WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

A THEORY OF INTER-INDUSTRY WAGE DIFFERENTIALS

by
Julio J. Rotemberg
Garth Saloner

WP# 1822-86 September 1986

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139
A THEORY OF INTER-INDUSTRY WAGE DIFFERENTIALS

by

Julio J. Rotemberg
Garth Saloner

WP# 1822-86 September 1986

Sloan School of Management and Department of Economics, M.I.T. respectively. We wish to thank William Dickens, Lawrence Katz and Lawrence Summers for helpful discussions and the NSF for research support.
ABSTRACT

The purpose of this paper is to present a model which broadly fits some of the salient features of inter-industry wage differentials. Several recent empirical papers have found wage differences across industries to be large and persistent. They also show high concordance across occupations and countries. High wages appear to be paid in industries that have high capital/labor ratios and are highly profitable. Our model explains these facts on the basis of firm-specific human capital accumulation by individual workers. We focus on the bargaining between experienced workers and the firm over the division of the surplus output an experienced worker produces over that produced by inexperienced workers. We show that this surplus, and therefore equilibrium wages of trained workers, depends on the the capital/labor ratio when the technology has putty-clay features. We also show that when there is multilateral bargaining between all firms and experienced workers, wages also depend on the profitability of the firm.
I Introduction

The purpose of this paper is to present a model which broadly fits some of the salient features of inter-industry wage differentials. These features are discussed in a number of recent papers including those of Krueger and Summers (1986a,b) and of Dickens and Katz (1986a,b). These papers find that inter-industry wage differences are large and persistent. They are also highly correlated across countries with substantial concordance across Western industrial countries and Eastern Block countries. Dickens and Katz (1986a) also show that the concordance is high across fairly narrowly defined occupational groups. One very striking feature reported both in Krueger and Summers (1986) and Dickens and Katz (1986b) is that industries in which, broadly speaking there are more variable profits, i.e. where the difference between revenues and variable costs is greater, tend to have higher wages. In particular, industries with high profits tend to have high wages even controlling for a host of other variables while the same is true, although to a lesser extent, for industries with a high capital/labor ratio.

Our model explains these differences on the basis of on-the-job accumulation of specific human capital by individuals. After an employee has accumulated specific knowledge about the important characteristics of his job, like the specific mechanical motions required of production workers, the idiosyncracies of particular
filing systems for administrative assistants or the tastes of
important clients for sales personnel he is more valuable to the firm
than are new trainees (see Becker 1964). This does not
automatically mean that a trained individual will receive more
compensation than his untrained counterpart. Suppose in particular
that all workers have the same marginal disutility of time. Then
workers would be willing to sign binding long-term contracts that
involve wages equal to this marginal disutility of time whether they
are trained or not. Our model assumes that such contracts are
infeasible. There is thus ex-post bargaining between the firm and
the worker over the division of the excess output produced by the
trained worker1.

Without more knowledge about the institutional framework for
this bargaining it is impossible to predict how this excess output
(or surplus) will be divided. Suppose the firm can make credible
take-it-or-leave it offers to workers; in other words, it moves
first at the beginning of each bargaining period, quotes a wage at
which it is willing to employ the worker and is committed to ignore
any counteroffers. The worker can only accept or reject the offer.
This gives the firm all the "bargaining power" so it receives all

1In the voluminous literature on specific human capital accumulation
this bargaining, which is often not discussed explicitly, plays an
important role nonetheless. In particular, whether in the presence
of specific human capital there will be too many quits, too many
layoffs or both depends critically on the division of the surplus
between workers and firms (see Stiglitz (1974), Salop (1979),
Mortensen (1978), Jovanovic (1979), Hall and Lazear (1984), Parsons
pick wages, Mortensen (1978) studies arrangements which lead to
efficient separations while Jovanovic (1979) assumes the worker pays
for all his training and receives the value of his ex-post marginal
the surplus\textsuperscript{2}. If instead the workers can make credible take-it-or-leave-it offers to firms they receive all the surplus. The models we present require only that workers receive at least some positive fraction of the surplus. This is intuitively reasonable for two reasons even granting the ability of firms to acquire reputations of being unbending in their negotiations with workers. First, in the model we present below, firms are ex-ante indifferent between having a reputation for hard-nosed bargaining and having one for softness, the present value of their payments to workers is the same in either event. Second, in a richer model in which workers are heterogeneous, a flexible attitude in bargaining can elicit information about workers' true alternative opportunities and thus avoid inefficient separations.

Once one accepts that workers receive some of the surplus from their ability to outperform trainees we still must explain why this difference is related to industry variables such as profits and the capital/labor ratio. The latter turns out to be a consequence of assuming both that trainees are less productive than incumbent workers and that the technology has some putty-clay features so that the capital labor ratio is fixed ex-post. The role of this latter assumption is to restrict the firm to replace any given incumbent worker with a single less productive trainee. The difference in output lost is the surplus to be divided. This output naturally has higher value if it is produced with more capital. To make the role product. Parsons (1984) surveys some of this literature
\textsuperscript{2}When the firm is uncertain about the alternative wages available to individual workers (as in Stiglitz (1974) and Salop (1979)) it will
of capital clear consider the following example. Output is produced by combining people and tractors. Ex-ante the capital/labor ratio is variable since the firm can buy tractors of different sizes. Ex-post only one worker works with a given tractor. If trainees make the same number of mistakes independently of the size of the tractor, the cost of their mistakes will be bigger the larger are the tractors. Thus the surplus that incumbents can extract depends on the size of the tractors.

The role of profits in a model of this type is less clear because, for any profit maximizing firm, the marginal revenue generated by the last unit sold equals its marginal cost. This might be seen to imply that the most a worker can receive is a fraction of marginal cost which is independent of the profits of the firm. We present two reasons why some of the rents of the firm can be shared by workers nonetheless. The first is that ex-post, some workers have a very substantial effect on firm output (or profits). This would be the case, for instance, for individual computer software developpers working on a large project. The departure of any one of them, slows down the whole project substantially. When the fall in output is this large, the lost revenues may not be well approximated by the units lost times marginal revenue. For instance, in an oligopolistic industry, one would also have to take into account that marginal revenue rises as output contracts. The faster marginal revenue rises the more surplus the worker produces. Insofar industries with steeper marginal revenue curves also have offer some of this surplus as wages to prevent quitting.
higher profits this induces a correlation between profits and wages.

It must be noted that there is an important shortcoming of viewing each worker as bargaining individually with the firm after he has acquired specific human capital. This is that the sum of the marginal contributions of trained workers can easily exceed the total value of output. Then, bilateral firm-worker bargaining can yield negative profits for the firm. A more appealing formulation would take into account the mutual interdependence of workers. It would recognize that if other workers have bargained and obtained more, less is available for the rest of the workers. Unfortunately, we do not have good solutions to this n-player noncooperative bargaining problem. Suppose for instance that all workers simultaneously make take-it-or-leave it offers to the firm subject to the constraint that the firm closes unless it makes nonnegative profits. If the sum of the marginal contributions of workers exceeds the value of the output this game has a multiplicity of equilibria. This occurs because if I believe that many other workers are demanding their marginal contributions I will be content to demand less and keep the firm alive. Conversely, if I expect others to be modest in their demands, I will ask up to the value of my marginal contribution. Many such beliefs about others, together with each individual workers resulting offer, constitute equilibria.

Given the absence of a satisfactory noncooperative solution to this problem, we consider also a cooperative bargaining solution. The solution we consider is due to Shapley (1953). This solution has a variety of advantages. First, it is unique and explicitly
takes into account the mutual interdependence of workers. Since all players are treated symmetrically by this solution, it can be thought of as the expected value of what they would obtain in a variety of noncooperative settings. Finally it is a natural generalization of the Nash bargaining solution for two-person games, for which a noncooperative rationalization does exist. Using this solution, we show that workers, because they cannot all be always regarded as "marginal" workers, receive some of the inframarginal profits generated by the firm.

There is an additional advantage of our model using the Shapley value over that of bilateral bargaining. This advantage is that our current model of bilateral bargaining considers only one homogeneous class of workers which obtains more when it is coupled with more capital. This is a drawback because the perplexing feature of inter-industry wage differences is that they apply to a variety of different workers. One attractive feature of the Shapley value is that all workers share in the rents that they help generate. So employees involved in marketing (who perhaps are not directly associated with much capital) still collect some surplus that is increasing in the capital/labor ratio if they are needed to sell goods produced by surplus-generating production workers.

Both the bilateral bargaining model as well as the one based on the Shapley value must be regarded as models in which individual workers bargain individualistically. They thus stand in contrast with models based on unions in which workers engage in collective bargaining (for examples see Aoki (1980, 1982), McDonald and Solow
As we discuss further in our concluding section, this is an advantage because the pattern of inter-industry wage differences cannot be explained by unionization alone.

The models we consider predict that trained workers will obtain higher wages than untrained workers and that this difference depends on profits and the capital/labor ratio. Yet, industries with high wages will naturally attract workers. Thus, if the labor market functions well, the present value of payments must be equalized across industries. This naturally requires that in industries in which incumbents extract high rents, trainees must receive very low (or even negative) payment. It is easy to show that in this case, at least under some circumstances the input and output decisions of the firm are efficient. This results from the fact that the present value of the payments to the different factors equals their opportunity cost.

Even in this case firms with higher wages for trained workers will also have higher average wages because, as the future is discounted, the reduction in the payment of trainees is lower than the undiscounted increase in the wages of incumbents. Yet, in computing average wages the wages of trained workers are not discounted relative to those of trainees. Nonetheless, this version of the model has the prediction that differences in average wages can be explained entirely by differences in the slope of wages over time. There is at least some evidence, discussed below, that appears inconsistent with this implication.

If, for some reason the wages of trainees cannot be lower than
some value, be it the minimum wage or some "fair" fraction of the wages of incumbent workers one obtains the type of unemployment that has been discussed in similar models by Stiglitz (1974), Salop (1979), Lindbeck and Snower (1984) and in "efficiency wage" models by Shapiro and Stiglitz (1984) and Bulow and Summers (1986). These barriers to low initial wages prevent the equalization of the present discounted value of worker compensation across jobs making some jobs more attractive and generating queues for these jobs. We show that these barriers raise not only the wages of trainees but also those of incumbents since these can now only be replaced with more expensive outside workers. This mean that these barriers only exacerbate the differences in average wages we obtain in our model with market clearing wages.

The effect of capital-labor ratios and profits on average wages naturally has the consequence that estimates of profitability or of price-cost margins will be biased across industries. Heavily monopolized industries with high prices will tend to have high wages as well so that the usual measures of price-cost margins will understate the difference between price and social marginal cost.

The paper proceeds as follows. We start with models in which the labor market clears. In section II we present a simple model in which each worker bargains unilaterally with the firm. In section III we analyze mutually interdependent bargains. In section IV we consider barriers that prevent the wages of trainees from being low. Section V presents some conclusions and argues that the model is also capable of explaining the positive effect of establishment and
firm size on wages that is discussed in Brown and Medoff (1986).

II Bilateral Bargaining

We consider a firm which produces output with labor and capital. Its production function has constant returns while decisions are taken in two stages. Initially capital as well as the amount of capital per worker is chosen. In the second stage employment can be varied but the maximum number of employees that can usefully be employed is given by the amount of capital times divided by the predetermined amount of capital per employee. This means that capital can be thought of as tractors or computer workstations whose size is ex ante variable but which can only be used by at most one worker ex post. Letting $k$ denote the level of capital, $k/n$ the amount of capital per worker and $e$ the ex post level of employment such a production function can be written as:

$$q = ef(k/n), \quad e \leq n$$

(1)

Note that this production function is a standard constant returns to scale production function when $e$ is equal to $n$. In our model, we also distinguish between skilled, experienced "insider" labor and inexperienced "outside" labor. As in Lindbeck and Snower (1984) when "outside" workers are first hired they are trainees for one period and during this period their productivity is only a fraction $\alpha$ of that of the existing "inside" workers$^3$. It is usefull
to define the number of "effective" employees \( z \) which is given by
\[
[(n-r) + \alpha r] \quad \text{where } n \text{ is total employment and } r \text{ is the number of trainees employed.}
\]

Analogously to (1) we consider the ex ante production function given by:
\[
q = (n - r)f(k/n) + \alpha f(k/n) = zf(k/n) \quad 0 \leq r \leq n \quad (2)
\]

so that \( z \) corresponds to \( e \). Once again, the firm must choose ex ante both the level of capital and that of the capital/labor ratio.

Since both \( k \) and \( k/n \) are now fixed, \( k \) and \( n \) are determined at this point. Our interpretation of (2) is that the firm expects the capital stock to be used mostly (if not exclusively) by trained workers. This means that it expects output to essentially equal \( nf(k/n) \) unless an incumbent worker is replaced by a trainee.\(^3\)

\(^3\)This is superficially similar to the assumption of Shaked and Sutton (1984), that new workers can join the firm only at predetermined intervals of time. In both cases a certain amount of output is lost if the incumbent worker and the firm cannot reach agreement. The difference is that if the wage is subject to periodic renegotiation in Shaked and Sutton's (1984) model, the wage falls to the reservation wage every time the firm is actually capable of hiring outsiders.

\(^4\)In this formulation it is implicitly assumed that firms never let the total number of employees fall below \( n \). This is not an important restriction since our model has perfect foresight so that the firm knows how many employees it will hire when it chooses \( k \) and \( k/n \). It would thus never purchase capital which would be unused ex post.

\(^5\)There is a somewhat more general formulation for this ex ante problem that yields essentially identical conclusions. In this alternative formulation the potential use of trainees is contemplated more directly when capital is chosen. Suppose the firm is allowed ex ante to pick \( k \), \( n \) and \( r \) the anticipated number of trainees that will work with the capital stock. Let \( z \) be \( [n-(1-\alpha)r] \), the anticipated number of effective workers. Then supposing
Ex post it is only possible to change r taking k and n as given. Instead of using (2) directly as our ex post production function we sometimes use a slightly more general formulation which assumes that ex post output is given by:

\[ q = g(z) \]  

(3)

where \( g \) is a nondecreasing function. For the particular case considered above the function \( g \) is \( zf(k/n) \).

Workers have access to a perfect capital market and have a reservation wage of \( w^* \) per period. First, we assume that they live forever. Thus they are willing to join the firm as long as the present value of their compensation equals at least \( w^*/(1-\rho) \) where \( \rho \) is a market determined discount factor. Obviously, the least a worker would be willing to accept to work in any given period is \( w^* \).

We thus focus on the most the firm would be willing to pay, i.e. on the compensation of workers if they are able to make take-it-or-leave-it offers. This is done with the understanding that, for most reasonable bargaining games, the worker will receive only a fraction that the number of trainees will equal \( r \), output is \( zf(k/z) \). Ex post, however the number of trainees \( r \) might exceed \( r \). Then output equals \( [z-(1-\alpha)(r-r)]f(k/z) \). With \( r \) set equal to zero, this formulation is identical to that in the text. Its advantage lies in that it allows the firm ex ante to hire relatively less capital for trainees (who use capital less productively). The idea of the two production functions is the same. Ex ante workers are substitutable for capital but ex post capital per worker is fixed so that substituting trained workers by trainees reduces output by more the higher is the capital/labor ratio. The reason for choosing the formulation with \( r \) equal to zero is that, if capital is long lived it will mostly be used by trained workers and the firm will lose if it taylors its capital to the abilities of trainees.
of the difference between what the firm is willing to pay and $w^*$. With this proviso, let both potential trainees and insiders make take-it-or-leave it offers to the firm at the beginning of each period. These offers specify a wage at which each worker is willing to work. Each of these offers can be either accepted or rejected. If it is accepted the worker receives the required wage and becomes (or remains) an insider at the end of the period. If it is rejected, the individual worker is not employed and maintains (or regains) the status of outsider.

Let the wage of trainees be denoted by $w_1$, that of insiders by $w_2$, $w_3$, $w_4$ ... where the subscript denotes the period of employment and $R(q)$ be the revenues of the firm when it sells $q$ units of output. Then the following proposition summarizes the equilibrium values of these wages:

**Proposition 1**

The unique perfect Nash equilibrium of the game has wages given by:

$$w_1 - w^* = \rho \beta$$  \hspace{1cm} (4)

$$w_i - w^* = (1 - \rho) \beta \quad \quad \quad i \geq 2$$  \hspace{1cm} (5)

where

$$\beta = R[g(n)] - R[g(n-(1-\alpha))]$$  \hspace{1cm} (6)
\[ z = (1-\alpha)[R'g' - (1-\alpha)[R''g'^2 + R'g']/2] \]  

(7)

and primes denote derivatives evaluated at \( r \) equal to zero.

**Proof:**

Outsiders will always lower the wages they demand during the traineeship until the present discounted payments of lifetime employment is \( w^*/(1-\rho) \). So suppose that insiders will demand a profile of wages whose present value at time \( \tau \) is \( S_\tau \). Paying this present value to an insider forever instead of turning to an outsider is acceptable to the firm only if

\[ S_\tau - w^*/(1 - \rho) \leq \beta = R[g(n)] - R[g(n-(1-\alpha))] \]  

(8)

that is if the difference in the present discounted value of wages equals at most the current difference in revenues obtained (since in future periods the revenues are the same). Since insiders make take-it-or-leave-it offers they demand a wage that makes (8) hold with equality. This equation implies that \( S_\tau \) is constant over time which implies that wages of insiders are constant over time. Thus without loss of generality we can write \( S_\tau \) as \( w_2/(1-\rho) \) and equation (5) follows. Equation (4) follows immediately from the requirement that the present discounted value of wages equal \( w^*/(1- \rho) \)
that the typical production worker reduces output so much when he is replaced by a trainee that marginal revenue itself is affected. On the other hand, there are clearly some production and some clerical workers who are pivotal in any organization and who reduce output significantly when they depart. For them, as well as for the members of a smoothly functioning string quartet, the \( R'' \) term may be important and its relevance even for only a minority of workers may end up reflected in the average wage.

We now turn to a discussion of the effect of \( R' \) on the wage differential. To do so we consider the ex ante decision of firms. For simplicity we focus on steady states in which \( R' \) is constant over time.\(^6\) Also we concentrate on the case in which (2) gives the ex-post production function. Let the firm add one additional worker as well as \( k/n \) units of capital forever. Thus the capital/labor ratio remains the same. The present value of the cost of these additional inputs is

\[
(w^* + vk/n)/(1-p)
\]

where \( v \) is the rental rate of capital. Since the capital/labor ratio remains the same, output increases by \( af(k/n) \) the first period and by \( f(k/n) \) thereafter. So, the present value of the revenues these inputs generate is:

\[\text{6Thus we are neglecting for instance the dynamics associated with the first period in which all the workers are trainees, output is low and marginal revenue is correspondingly higher than in later}\]
Equation (7) results from a second order Taylor expansion with respect to r. QED

Note that more generally, when \( w_2 - w^\infty \) equals approximately \( (1- \rho) \chi(1-\alpha)[R'g'-(1-\alpha)[R"^2 + R'g"]/2] \), the difference between \( w_2 \) and \( w_1 \) is given by:

\[
w_2 - w_1 = (1-\alpha)\chi[R'g' - (1-\alpha)[R"^2 + R'g"]/2] \tag{9}
\]

Several comments deserve to be made about (5), (6) and (9). First, the premium earned by incumbent workers over the reservation wage is proportional to the premium they earn over the trainee wage, the latter is larger because trainees must "pay" for their jobs by earning less than \( w^\infty \). Second, the larger are \( (1-\alpha) \) and \( g' \), i.e. the larger is the fall in output from replacing an incumbent worker with a trainee, the larger is the wage premium earned by incumbent workers.

Now consider the second order term and ignore \( g" \). Under perfect competition \( R' \) is equal to the competitive price and \( R" \) is zero. For a firm with market power \( R" \) is negative and gives the rate at which marginal revenue decreases as output increases. If the firm perceives a linear demand curve for its product, \( R" \) gives the absolute value of its slope. It is thus natural to expect that \( R" \) will be larger in absolute value the more monopolistic the firm is and that \( R" \) will therefore be correlated with the profitability of the firm. It must be pointed out that the relevance of this second order term is unclear. On the one hand it is difficult to imagine
\[ \frac{\alpha + \rho}{1-\rho} f(k/n) R' \]  

(11)

If the firm is maximizing profits, (10) must equal (11) in a steady state so that

\[ R' = \frac{(w^* + vk/n)}{f(k/n)[\rho + (1-\rho)\alpha]} \]  

(12)

which means that, if we can neglect the second order terms, the premium received by incumbent workers is proportional to \((w^*+vk/n)\); in particular (6) becomes:

\[ w_2 - w^* = (1 - \rho)(1-\alpha)(w^* + vk/n)/[\rho + (1-\rho)\alpha] \]  

(13)

and is thus increasing in the capital/labor ratio and independent of the form of the function \(f\). It is worth noting that (13) does not prove that a single firm, by raising the capital/labor ratio would raise wages of incumbent workers although this follows immediately from (6) and (7) as \(f\) (and therefore \(g'\)) goes up when \(k/n\) goes up. Instead (13) proves that different firms with different production functions will pay higher wages the higher is the optimized value of their capital/labor ratio. We show below that under certain circumstances this optimized value is actually efficient.

The reason for this result is that one can always define the unit of output as the amount produced by one worker and the amount periods.
of capital given by the capital/labor ratio. The marginal revenue of the unit so defined must equal its marginal cost which depends only on \( w^* + vk/n \). Replacing one incumbent worker by one trainee leads to the loss of \((1-\alpha)\) units so defined hence the surplus is proportional to \((1-\alpha)(w^* + vk/n)\). From (13) one can conclude that unless the second order effects are important, only the capital/labor ratio and not the level of profits affects wages.

**Efficiency of the capital labor ratio**

It might be thought at this point that since insiders capture some of the compensation due to capital they would bias downward the capital labor ratio of firms. This does not follow necessarily from a model of this type as can be seen by considering the ex ante decision problem of firms. At the moment firms pick capital and the capital labor ratio they know that the present value of its expenditures per employee will be \( w^*/(1-\rho) \). Thus the optimal choice of \( k/n \) satisfies:

\[
\frac{w^*f'}{v(\rho + (1-\rho)\alpha)(f - f'k/n)} = 1
\]

which is the same as would obtain if workers could precommit to receiving \( w^* \) per period. The possibility that the capital/labor ratio is efficient should not be overstressed however since it relies heavily on a single ex-ante capital/labor choice. More realistically some of the capital is purchased after some trained
workers are on the premises. This may have the implication that capital will be chosen taking the dependence of wages on capital intensity into account and that capital choice will be distorted.\(^7\)

**Putty-clay vs. Putty-putty**

At this point it is worth commenting on the importance of the assumption of putty-clay technology. If the technology were putty-putty, both the ex-ante and the ex-post production functions would be given by:

\[
q = zf(k/z)
\]

which corresponds in (1) to allowing \(e_1\) and \(e_2\) to be picked ex post subject to the constraint that they be equal to each other. In this case, even with a fixed capital stock, the loss of output from replacing an incumbent worker with a trainee would only be \((1-\alpha)(f-f'k/n)\) instead of the \((1-\alpha)f\) which occurs in our putty-clay model. The loss is smaller because capital can be put to more productive uses by combining it with the more productive incumbent workers. Moreover, we now argue that if the same substitution is possible ex post as ex ante the capital labor ratio plays no role in the premia collected by experienced worker. This is so because, by an argument analogous to the one used above, the marginal revenue produced by

\(^7\)A similar ex-post inefficiency may arise with respect to labor turnover. This turnover may be inefficiently low due to the presence of high wages after the worker is trained.
\[ \frac{\alpha \rho}{(1-\rho)} (f-f'k/n) \] units of output, the extra output produced by
one extra worker that is kept forever, must equal their marginal
cost which equals \( w^*/(1-\rho) \). This means that the loss in output from
replacing one incumbent worker by a trainee, \((1-\alpha)(f-f'k/n)\) leads to
a loss in revenue proportional to \( w^* \) and independent of \( k/n \).

**Average wages and finite horizons**

So far we have concentrated only on wage premia for experienced
workers rather than on average wages. Of course with infinitely
lived workers, the average wage is essentially equal to the wage for
experienced workers. We now show that industries with high wages
for such workers also have high average workers even if we
arbitrarily assume that workers remain with the firm only \( T \) periods.

Suppose that, once an employee joins the firm, he remains with
the firm only \( T \) periods. For simplicity we revert to the technology
in (1) in which first order approximations to the loss in revenue
from switching workers are sufficient. Again, let workers make
credible take-it-or-leave-it offers.

**Proposition 2**

If workers join the firm only for \( T \) periods, the unique Nash
equilibrium has constant wages for trained workers equal to \( w_2 \).
These as well as the wages of trainees satisfy:
\[ w_1 - w^* = -(\rho - \rho^T)\beta/(1 - \rho^T) \]  \hspace{1cm} (14)

\[ w_2 - w^* = (1 - \rho)\beta/(1 - \rho^T) \]  \hspace{1cm} (15)

**Proof**

Outsiders will now bid down the entering wage \( w_1 \) until the present discounted value of compensation equals the present value of the foregone reservation wage \( w^*(1-\rho^T)/(1-\rho) \).

Insiders who expect to work for another \( j \) periods will remain employed can, at most expect to bid a path of wages whose present value \( S_j \) leaves the firm indifferent between keeping them and replacing them with a sequence of new workers. For the firm to be indifferent:

\[ S_j - w^*(1 - \rho^j)/(1 - \rho) \leq \beta(1 - \rho^j)/(1 - \rho^T) \]  \hspace{1cm} (16)

where the LHS of (13) represents the extra disbursements while the RHS represents the gain in revenue. For \( j \) equal to 1 (16) gives (15). To demonstrate that this is true for all \( j \) one proceeds by backwards induction. To obtain (14) it is enough to remember that the present discounted value of payments to a worker \([w_1 + (\rho - \rho^T)w_2/(1-\rho)]\) must equal \( w^*(1 - \rho^T)/(1-\rho) \) and apply this to (15). QED

Note that the premium paid to experienced workers rises as \( T \) becomes shorter because the strategy of replacing current experienced workers with trainees means that the replacement for the
trainees themselves will have to occur sooner. This makes this strategy more expensive and increases the value of existing workers.

To compute average wages we assume that the number of workers leaving the firm is constant over time. Then average wages \( w \) are given by:

\[
w = \frac{[w_1 + (T-1)w_2]}{T}
\]

\[
w - w^* = \beta [(T-1)(1-p) - (p-p^T)]/[T(1-p^T)]
\]

That this expression is increasing in \( \beta \) for \( T \) greater than one can be seen as follows. For \( T \) equal to 2, it is equal to \( \beta (1-p)/2(1+p) \) and thus increasing in \( \beta \). Moreover the numerator of (14) is increasing in \( T \) and the denominator is positive so the expression multiplying \( \beta \) remains positive for all \( T \).

Comparison with efficiency-wage models

The reason why workers obtain more than their reservation wage in this model is that they are more valuable to the firm than outside workers. Instead, in efficiency-wage models with workers of homogeneous productivity wages are high to elicit a change in behavior by existing workers\(^8\). This change in behavior changes the "efficiency" of workers. It does this either by directly increasing effort (Solow (1979)), reduce shirking (Shapiro and Stiglitz (1984),

\(^8\)Efficiency wage models also include adverse selection models (Weiss
or reduce turnover (Stiglitz (1974, 1985), Salop (1979)). In the first of these models the wage premium depends only on the elasticity of supply of effort with respect to the wage and the capital/labor ratio has no role. The shirking model of Shapiro and Stiglitz also has the property that all firms for which efficiency wages are an issue pay a wage that prevents shirking independent of firm characteristics. This conclusion probably depends in part on the fact that the substitutability of effort (or lack of shirking) for number of employees is modelled as independent of the capital stock.

On the other hand the turnover-based efficiency wage models are closely related to the model of this Section. In Stiglitz (1974, 1985) there is a training cost $T$. In contrast to the model presented so far, these models assume that quitting by workers is a nondegenerate decreasing function of the wage they are paid. These models also assume that trainees and experienced workers must be paid the same wage and conclude that wages can exceed the market-clearing level. The reason for this is that each firm raises the wage it pays until the higher payroll costs are balanced by reductions in quits which translate into reductions in training costs. Moreover, if the second order conditions are satisfied an

(1980) in which workers differ in their true (unobservable) ability.

9On the other hand a model in which this substitutability depends on the capital stock is really not very different from a model in which workers are of differing qualities and high quality workers have a comparative advantage in the production of goods with high capital/labor ratios.

10Salop (1979) assumes that the training costs are an increasing convex function of the number of trainees which makes it difficult to analyze the cross-industry effects of different training costs.
Increase in \( T \), which increases the value of keeping incumbent workers raises wages.

With the putty-clay technology assumed here, \( T \), the training cost equals \( (1-\alpha)f(k/n) \) which, as we have shown is increasing in \( k/n \). Thus their turnover model can be adapted to yield higher wages for more capital intensive industries. This result would emerge even if workers were allowed to pay for their training because, as above, keeping the present value of wages constant, average wages rise when incumbent's wages rise.

Thus the putty-clay technology implies high wages in capital intensive industries even when firms unilaterally set the wage, as long as quits fall smoothly with increases in wages. Our model assumes instead that workers all quit if they are paid less than \( w^* \) while they remain otherwise. In this case, workers must have some bargaining power for wages to increase with capital intensity.

### III Mutually Interdependent Bargains

The model of bilateral bargaining presented above assumes that while each individual worker bargains with the firm (and threatens to quit) he expects all the other workers to remain with the firm. While this may appear empirically reasonable its logic is somewhat suspect. A more symmetric treatment of the threats of each individual worker would recognize that other individuals are also threatening the firm and that, were they to leave, the individual's
bargaining position would be altered. To put this differently, suppose that some of the other workers in the firm in fact quit, then the additional departure of an individual worker is more costly to the firm as long as \( R'' \) is negative. This means that the individual can, in principle, demand to be paid taking into account this eventuality.

Consider for instance an Economics Department. Having one experienced econometrician is clearly essential to guarantee the quality and continuity of the Department's activities. Even if the department has three econometricians each one ought to be able to capture some of the surplus he would provide if the other two were to leave.

To model this ability to capture some inframarginal rents we model a firm with \( n \) incumbent workers as an \((n+1)\) person game in which the firm (associated with capital) is one of the players. Since no convincing extensive forms for games of this kind have been worked out we use an axiomatic solution to this game. In particular, since the game is well described by its characteristic function (see Shubik (1982)), the Shapley value (Shapley (1953)) is a natural solution concept.

A second problem with the bilateral approach of the previous section is also solved by the Shapley value. This problem arises when \( g' \) is very large, so that the departure of any one worker is very costly. Then the sum of all the payments to workers using our bilateral bargaining approach may exceed the revenues of the firm. In some sense this problem is due to the fact that workers are only
bargaining with the firm instead of bargaining with other workers as well. This is naturally taken care of when the n workers and the firm are being viewed as bargaining at once.

We simplify the analysis somewhat by considering only a single period and considering the ex-post technology (1) with the function g given by the product of f(k/n) times its argument. The firm sells its output in an imperfectly competitive markets and its perceived demand curve is given by:

\[ q = a - bp \] (18)

where p is the price for the good.

Total revenues from employing r trainees and (n-r) experienced workers is:

\[ R(r) = [a - f(n-(1-\alpha)r)][f(n-(1-\alpha)r)/b \] (19)

The characteristic function, denoted v, is a function from subsets S of players to real numbers which gives the maximum revenue that the subset of players can obtain by cooperating. Since all the incumbent workers are symmetric in our example we consider only coalitions with j workers with capital, for whom the value of the characteristic function is v(k,j) and coalitions of j workers without capital for which it is v(0,j). Consider first the latter. The most this coalition can obtain is jw since their special skills can only be exercised together with the capital they are familiar
with.

Now consider the coalition of capital with \((n-r)\) experienced workers. This coalition will find it in its interest to hire \(r\) trainees and pay \(w^*\) to each. It thus obtains the revenue given by the expression in (16) minus \(rw^*\). The marginal revenue from replacing an additional incumbent worker with a trainee can be approximated (to first order) by:

\[
\Delta = -m - \xi r
\]
\[
m = (1-a)f(a - 2n)/b
\]
\[
\xi = 2[f(1-a)]^2/b
\]

where \(r\) is the current number of trainees. Note that \(m\) is \((1-a)\) times the marginal revenue associated with one additional trained worker and \((k/n)\) units of capital. By the argument of the previous Section it is equal to \((1-a)\) times the marginal cost of these inputs. Thus for \(r\) equal to zero, this approximation neglects the second order terms considered in Section II. The approximation used in (17) can also be used to approximate (16) for any value of \(r\). It gives

\[
R(r) = v(k, n-r) = R(n) + \sum (m + \xi) - rw^*
\]  
\[
i=r
\]

where (21) is used to define the value of the characteristic
function associated with any coalition of capital together with \((n-r)\) incumbent workers.

Now that we have expressed the characteristic function we can calculate the Shapley value for each player \(i\). This value can be written, for example, as \(\phi_i\):

\[
\phi_i = \sum_{j=1}^{n+1} \left( \frac{1}{c(j)} \right) \sum_{j=1}^{n+1} \frac{v(S) - v(S-\{i\})}{(n+1)}
\]

where \(v(S)\) is the value of the characteristic function associated with the coalition \(S\). Thus each player receives the average of his marginal contributions to all the coalitions to which he contributes. To compute this value we use the procedure proposed by Shapley (1953). This procedure considers all \((n+1)!\) ways of ordering the players. For each ordering it attributes to any given player \((1/(n+1)!\) of the contribution he makes to the coalition of the players that come before him.

We obtain \(\phi_w\) for workers and \(\phi_k\) for capital:

\[
\phi_w = w^* + m/2 + \xi(n-1)/6
\]

\[
\phi_k = R(n) - nw^* + nm/2 + n(n-1)\xi/3
\]
When workers come "before" capital they receive $w^*$; when they come "after" they contribute $w^* + m$ even when $\xi$ is zero. The average thus always includes $w^*$. Since each worker comes before capital half the time and after half the time $m$ is split 50/50 with capital. Even without any incumbent workers, capital earns $R(n) - nw^*$ by using $n$ trainees, it is thus guaranteed at least this amount.

The term $\xi$ captures the increased marginal revenue of inframarginal units. In particular it is higher for steeper demand curves and zero for horizontal demand curves. Workers as a whole receive only 1/6 of the total revenues that depend on $\xi$ while capital receives 1/3. The reason for this uneven splitting is that whenever capital comes after at least some workers it receives the revenues of the early, most valuable, units.

One interesting feature of (23) is that as $n$ rises the compensation of each worker rises proportionally to $n\xi$. This may seem paradoxical since it says that a larger level of employment leads to higher wages. It must be kept in mind however that $n$ represents the optimized value of employment. A higher $n$ thus means that the demand curve of the firm has a higher intercept which leads it to hire more workers. This larger demand naturally translates into more equilibrium surplus some of which goes to the workers.

Since firms with a higher $\xi$, which can be interpreted as degree of monopoly power, have both higher profits (compensation to capital) and higher wages, our model predicts a positive correlation between profits and wages. This means that the level of accounting
profitability may considerably understate the level of actual profits.

At this point the reader may wonder what is special about capital. Aren't we simply asserting that the payment to one factor depends on the worth of the other factors with which he is combined? If so, why should this worth vary across goods? The answer to this question is that it is capital which becomes less productive if combined with trainees while each individual worker has the same option as any trainee, namely to leave and earn his reservation wage. This does not mean that incumbent workers need to be interact with capital directly in order to receive premia for their experience. It is enough that they enhance the productivity of other workers who do interact with capital. To see we turn to a simple three agent example.

The three agents or factors we consider are capital, a worker denoted m for manager whose reservation wage is \( w_m \) and a worker denoted s for secretary whose reservation wage is \( w_s \). Trainees for each are available at the same wages. We assume the characteristic function of the game is given by:

\[
\begin{align*}
v(k,m,s) &= p \\
v(k,m) &= \beta p - w_j \\
v(k) &= \alpha p - w_m - w_j \\
v(m) &= w_m \\
v(m,j) &= w_m + w_j \\
v(k,j) &= \alpha p - w_m \\
v(j) &= w_j
\end{align*}
\]

where \( p \) can be thought of as the value of the output. In this
example the replacement of incumbent s by a trainee does reduce output by \((1-\beta)\). This however can be interpreted as a loss in the output of m because when m is replaced by a trainee output falls to a regardless of whether s is an incumbent or a trainee. Note that in this example the sums of the marginal contributions of m and s, \((2-\alpha-\beta)\) can easily exceed 1 which makes the solution proposed in Section II problematic.

The values of the game to k, m and s are given by \(\phi_k\), \(\phi_m\) and \(\phi_s\) respectively:

\[
\begin{align*}
\phi_k &= v(k) + p[(1-\alpha)/2 - (1-\beta)/6] \\
\phi_m &= v(m) + p[(1-\alpha)/2 - (1-\beta)/6] \\
\phi_s &= v(s) + p(1-\beta)/3
\end{align*}
\]

The reason why s earns more than his reservation wage is that when he comes in last (which happens in \(1/3\) of the orderings of players) he receives \((1-\beta)p\). This compensation to s is shared by k and m equally as is the surplus \((1-\alpha)p\) from having an incumbent team. Note however, that the dollar surplus of s depends on p which, once again, depends in equilibrium on the amount of the factor without alternative uses k.

IV Constraints on Trainee Wages

So far we have assumed that the labor market clears in the sense that trainees upon joining the firm can expect to earn no more
than the present value of their reservation wages. This however requires that trainees earn very little in high wage industries, a pattern that is not borne out by any available evidence.\textsuperscript{11} This is a particularly severe problem under one natural reinterpretation of our model. In this reinterpretation trainees are no less productive than incumbents but (1-\alpha) units are lost in the transition, perhaps because it takes time to recruit trainees. This would imply that all wages are the same but that new hires would have to make a lump sum payment to the firm, much as in efficiency-wage models\textsuperscript{12}.

The extent to which trainees pay for high paying jobs is to some extent an open question. Yet, the empirical evidence on tenure, while not narrowly focused on this question, does not support the notion that trainees earn particularly little in high wage industries. First, Abraham and Farber (1986) find the effect of tenure on wages to be relatively small once one controls for the ultimate length of the job\textsuperscript{13}. Second, Katz (1986) reports that

\textsuperscript{11} An alternative which may be more realistic in some environments is for trainees to work hard at a host of more menial activities in addition to performing the tasks in which they are replacing a (more productive) incumbent. The value of these more menial activities would then have to be subtracted from the trainee wage to obtain our measure of \( w_1 \).

\textsuperscript{12} In efficiency-wage models it is often argued (Akerlof and Katz (1986)) that capital market imperfections prevent such lump-sum payments. In the presence of these imperfections efficiency-wage models tend to imply that firms can gain by paying entering workers low salaries and making them perform tasks for which efficiency-wages are not important. Thus insofar new hires do not receive low wages in high wage industries, some doubt is cast on efficiency wage models as well.

\textsuperscript{13} Admittedly, for a large fraction of the sample, the ultimate length of the job is measured to be an increasing function of actual tenure. Their paper is closely related to Altonji and Shakotko (1985) who show that there if one concentrates only on the effect of the growth of tenure on wages (and ignores the effect of the level
inter-industry wage differences for workers with less than one year of tenure and those with more than ten years are highly correlated.

In this section we consider briefly the behavior of inter-industry wages in our model if for some reason it is impossible to pay entering workers low wages. As is apparent from the analysis of Salop (1979) and Lindbeck and Snower (1984) such constraints also induce unemployment. Another obvious consequence is that, as these constraints make labor more expensive they distort the capital/labor ratio.

We consider two such constraints on entering wages. The first is that entering wages must equal, perhaps for legal reasons, at least some minimum level $w$. The second is that, perhaps for sociological reasons, the entering wage must equal at least a fraction $\lambda$ of the wage paid to incumbent workers. In both cases we conclude that the imposition of this constraint raises wages for experienced workers. Moreover, average wages continue to be higher the higher would have been wages for experienced workers in the absence of the constraint. These increases in wages naturally lead to the question of the what other constraints prevent wages from growing without bounds. Two such constraints are discussed at the end of this section.

For purposes of this discussion we adopt again the analysis of the beginning of Section II, though little would be altered if we considered the multilateral bargaining of Section III. Before computing equilibrium wages under the two constraints we it is

of tenure on wages) one finds only small effects.
useful to prove the following lemma that relates wages of experienced workers to the present value of payments to new workers W.

Lemma 1

Suppose new workers earn a present value of W. Then, if they make take-it-or-leave-it offers to the firm, wages of experienced workers, w₂, are at most given by:

\[ w₂ = (1-p)(β + W) \]

where β, is given in (7).

Proof

The firm can either keep the insider or hire an outsider who earns W in present value. It will keep the insider only if:

\[ \frac{w₂}{1-p} - W \leq β. \]  \hspace{1cm} (24)

So that an insider who makes take-it-or-leave-it-offers he demands a wage given at most by (24). QED

Suppose that, starting in the second period of employment workers earn \( w₂ \). Then, whatever the value of the wage in the first period, \( w₁ \), W is given by \( [w₁ + \frac{p}{1-p}] \) so that, at most:
\[ w_2 \leq (\beta + w_1)/(1-\rho). \] (25)

Three conclusions emerge from (25), assuming it holds as an equality. The first is that increases in \( w_1 \), due for instance to increases in \( w \), raise \( w_2 \) and thus raise average wages. The second is that increases in \( \beta \) continue to be reflected in higher average wages. Finally, we see that even constraints on \( w_1 \), do not eliminate the prediction of the model that there must be a difference between \( w_2 \) and \( w_1 \) and that the difference between \((1-\rho)w_2\) and \( w_1 \) must be increasing in \( \beta \).

Now suppose that the entering wage is given by \( \lambda \) times \( w_2 \). This gives a wage for experienced workers of at most:

\[ w_2 = \beta/(1-\rho-\lambda) \] (26)

which is finite only if \( \lambda \) is less than \((1-\rho); otherwise incumbent workers are always cheaper than trainees and can thus demand up to an infinite wage. Again (23) implies higher wages the higher are \( \lambda \) and \( \beta \).

Since (22) and (23) can give rise to very large wages indeed we now discuss two limits on wages. First, wages of workers can never exceed their ex-post value marginal product \( R'f(k/n) \) which is of course increasing in \((k/n)\). So one possibility that is consistent with this analysis is that \( \lambda \) is one and that both trainees and incumbent workers receive some fraction of \( R'f \) with the other
fraction going to the firm. In this case there would, of course, be no slope in the wages received by workers over time.

Second, the firm might hire some workers as "supernumeraries" whose main function is to be already trained (or semitrained) and thus reduce the bargaining power of existing workers. The presence of these workers has the effect of raising $\alpha$ and thus lowering $\beta$ and the wage firms must pay to existing workers. Yet the analysis of the equilibrium with such supernumerary workers is a far from trivial task.

For instance, suppose that the firm has two workers for each of the $n$ tasks. What should each worker demand if he can make take-it-or-leave it offers? Suppose there is a last period after which the firm will cease to exist. Then the most that each worker will demand, taking the other worker's offer as given is $w^*$; they are effectively competing in Bertrand fashion. This should not be viewed as predicting that wages to a single worker are constrained to equal at most twice the reservation wage. To see this, consider the period before the last period. Even if one worker is demanding $w^*$ the other can demand more than $w^*$. He can do this because he knows that if the firm fires him at this point it will have to pay more to the remaining worker in the last period. The firm is willing to pay a premium above $w^*$ to the supernumerary worker in order to maintain the threat against the other worker.

Nonetheless we expect that in a realistic model of this kind, some supernumerary workers will be hired, particularly if these are able to contribute to production somewhat. A richer model might
have firms with high capital/labor ratios and high rents spending more on labor both through high wages and through excessive employment.

V Conclusions

The basic message of this paper is that the pattern of inter-industry wages appears consistent with a model in which three conditions are met. First, trainees are not as productive as incumbent workers. Second, technology has some putty-clay features. Third workers have some bargaining power so that the wage is somewhere between the wage a firm would quote if it could make take-it-or-leave-it offers and the worker would quote if he could. While replacing this third assumption with the requirement that quits are a smooth function of the wage also implies that capital/labor ratios are correlated with the wage, our form of bargaining is probably required to obtain correlation between wages and profits.

This naturally leads to the question of the relationship between our model and models based on collective bargaining. One natural advantage of considering only individualistic bargaining as we do is that the public goods problem inherent is the formation of a union is avoided. On the other hand some workers are clearly

14In Aoki (1980, 1982) this problem appears to be circumvented by letting the "representative (incumbent) worker" bargain in a two-person game with management. Yet, unlike what is true when each worker bargains individually, he does not take into account that the demands of other employees alter the opportunities of any given employee. Aoki requires instead that the representative worker assume, when bargaining, that other workers will act exactly as he
capable of overcoming these problems. In models in which unions maximize some measure of member welfare it is of course not surprising that workers can extract some monopoly rents or even some of the quasi-rents produced by capital. The difficulty with ascribing wage differences to collective action on the part of workers is that it is difficult to imagine collective bargaining in the absence of institutions such as unions. Yet, most U.S. workers, including many high wage workers are not unionized.

This has led Dickens (1986) to postulate a model in which the threat of unionization can lead to high wages. The idea is that workers in some sense prefer to avoid the coordination costs of forming a union as long as they do not earn wages much below those they could obtain if they did organize a union. Firms, in response, will take actions that prevent union formation. This includes either raising wages or making it more difficult to form a union by raising employment or both. Thus this model has the potential of explaining why high wherever unions could obtain high wages. The empirical difficulty with this model is that, as shown in Dickens and Katz (1986) high wage industries also pay high wages to managers and professionals who, in the US, are very unlikely to unionize. Similar difficulties arise when attempting to explain the relatively high correlation of inter-industry wage differences across countries in which unionization is easy and the Eastern Block countries in which it is essentially impossible (see Krueger and does. This ensures that workers obtain what is in their collective self interest much as conjectural variations equal to 1 bring about the monopoly outcome in oligopolistic markets.
Summers (1976b) who report a correlation between log wages of .78 between Germany and the USRR). It would seem that our model which requires only bargaining with the individual worker, plus a modicum of self interest on the part of the firm, is better suited to Eastern Block countries.

Our paper has focused exclusively on inter-industry wage differences yet there is one other source of wage differences which appears just as difficult to reconcile with standard models. This is the widely documented effect of employer size on wages which is discussed in Brown and Medoff (1986). They show that both the size of the establishment and the employer for which a worker works are positively correlated with wages. They argue that these effects actually occur within certain industries are are thus not directly due to inter-industry wage differences.

To explain this fact one would naturally need to know first why establishments and companies within industries differ in size. If all firms are competitive and have access to a constant returns to scale technology such as (2), firm and establishment size are indeterminate and nothing ought to depend on these sizes. Now suppose that there is market power in the industry. Clearly large firms now have larger surplus but not necessarily larger surplus per employee. Suppose however that the products of the firms are actually differentiated. Then our model based on the Shapley value has the potential for explaining these size effects. We have shown that if the slopes of the demand for various goods produced with the same technology are the same, larger product lines generate more
surplus for each worker. Brown and Medoff (1986) argue that larger firms actually have a smaller elasticity of demand. Yet, their estimates suggest a steeper slope for larger firms which is consistent with more surplus in these firms.

A second explanation for differences in company and establishment size is the presence of some entrepreneurial ability which has diminishing returns. This means that there are inframarginal rents which are bigger in larger firms (establishments). In a model based on the Shapley value workers will extract some of these rents so they will be paid more in those firms (establishments). It must be noted that both these explanations are consistent with the collection of facts Brown and Medoff (1986) present to reject other theories of the size effect. In particular, workers in these high paying jobs will quit less and they will lose their high wages if they go to smaller firms or establishments.

REFERENCES


Becker, Gary S. Human Capital, New York, NBER, 1964


_________ and Katz, Lawrence: "Industry Wage Patterns and Theories of Wage Determination" unpublished 1986a

_________: "Inter-industry Wage Differences and Industry Characteristics" unpublished 1986b


Katz, Lawrence F.: "Efficiency Wage Theories: A Partial Evaluation" forthcoming

Krueger, Allan B. and Lawrence H. Summers: "Efficiency Wages and the Inter-Industry Wage Structure" unpublished 1986a


Salop, Stephen, :"A model of the Natural Rate of Unemployment", American Economic Review 69, March 1979, 117-26

Shaked, Avner and John Sutton: "Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model" Econometrica, 52,
November 1984, 1351-65

Shapiro, Carl and Joseph E. Stiglitz: "Equilibrium Unemployment as a Worker Discipline Device" American Economic Review 74, (June 1984), 433-44


