wages per efficiency unit of labour are denoted by $w_k$ so the remuneration of a worker with human capital $h_i$ working on task $k$ is given by $f(\phi_{ik})h_iw_k$. Note that a high-ability worker will gain more – in absolute terms – from participating in training. However, when two workers of different ability enter the same training programme and receive the same intensity of training, the ratio of their human capital – innate as well as acquired – stays constant.\(^7\)

The $k$-th firm receives profits $P$ from the value-added net of training costs of the tasks performed within its ranks. Value-added is the difference between the value of output (measured

---

\(^7\) As shown in Booth and Zoega (2004), our formulation gives wage compression as defined by Acemoglu and
in terms of output from the most basic task) and wages per efficiency unit of labour. The $k$-th firm’s expected profits from employing worker $i$ can thus be written as

$$P^i_k = (1 - q)f(\phi_{ik})(y_k - w_k)h_i - (k + c(\phi_{ik}))$$  \hspace{1cm} (3)

where $q$ denotes the quit rate and $1-q$ the retention rate (more on that below). Profits from task $k$ can then be calculated by adding up the profits of workers that have been allocated to the task. Finally, the firm’s total profits equal the sum of profits emanating from all tasks performed.

### 2.6 ‘New Monopsony’

We now introduce into the model a labour market imperfection based on the assumption of idiosyncratic match values. Following Stevens (1994, 1996), we assume that workers have independent and randomly distributed tastes affecting the probability they will stay with a firm. These frictions make it costly for workers to change jobs. This is a simple way of capturing ‘new monopsony’, whereby firms face upward-sloping supply curves arising from labour-market frictions (rather than market structure), such as mobility costs (see inter alia Stevens, 1994, Manning, 2003, and Hirsch and Schumacher, 2003). Once trained, workers choose either to stay with the firm that provided training and produce, or quit to work in other firms in the skilled sector. The probability for a given firm of retaining current workers is captured by the reduced form quit function $q(w - w^A, M_k)$, where $w^A$ is the wage differential between the firm under consideration and other firms ($w^A$). By offering higher wages, a firm might attract an employee from a competing firm. The probability of that same firm poaching a worker trained by other firms is captured by the poaching function $p(w - w^A, M_k)$.

The two functions, and their first derivatives, affect the optimal level of wages and hence also the optimal level of training. Following Stevens (1996) we also assume that a firm’s poaching and quitting functions have the following properties:

---

Pischke (1999): increased training raises expected productivity by more than expected wages.
\[
\lim_{M_k \to \infty} p = 0, \quad \lim_{M_k \to \infty} p_w = \infty, \quad \lim_{M_k \to \infty} q = 1, \quad \lim_{M_k \to \infty} q_w = -\infty
\]

As the number of firms \(M_k\) rises, both poaching and quitting become more responsive to changes in wages. Similarly, a larger number of firms create a greater temptation for more workers to move in the absence of wage incentives. This is because there are now more outside opportunities – the quit rate approaches unity as the number of firms approaches infinity. It becomes more difficult to poach for the same reason – the poaching probability approaches zero. Booth and Zoega (2001) provide micro-foundations for the properties of these two functions based on worker preferences for job characteristics.\(^8\)

3. The Intra-firm Allocation of Tasks

Firms’ task heterogeneity is determined by the innate ability of workers selected from those who show up. Not all workers who turn up are selected due to the cost of training them. Firms select the best quality workers and then begin training and production. We now consider the intra-firm allocation or assignment of workers across tasks for given optimally chosen levels of wages and training intensity for each task.

Each firm undertaking task \(k\) uses a technology that is constant returns to scale with respect to the human capital it employs. Each firm takes the other firms’ actions as given, and chooses the number of workers to allocate to task \(k\), \(n_k\), by maximising the sum of expected per worker profits shown in equation (3) subject to the number of available workers of each skill type. The firm anticipates setting wages and training optimally, and the reservation wage is normalised to zero. Remuneration and marginal productivity vary with a worker’s innate productivity, \(h_i\).

A firm will only hire a worker as long as it profits from doing so. Equation (3) implies that the following condition must hold for each worker (optimally) trained (\(\phi^*_i\)) to do task \(k\):

\[
(1 - q)f(\phi^*_i)\left(y_k - w^*_k\right)h_i - \left(k + c(\phi^*_i)\right) \geq 0
\]
The first term is the expected profits from employing worker $i$, and the second term is the cost of training her. Task complexity $k$ determines both the cost of training – by assumption linearly – as well as the monopsony profits $(y-w)$ as derived below. All workers that satisfy equation (4) can be trained to do task $k$. However, this is only a necessary condition, not a sufficient condition, since the firm might profit more from training them to do some alternative task. Clearly, a given worker is hired if she satisfies equation (4) for at least one task within the firm.

Turning to the intra-firm allocation of workers across tasks, should a firm use a few of the best workers for the most difficult task or should it use many less able workers for that task? We will show that the best workers should be allocated to the most advanced tasks. Assume there is a worker with ability $h_j$ doing task $r$ – the most sophisticated within firm $r$ – and another doing the least advanced task – task 1 – with a higher ability $h_i$; $h_i > h_j$. If the firm could relocate the two workers at zero cost so that the latter would be trained to do task $r$ and the former to do task 1, total training costs would be unchanged and the benefit to the firm would be

$$
(1-q) \left[ f\left(\phi^*_r\right)(y_r-w^*_r) - f\left(\phi^*_1\right)(y_1-w^*_1)\right] (h_r-h_j) > 0
$$

if and only if

$$
f\left(\phi^*_r\right)(y_r-w^*_r) > f\left(\phi^*_1\right)(y_1-w^*_1)
$$

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 $r$ (proved in the following section) and a higher level of training $\phi_r > \phi_1$. It will then become clear that the best workers will be allocated to the most advanced task performed with the firm and the worst to the least advanced one.\(^8\)

\(^8\) Following an approach similar to Bhaskar and To (1999), Hamilton, Thisse and Zenou (2000), and Thiesse and Zenou (2000).

\(^9\) If there is a shortage of workers at least as able as this marginal (minimum ability) worker – this is the pool of desired workers – the firm cannot hire as many workers for each task as it desires. (Recall there is a finite number of each ability type seeking employment.) Faced with this constraint, the best the firm can do is to allocate workers efficiently across tasks, leaving unfilled vacancies throughout the firm.
Now consider the firm ranked $R$ – which can perform all tasks of complexity $R$ and below. If training costs are not too high, the firm will have task $R$ performed. It follows from equation (4) that the threshold ability level is:

$$h_i (1 - q) f(\phi^*_i) (y^*_R - w^*_R) \geq R + c(\phi^*_i)$$

(6)

The left-hand side of (6) shows the benefit of training worker $i$ to do task $R$. This is increasing in innate ability $h_i$: the more able the worker, the greater the benefits from training her to do the task. The right-hand side shows the cost of training the worker to do task $R$. This is also, by assumption, increasing in task complexity. Clearly, if any workers should be assigned to this advanced task $R$, it should be those with the highest innate ability $h_i$. But the firm has first to calculate the opportunity cost - that is the benefits and the costs of training that worker to do other tasks – and the worker will be assigned to whatever task yields the greatest net benefit. This may not be task $R$ if the cost of training is sufficiently high.

Equation (6) demonstrates that the benefits the firm obtains from its monopsony power $y-w$ are an increasing function of workers’ innate human capital $h_i$. This underlies the insight of equation (5) - that it is better to train the high-ability workers to do the more advanced tasks. But now, as we move down the ability ladder to less and less able workers, it becomes optimal to assign them to simpler tasks with low training costs because the firm cannot reap large monopsony profits from their employment. Therefore, as we move from the most sophisticated task – perhaps task $R$ – to less sophisticated ones, we are bound to meet workers of lower ability. A worker who does not meet the grade for the least advanced task 1 is then turned away:

$$h_i (1 - q) f(\phi^*_i) (y^*_1 - w^*_1) < 1 + c(\phi^*_i)$$

(7)
4. Equilibrium Wages and Training Intensity

In this section we investigate the determination of wages and training intensity. Firms augment each worker’s human capital through training that, while specific to a particular task, can be used at other firms also undertaking that task.

We now return to equation (1). Firm $R$ is the only firm in the market for labour trained to perform its most sophisticated task – subject to the caveats spelled out in Section 3 – and produce units of output valued at $y_R$ 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 output valued at $y_{R-1}$. The degree of labour-market competition is then rising until it reaches its maximum for the first units valued at $y_1$.

We now consider the firm’s profits from task $k$ in equation (3). Each firm, taking the other firms’ actions as given, sets wages and the level of training by maximising profits net of training costs from each worker performing the task $P_k'$

$$\max_{w_i, w_k} (1-q)f(\phi_k)(y_k - w_k)h_i + pf(\phi^A_k)(y_k - w_k)h^A - c_k(\phi)$$

where $h^A$ is average innate human capital in other firms and $\phi^A$ is the average level of training in other firms. The first term gives expected profits from employing a retained worker trained in this firm. The second term gives the expected profits from workers trained by all other firms performing this task. Finally, the last term has the cost of training given by the strictly convex function $c_k(\phi)$. To make things simple, we assume that the ‘poachees’ are of the same ability as
worker $i$, $h_i = h^A$. This has the consequence that optimal wages per efficiency unit of labour $w_k^*$ become independent of the worker’s ability $h_i$.\footnote{In a preliminary but more complex version of this paper we relaxed this assumption and allowed for asymmetries between current and prospective workers. The results are not overturned – although optimal wages now become} We now come to the paper’s first proposition.

**Proposition 1**: The equilibrium wage rate $w_k$ and training intensity $\phi_{ik}$ are symmetric across firms performing task $k$ for workers of identical ability.

**Proof**: See the Appendix.

We now consider the determination of wages and training intensity in some detail.

### 4.1 Wage-setting

Wages are set to reduce quitting as in the efficiency wage model of Salop (1979) and in Stevens (1994, 1996). Raising wages can increase profits through increased retentions because quitting is costly due to the monopsony power firms enjoy viz-à-viz their employees. At the optimum, the marginal cost and benefit of raising wages are equal. The wage is determined by the following equation, which is the first-order condition of (8) with respect to wages;

$$\left(p' f(\phi^A_k) - q' f(\phi^A_i)\right)y_k - w_k = pf(\phi^A_i) + (1 - q)f(\phi^A_i)$$

where $\phi^A_k$ is the average training intensity elsewhere. The left-hand side represents the marginal benefit of raising wages – in the form of fewer quits and more poachees – and the right-hand side denotes the marginal cost – in terms of a higher wage bill.

Equation (9) can be solved for wages paid per efficiency unit for task $k$:

$$w_k^* = y_k - \frac{pf(\phi^A_k) + (1 - q)f(\phi^A_i)}{p' f(\phi^A_k) - q' f(\phi^A_i)} = y_k - \Psi_k$$

In symmetric equilibrium, the wage is independent of training intensity $\phi$ and can be written as:

$$w_k^* = y_k - \frac{p + (1 - q)}{p' - q'} = y_k - \Psi_k$$

\footnote{In a preliminary but more complex version of this paper we relaxed this assumption and allowed for asymmetries between current and prospective workers. The results are not overturned – although optimal wages now become}
Note that equation (10) does not give the actual remuneration received by a worker, which is given by $f(\phi^*_{ik})h_iw_k^*$ where $\phi^*_{ik}$ is the optimal training intensity to be determined below. We now return to the properties of the poaching function and the quitting function. In particular, the term $\Psi_k$ is decreasing in the number of firms $M_k$ and approaches zero as $M_k$ approaches infinity. This gives us the following proposition.

**Proposition 2:** The gap $\Psi_k$ between wages per efficiency unit and productivity is declining in the number of firms. Hence the more sophisticated a task, the larger is $\Psi_k$.

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 task, hence the degree of competition is highest for this task. But as firms become ‘better’, they gain elements of monopsony power in the market for labour trained to do the more complex tasks within their ranks.

Note that a generally bad firm – measured in terms of the collective ability of its workers – cannot assign the few good workers it has to complicated tasks because these cannot be performed within its ranks. In contrast, a generally good firm will assign the few bad workers it has to do some simple tasks. This is, intuitively, the reason why there is less competition in the labour market as we move to more complex tasks. It is not difficult to imagine scenarios of people sweeping floors in the most advanced firms, while there is no one doing the most challenging tasks imaginable in low-ability firms that focus on basic tasks. Note that this does not mean that every advanced firm has people sweeping floors only that this scenario is more likely than the latter.

### 4.2 Training Intensity

The first-order condition for the level of training per efficiency unit for task $k$ is:

worker specific – as long as firms keep some of the rent from employing the workers. For the higher-level tasks the
The left-hand side gives the expected marginal benefit from increased training while the right-hand side gives the marginal cost.\(^{11}\)

Equation (11) has two interesting features. First, the level of training depends on a worker’s ability \(h_i\): the more able the worker, the more training he receives from his employer. It follows that workers doing task \(k\) will not all receive the same level of on-the-job training, since they will not necessarily all have the same innate ability. Moreover, all workers with the same ability will not necessarily be trained for the same task, since this depends on which firm employs them, i.e. the ability of their colleagues.

Second, according to equation (11) there is under-investment in training because the firm discounts the marginal benefit from training \(f'(\phi^*_k)\) by \((1-q)\). In contrast, the first-best level \(\phi^{**}\) is set to maximise the value of total output produced by all trained workers doing the \(k\)-th task – both those retained by any firm, plus those who quit to work in other firms – less the costs to society of training (including the opportunity cost of labour). In short, \(\phi^{**}\) solves \( f'(\phi^{**}_k) (y_k - w^{**}_k) h_i = c'_k (\phi^{**}_k) \), where \(w^{**}_k = 0\) when the opportunity cost of labour is zero. It follows that \(\phi^* < \phi^{**}\). We now have our next proposition.

**Proposition 3:** The amount of training set by the firm \(\phi^*\) is always positive, but is less than the efficient level given by \(\phi^{**}\).

This result arises quite naturally in models with imperfect competition in labour markets, such as ours, and can be traced back to the work of Pigou (1912).\(^{12}\)

\(^{11}\) The second-order condition for maximum is satisfied because the Hessian matrix is negative definite.

\(^{12}\) See also Stevens (1994, 1996). See Booth, Chen and Zoega (2002) for a comparison of training and firing costs.
4.3 The Financing of General Training

We now consider the financing of this general training, where the firm’s choice of training intensity is denoted by $\phi^*$. Proposition 2 and equation (11) imply our Proposition 4.

**Proposition 4:**

(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.*

This is an important result. The firm is here willing to finance its chosen level of general training because it augments workers’ productivity in a multiplicative way. From equation (11) it follows that the larger is $\Psi$ – the difference between output and wages – the greater is the level of $\phi$. It also follows that the firm finances more extensive training for the sophisticated tasks. Figure 2 shows the optimal level of training as a function of the number of firms performing the task.

![FIGURE 2](image)

Now returning to equation (11) 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 4 and Figure 2, these workers also have greater ability or innate human capital $h_i$ and receive more training for that reason.

5. Gross Wages, Training and Task 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 gross wage – or total remuneration – of worker $i$ performing the $k$–th task is given by $W_{ik}$:

$$W_{ik}^* = f\left(\phi_{ik}^*\right)h_iy_{ik}^* = f\left(\phi_{ik}^*\right)h_i\left(y_{ik}^* - \Psi_{ik}\right)$$

(12)
While a worker’s inherent human capital \( h \) is exogenous to our model, we have modelled the determination of both training and wages per efficiency unit. Gross 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; \( \Psi \). This difference then makes firms choose a higher level of training. To see this, take the total differential of equation (11) with respect to \( \phi \) and \( \Psi \) to obtain:

\[
\frac{d\phi^*_k}{d\Psi_k} = \frac{(1-q)\frac{\partial f}{\partial \phi^*_k}}{c_k(\phi^*_k)(1-q)f''(\phi^*_k)h_k\Psi_k} > 0 \tag{13}
\]

Not surprisingly, the effect of changes in the level 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 ability of workers. It follows that if both the cost function \( c_k \) and the training function \( f \) are close to being linear, the effect of enhanced monopsony power on training can be very significant, especially for the better workers.

Importantly, the effect of an increase in monopsony power on training is decreasing in the level of \( \Psi_k \) due to the concavity of the training function \( f \) – decreasing returns to training intensity – as shown in Figure 2. The effect of enhanced monopsony power on wages going through training is then also decreasing in the level of \( \Psi_k \). Depending on the value of the model’s parameters, the stage is set for the possibility of a hill-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 – which comes on top of the rising value of output – may dominate its direct effect so that wages are increasing in task complexity. The indirect effect may then diminish more rapidly – due to decreasing returns to training intensity – so that the direct effect starts dominating and wages start falling in the
complexity of the task. In summary, wages are increasing in task complexity because of increased training and also higher output per efficiency unit of labour $y_k$. Wages are decreasing in complexity because of higher monopsony power $\Psi_k$.

6. Applications to Labour Markets

Our model shows how wage compression arises quite naturally in a market economy with heterogeneous workers. The level of monopsony power is increasing in task complexity, which explains why wages (per efficiency unit) rise less than productivity as we move from one task to the other. In showing that training intensity is rising in task complexity, as also is the value of output, we have illustrated how the distribution of gross wages or total remuneration can take a hill-shaped or concave form.

Lucrative consulting opportunities among university professors may provide one example of the relevance of a hill-shaped relationship between task 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 earned – often for more demanding tasks – 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 our model. Universities enjoy monopsony power in the market for many forms of research and teaching. They can therefore pay a salary that is lower than the value of output, which then makes them benefit from any accumulation of human capital (and in many instances patent rights) of their employees.

We have shown how the level of training is increasing in task complexity. The more advanced the task, the greater the employers’ monopsony power and the greater is her willingness to finance the (general) training of her employees. Once again, examples from our

---

13 The hill-shaped relationship depends crucially on productivity not increasing too rapidly in task complexity ($y_k$ exceeding $y_{k-1}$ by too much) and the shape of the training and cost functions being such that the second derivative of
own profession abound. We may be able to explain the generous treatment given to many professors in the form of teaching reductions, financial support for research activities, extensive sabbaticals and a laissez-faire attitude to apparently self-centred behaviour and activities. In contrast, a stricter regime is imposed on secretaries and junior faculty.

Fortunate university graduates sometimes face the trade-off whether to accept higher-paying jobs at average universities or lower-paying jobs in the better universities. More often than not, they choose the latter. Our model gives a very plausible rationale. By accepting a job in a better university, the graduate has a chance to be assigned more sophisticated tasks (i.e. taking part in more advanced and demanding research projects), an opportunity that he does not have elsewhere. It follows that he may initially not be paid well due to the monopsony element. However the university will provide good training – formal and informal through collaborations with colleagues – that will augment the graduate’s innate human capital. And eventually this may raise his pay above the level he could hope for in a less prestigious institution.

When looking across workers performing the same task, workers with higher human capital also receive higher remuneration. The reason is twofold. First, they embody more units of innate human capital to start with. Second, because their endowment of human capital is greater, they also receive more training, which raises their wage further. Again, we can take examples from our own profession. In a university department, the more productive professors might both be paid more than their less productive colleagues due to their initially higher creativity but are also often given less teaching and more opportunities to enhance their skills further.

7. Conclusion

Workers’ output depends not only on their own ability but also on that of colleagues, who can enable each individual to perform tasks they cannot accomplish on their own. We have shown that this simple truism gives some employers an element of monopsony power. An employer

\[ \phi \text{ with respect to } \Psi \text{ is sufficiently negative.} \]
whose workers are of high collective ability – and thus enhance each other’s individual productivity through facilitating the performance of a more complex task – can offer wages that fall short of marginal (within-firm) productivity. Since firms performing the more basic tasks are more numerous than those performing the more advanced tasks, workers doing the more advanced tasks will have more limited outside opportunities. Alternative firms may not perform the task for which they have been trained because they might have workforces of lower ability.

Because of this ‘new’ monopsony, employers find it in their interest to expend resources on the training of the workforce. Employers, knowing they will keep part of any extra surplus from training, benefit from financing training, even training that is perfectly general across firms. This helps explain why so much general training appears to be financed by employers.
Figure 1. Increasing returns to skills

Figure 2. Training, task complexity and the level of monopsony

\[ \text{wages} = w_h f(\phi) h \]

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 for the $k$-th task can written as

$$ P_{1k} = (1-q_{1k}) f(\phi_{ik1})(y_k - w_{1k}) h_i + p_{1k} f(\phi_{ik2})(y_k - w_{1k}) h_i - c_k(\phi_{1}) $$  \hspace{1cm} (A1)

where $\phi_{ik1}$ denotes the training of worker $i$ doing task $k$ in firm 1 and we have assumed that a potential poachee is of the same ability as the worker whose wages are being set. Maximisation of (A1) w.r.t. training intensity $\phi_{1k}$ and wages $w_{1k}$ respectively gives:

$$ (1-q_{1k}) f'(\phi_{ik1})(y_k - w_{1k}) h_i = c'_k(\phi_{ik1}) $$  \hspace{1cm} (A2)

$$ [p_{1k} f(\phi_{ik2}) - q_{1k} f(\phi_{ik1})](y_k - w_{1k}) = p_{1k} f(\phi_{ik2}) + (1-q_{1k}) f(\phi_{ik1}) $$  \hspace{1cm} (A3)

The analogous expected profit equation for firm 2 is:

$$ P_{2k} = (1-q_{2k}) f(\phi_{ik2})(y_k - w_{2k}) h_i + p_{2k} f(\phi_{ik1})(y_k - w_{2k}) h_i - c_k(\phi_{2}) $$  \hspace{1cm} (A1b)

The first-order conditions of firm 2 are:

$$ (1-q_{2k}) f'(\phi_{ik2})(y_k - w_{2k}) h_i = c'(\phi_{ik2}) $$  \hspace{1cm} (A2b)

$$ [p_{2k} f(\phi_{ik1}) - q_{2k} f(\phi_{ik2})](y_k - w_{2k}) = p_{2k} f(\phi_{ik1}) + (1-q_{2k}) f(\phi_{ik2}) $$  \hspace{1cm} (A3b)

We now subtract (A3b) from (A3) to obtain after some manipulation the following function, assumed to be continuous from continuity of the underlying functions:

$$ \Phi(w_{1k} - w_{2k}) = (w_{1k} - w_{2k}) + \left[ \frac{p_{1k} f(\phi_{ik2}) + (1-q_{1k}) f(\phi_{ik1})}{p_{1k} f(\phi_{ik2}) - q_{1k} f(\phi_{ik1})} \right] \Phi - \left[ \frac{p_{2k} f(\phi_{ik1}) + (1-q_{2k}) f(\phi_{ik2})}{p_{2k} f(\phi_{ik1}) - q_{2k} f(\phi_{ik2})} \right] $$  \hspace{1cm} (A4)

We now show existence of an industry equilibrium by showing that equation (A4) only holds $w_{1} = w_{2}$.

Case 1: $w_{1k} - w_{2k} \rightarrow \infty$

Here the first term in (A4) is $\infty$. Clearly $\Phi(w_{1k} - w_{2k}) \rightarrow \infty$ as $w_{1k} - w_{2k} \rightarrow \infty$. 

24
**Case 2:** \( w_{1k} = w_{2k} \)

Here the first term in (A4) is zero. The second and third terms are only equal to zero when \( \phi_{1k} = \phi_{2k} \). This implies \( \Phi(w_{1k} - w_{2k}) = 0 \) if and only if \( w_{1k} = w_{2k} \) and \( \phi_{1k} = \phi_{2k} \).

**Case 3:** \( w_{1k} - w_{2k} \to -\infty \)

As \( w_{1k} - w_{2k} \to -\infty \), the first term in (A4) \( \to -\infty \). Because there is consequently no quitting or poaching, \( \Phi(w_{1k} - w_{2k}) \to -\infty \) as \( w_{1k} - w_{2k} \to -\infty \).

To guarantee uniqueness of the symmetric equilibrium \( w_{1k} = w_{2k} \), note that

\[
\Phi'(w_{1k} - w_{2k}) > 0.
\] (A5)

if \( \phi_{1k} = \phi_{2k} \) and \( p'' = -q'' \).
REFERENCES


INSTITUTE OF ECONOMIC STUDIES WORKING PAPERS 1987-2006
Formerly Iceland Economic Papers Series

Editor Sveinn Agnarsson

A complete list of IoES Working Papers and Reprints can be accessed on our World Wide Web site at
http://www.ioes.hi.is

W07:11 Alison L. Booth and Gylfi Zoega: Worker heterogeneity, new monopsony and training.
W07:10 Yu-Fu Chen and Gylfi Zoega: Aging and job security.
W07:09 Ron Smith and Gylfi Zoega: Keynes, investment, unemployment and expectations.
W07:07 Helgi Tomasson: Likelihood based surveillance of continuous-time processes.
W07:05 Anna Gunthorsdottir, Roumen Vragov, Kevin McCabe and Stefan Seifert: The meritocracy as a mechanism to overcome social dilemmas.
W07:04 Gylfi Zoega: Endogenous employment cycles in Euroland.
W07:03 Thorvaldur Gylfason: The international economics of natural resources and growth.
W07:02 Helga Kristjansdottir: Talking trade or talking aid? Does investment substitute for aid in the developing countries?
W06:13 Brynhildur Davidsdottir: Sustainable energy development.
W06:12 Helga Kristjansdottir: Substitution between inward and outward foreign direct investment.
W06:11 Helga Kristjansdottir: Evaluation of Icelandic trade flows, the gravity model approach.
W06:10 Brynhildur Davidsdottir: Capital constraints and the effectiveness of environmental policy.
W06:09 Gylfi Zoega: Market forces and the continent’s growth problem.
W06:08 Fridrik M Baldursson and Nils-Henrik M von der Fehr: Vertical integration and long-term contracts in risky markets.
W06:07 Ragnar Arnason: Conflicting uses of marine resources: Can ITQ’s promote an efficient solution?
W06:06 Thorvaldur Gylfason and Gylfi Zoega: A golden rule of depreciation.
W06:05 Ron Smith and Gylfi Zoega: Global factor, capital adjustment and the natural rate.
W06:04 Thorolfur Matthiasson: To whom should the rent accrue?
W06:03 Tryggvi Thor Herbertsson and Gylfi Zoega: Iceland’s Currency Dilemma.
INSTITUTE OF ECONOMIC STUDIES

W06:02 Thorolfur Matthiasson: Possible stakeholder conflicts in quota regulated fisheries, contribution to the political economics of fisheries.

W06:01: Eyjolfur Sigurdsson, Kristin Siggeirsdottir, Halldor Jonsson jr., Vilmundur Gudnason, Thorolfur Matthiasson, Brynjolfur Y Jonsson: Early discharge and home intervention reduces unit costs after total hip replacement: Results of a cost analysis in a randomized study.

W05:14 Gylfi Zoega and J Michael Orszag: Are Risky Workers More Valuable to Firms?

W05:13 Friðrik Már Baldursson: Fairness and pressure group competition.

W05:12 Marías H. Gestsson and Tryggvi Thor Herbertsson: Fiscal Policy as a Stabilizing Tool.

W05:11 Tryggvi Thor Herbertsson and Gylfi Zoega: On the Adverse Effects of Development Aid.

W05:10 Thráinn Eggertsson and Tryggvi Thor Herbertsson: Evolution of Financial Institutions: Iceland’s Path from Repression to Eruption.


W05:08 Ron Smith and Gylfi Zoega: Unemployment, investment and global expected returns: A panel FAVAR approach.

W05:07 Gylfi Zoega and Thorlakur Karlsson: Does Wage Compression Explain Rigid Money Wages?

W05:06 Thorvaldur Gylfason: India and China

W05:05 Edmund S. Phelps: Can Capitalism Survive?

W05:04 Thorvaldur Gylfason: Institutions, Human Capital, and Diversification of Rentier Economies

W05:03 Jón Danielsson and Ásgeir Jónsson: Countercyclical Capital and Currency Dependence

W05:02 Alison L. Booth and Gylfi Zoega: Worker Heterogeneity, Intra-firm Externalities and Wage Compression

W05:01 Tryggvi Thor Herbertsson and Martin Paldam: Does development aid help poor countries catch up?

W04:12 Tryggvi Thor Herbertsson: Personal Pensions and Markets

W04:11 Fridrik M. Baldursson and Sigurður Johannesson: Countervailing Power in the Icelandic Cement Industry

W04:10 Fridrik M. Baldursson: Property by ultimatum: The case of the Reykjavík Savings Bank

W04:09 Ingólfr Arnarson: Analyzing Behavior of Agents of Economic Processes in Time

W04:08 Otto Biering Ottosson and Thorolfur Matthiasson: Subsidizing the Icelandic Fisheries


W04:06 Ingolfur Arnarson: Modelling Fishery Management Schemes with an Olympic System Application

W04:05 Ingolfur Arnarson and Pall Jansson: Adding the Sales Markets Dimension to Bio-Economic Models. The Case of Fishery Management
W04:04 Edmund S. Phelps: Changing Prospects, Speculative Swings: Structuralist Links through Real Asset Prices and Exchange Rates

W04:03 Ingolfur Arnarson: Analysing Behavior of Agents of Economic Processes in Time

W04:02 Ron Smith and Gylfi Zoega: Global Shocks and Unemployment Adjustment

W04:01 Fridrik M. Baldursson and Nils-Henrik M von der Fehr: Prices vs. quantities: public finance and the choice of regulatory instruments

W03:07 Sveinn Agnarsson and Ragnar Arnason: The Role of the Fishing Industry in the Icelandic Economy. A historical Examination

W03:06 Thorolfur Matthiasson: Paying paper by paper, the wage system of Icelandic University teachers explained

W03:05 Gur Ofur and Ilana Grau: Bringing the Government hospitals into line: The next step of reform in the healthcare sector

W03:04 Ingolfur Arnarson and Pall Jensson: The Impact of the Cost of the Time Resource on the Efficiency of Economic Processes

W03:03 Torben M. Andersen and Tryggvi Thor Herbertsson: Measuring Globalization

W03:02 Tryggvi Thor Herbertsson and J. Michael Orszag: The Early Retirement Burden: Assessing the Costs of the Continued Prevalence of Early Retirement in OECD Countries

W03:01 Eirik S. Amundsen, Fridrik M. Baldursson and Jørgen Birk Mortensen: Price Volatility and Banking in Green Certificate Markets

W02:10 Tryggvi Thor Herbertsson and Gylfi Zoega: A Microstate with Scale Economies: The Case of Iceland

W02:09 Alison, L. Booth and Gylfi Zoega: Is Wage Compression a Necessary Condition for Firm-Financed General Training

W02:08 Asgeir Jonsson: Exchange rate interventions in centralized labor markets

W02:07 Alison, L. Booth, Marco Francesconi and Gylfi Zoega: Oligopsony, Institutions and the Efficiency of General Training

W02:06 Alison L. Booth and Gylfi Zoega: If you’re so smart, why aren’t you rich? Wage inequality with heterogeneous workers