A Multisector Model of Efficiency Wages

Frank Walsh, University College Dublin

The pattern of effort and wages is derived in a multisector efficiency wage model. Firms choose effort endogenously. Easily monitored or low-turnover jobs have high effort and may have low wages in equilibrium. Empirical wage differentials from a measure of supervision are smaller than observed industry differentials that have been attributed to efficiency wage models and are closer to those predicted by the model. Workers can search for and avail of on-the-job offers. If sectors grow at different rates or the unemployment rate changes, the pattern of wage differentials is unaffected.

I. Introduction

Most of the empirical literature on efficiency wages concentrates on wage differentials across sectors. However, theoretical efficiency wage models typically have only one or two sectors and do not provide an adequate theoretical framework to examine interindustry wage differentials.¹ The theoretical model outlined below, as well as empirical results from a measure of monitoring, indicate that it is a mistake to attribute large industry wage differentials to this class of efficiency wage models.

I embed Shapiro and Stiglitz’s (1984) efficiency wage model in a

multisector framework and endogenise the effort decision using Solow’s (1979) model where firms choose effort optimally. These generalizations have important consequences for the predicted pattern of equilibrium wages and effort across sectors, for the pattern of wage differentials over the business cycle, and as industries grow at different rates.

The standard prediction of the monitoring/turnover efficiency wage model has been that badly monitored or high-turnover sectors pay higher wages.\(^2\) By contrast, I show that whether higher or lower wages are chosen by firms with more intensive monitoring of workers or lower turnover depends on the shape of the worker’s effort supply curve. In particular, the relationship between wages and monitoring intensity or turnover depend on whether the elasticity of the disutility of effort with respect to effort is increasing or decreasing.\(^3\) If the elasticity is decreasing/increasing, firms with more intensive monitoring or low turnover will pay lower/higher wages and will always have higher effort. It is only when the elasticity is increasing that we will observe higher effort in the high-wage sectors. The intuition is that in sectors with more intensive monitoring it is cheaper to elicit effort, firms choose higher effort, and depending on the shape of the effort supply curve, may do this up to a point where the wage is higher than in sectors with less intensive monitoring. This point has often been lost in single sector models where the assumption is that empirically we should expect high wages to be associated with high effort and higher worker productivity.\(^4\)

Simulations of the model allow a direct comparison between the predictions of the model and empirically observed industry wage differentials. The simulations show that the model does not predict the large industry wage differentials observed in empirical studies such as Krueger and Summers (1988). The ratio of supervisors to workers as a measure of supervision predicts smaller wage differentials than observed across industries. These supervision wage differentials are closer in size to those predicted by the simulations of the theoretical model.

Letting employment grow at different rates across sectors leaves the pattern of wage differentials unchanged.\(^5\) A fall in unemployment increases wages in all sectors by the same amount, so wage differentials do not change over the business cycle, although wages in all sectors are procyclical.

The rent associated with a job in any sector is increasing in effort.

---

\(^2\) See Katz (1986) for a survey of these models.

\(^3\) Note this is not the elasticity of effort with respect to the wage that is often referred to in the efficiency wage literature and that I will come to below.


required and decreasing in monitoring intensity. High-turnover sectors have higher wages to compensate for the increased probability of job termination, but rents are not higher in these sectors, and empirically we should not expect job queues to depend on turnover even though wages may.\footnote{In other versions of the efficiency wage model (see, e.g., Holzer, Katz, and Krueger [1991]) firms with higher turnover costs may pay a wage premium to reduce turnover. These premiums will be rents.}

While efficiency wage models are often used as a rationale for involuntary unemployment (see Calvo 1985 or Shapiro and Stiglitz 1984), the model presented here is consistent with full employment or unemployment.\footnote{Katz (1986, p. 237) says, “If efficiency wage considerations are equally important in all sectors of the economy involuntary unemployment can arise.” The model presented here is consistent with an unemployment equilibrium where there are many sectors with different monitoring and turnover rates.} When workers can receive on-the-job offers and engage in on-the-job search the pattern of wage differentials remains unchanged qualitatively, but a full-employment equilibrium becomes possible. Equilibrium wages in all sectors will be a function of wage contracts in all other sectors. Otherwise the full-employment equilibrium illustrated is similar to that in Bulow and Summers (1986).

**II. The Effort Supply Curve**

In this section I derive the worker’s effort supply curve. Later it will be combined with effort demand from the firm’s problem to solve for equilibrium. Identical workers have the following utility function:

$$U = w - g(x), \quad x = \text{effort}, \quad w = \text{Wage},$$  \hspace{1cm} (1)

where $g(x)$ is the disutility of effort that is assumed to be convex. A worker who gets a job in sector $i$ receives a wage stream $w_i$ over time in return for a specified effort level $x_i$. A worker caught shirking is fired immediately. The values of a job in sector $i$ for shirkers and nonshirkers are, respectively,

$$rV_i^N(t) = w_i - r(x_i) + b_i[V_u(t) - V_i^N(t)] + \dot{V}_i^N(t)$$  \hspace{1cm} (2)

and

$$rV_i^N(t) = w_i - g(x_i) + b_i[V_u(t) - V_i^N(t)] + \dot{V}_i^N(t),$$  \hspace{1cm} (3)

where $b_i$ is the rate of job separations due to exogenous factors (layoffs,
etc.) and \( q_i \) is the detection rate (or monitoring intensity). Since \( q_i \) is the same for any effort level below the specified level \( x_i \), a worker who shirks will exert no effort. The discount rate is \( r \).

\( V_u(t) \) is the flow value of unemployment that is defined below. The right hand side of equation (2) is the dividend (wage) plus the probability of job loss times the change in asset value of losing a job, plus the change in value of being an employed shirker over time. These last two terms are the expected rate of capital gain (loss) associated with this state. The equation for a non-shirker is of the same form except the disutility of effort is subtracted out and the probability of losing the job if supervised is zero. The lowest wage that satisfies the no-shirking condition is when \( V_i^N(t) = V_i^N(t) \). We can use this condition to solve for the rent associated with a job in any particular sector:

\[
V_i(t) - V_u(t) = \frac{g(x_i)}{q_i}.
\]

A job that specifies higher effort, but where monitoring is more difficult gives higher rent. A sector where monitoring is perfect, so that \( q = \infty \), will have no rents. The exogenous probability of job loss will not affect the value of a job. High-turnover jobs may pay more, but this just compensates for the higher probability of being fired. We should not expect longer queues for high-turnover jobs.

The value of unemployment (welfare plus other job offers) is

\[
rV_u(t) = B + \sum_{j=1}^{n} a_j [V_j(t) - V_u(t)] + \dot{V}_u(t),
\]

where \( B \) equals welfare payments and \( a_j \) is the arrival rate of job offers from sector \( j \) to an unemployed worker. Initially we assume a worker must enter the unemployed pool to switch jobs. Using equation (4) we can rewrite the above equation as

\[
rV_u(t) = B + \sum_{j=1}^{n} a_j \frac{g(x_j)}{q_j} + \dot{V}_u(t).
\]

Once the no-shirking condition is satisfied workers will not shirk in equilibrium, so we treat the expression for \( rV_i^N(t) \) as the equilibrium value of a job. Next we can use equation (4) to solve for \( V_i(t) \) in terms of \( V_u(t) \). Using equation (6) above for \( V_u \) and noting that since (4) holds at each point in time \( \dot{V}_i(t) = \dot{V}_u(t) \), we can solve for the relationship between wages, effort, and the exogenous parameters such that the worker will not shirk. This yields the worker’s effort supply relationship:
\[ w_i(t) = g(x_i)A_i + C, \]  

(7)

where \( A_i = [1 + (b_i + r)/q_i] \) and \( C = B + \sum_{j=1}^{n} a_j \left[ g(x_j)/q_j \right] \). \( A_i \) is increasing with turnover in sector \( i \) and decreasing with monitoring intensity. \( C \) gives the value of welfare benefits plus the value of job offers the worker receives in unemployment. The value of these offers is constant across sectors. Workers are assumed to be identical. So, for example, a janitor who loses a job in a large manufacturing plant where monitoring is difficult has no reason to do better searching for jobs than a janitor who worked in a small restaurant. Figure 1 graphs equation (7), illustrating a worker's effort supply curve as he moves from sector 1 to sector 2, where monitoring intensity is higher. At any given effort level the worker prefers the \( q_1 \) effort supply curve where monitoring intensity is lower.

III. Profit Maximization

In many cases firms can react to the monitoring difficulty not just by paying higher wages, but also by specifying lower effort levels for the job. This is an alternative (and up to a point cheaper) way of making workers value their job so that they will not risk losing it by shirking. Competitive firms maximize profits:

\[ \text{Max}_w, N P_i F_i(x(w)N) - wN. \]  

(8)
$P_i$ is the output price in sector $i$, $F(\cdot)$ is the production function, and $N$ is the number of workers. As shown in Solow (1979), the first-order conditions imply

$$\omega^* = \frac{x(w^*)}{x'(w^*)}. \quad (9)$$

The elasticity of effort with respect to the wage will be one. Firms increase the wage to a point where the marginal cost of an increase in efficiency units is the same, whether it comes from hiring more workers, or increasing effort.\(^8\) I use equation (7) (the effort supply curve) as the $x(w)$ function that the firm faces in choosing the wage in equation (8) above. Setting the elasticity of effort from a change in wages equal to one we get the equilibrium illustrated in figure 2 where a ray from the origin (effort/wage) equals the slope of the effort supply curve.

\(^8\) Effort and workers will not always be perfect substitutes. Ramana and Rowthorn (1991) look at a more general case where the technology is $F(e(w), N)$. It may be that lower effort by one worker (an airline pilot is the example used by Ramana and Rowthorn) may do a lot of damage to the firm. Equilibrium in this case is where the elasticity of effort with respect to wages equals the ratio of the output elasticity with respect to labor and the output elasticity with respect to effort.
IV. Wages and Effort across Sectors

Firms in each sector choose both wages and effort optimally in equilibrium. In this section I compare the wage and effort level chosen by a firm in a given sector with the wage and effort level chosen by a firm in a different sector with more intensive monitoring or lower turnover. The combination they choose depends on how expensive it is to get workers to exert additional effort (the shape of the effort supply curve) relative to the cost of hiring additional workers. Firms in sectors with higher monitoring intensity or lower separation rates face a higher effort supply curve as shown in figure 1. I will show below that this implies firms in these sectors will choose higher effort in equilibrium.

Whether firms with higher monitoring intensity (or lower turnover) choose to pay higher or lower wages depends on the shape of the effort supply curve. If it is increasingly difficult to get workers to exert additional effort (the elasticity of the disutility of effort with respect to effort is increasing), firms facing a higher effort supply curve will choose a little more effort and a lower wage than firms facing a lower curve. If the elasticity is decreasing, it is easier to get workers to exert additional effort, and firms facing a higher effort supply curve choose more effort to the extent that they pay a higher wage than firms facing the lower curve. The implication of this is that the shape of the effort supply curve determines whether firms with more intensive supervision pay higher or lower wages.

We know from the Solow condition (eq. [9]) that the elasticity of effort with respect to wages equals one, which implies inversely that the elasticity of wages with respect to effort equals one. Imposing this condition on the effort supply curve (eq. [7]) we get the following relationship:

\[ w_i = g(x_i)A_i + C = \frac{dg(x_i)}{dx_i} A_i x_i. \]  

(10)

This in turn implies

\[ \frac{C}{A_i} = x_i g'(x_i) - g(x_i). \]  

(11)

Recall that the parameters of \( A \) and \( C \) are exogenous to the firm while they choose wages and effort. Higher monitoring \( q_i \) or lower turnover \( b \), is associated with a higher value of the left-hand side of equation (11). This implies higher effort since the right-hand side is increasing in effort if the disutility of effort \( g(x) \) is convex (which is the second order condition for the firm’s problem). In other words, if effort is more expensive (firms face a lower effort supply curve in fig. 2), we expect firms to choose lower effort in equilibrium.
From equations (10) and (11) we can show that the elasticity of the disutility of effort with respect to effort \( E \), satisfies the following equation:

\[
\omega_i = \frac{C}{1 - \frac{1}{E_i(x_i)}}. \tag{12}
\]

Since we know that higher monitoring intensity means higher effort, equation (12) tells us that whether wages are higher or lower in a sector with higher monitoring intensity depends only on whether the elasticity of the disutility of effort with respect to effort is increasing or decreasing in effort. As long as it is not increasingly easy to elicit additional effort, wages will be lower in the closely monitored sectors.\(^9\)

We can look at three examples where wages are lower, the same or higher in sectors with higher monitoring. The \( g(x) \) functions for these three cases are, respectively,

\[
g(x) = x + x^\alpha, \tag{13}
\]

\[
g(x) = x^\alpha, \tag{14}
\]

and,

\[
g(x) = x \ln(x). \tag{15}
\]

Assume \( \alpha \) is a parameter \( > 1 \). Notice that the elasticity of the disutility of effort with respect to effort is increasing, constant, and decreasing in effort, respectively. Using the above functional forms and imposing the Solow condition as in equation (10), we get the equilibrium wage implied for each case:

\[
\omega_i = \frac{\alpha}{\alpha - 1} C + \left( \frac{C}{\alpha - 1} \right)^{\frac{1}{\alpha A}} \frac{\alpha - 1}{\alpha}, \tag{16}
\]

\(^9\) An imperfect monitoring technology may falsely identify nonshirkers as shirkers and vice versa. As a referee has pointed out, the implication is that part of the exogenous separation rate represents these false positives. It would be reasonable therefore to think of the separation rate falling as monitoring intensity and effort increases since the rate of false positives might be expected to fall. Weiss (1991) discusses the effects of an increase in the level of monitoring as against an increase in the precision of monitoring.
\[ w_i = \frac{\alpha}{\alpha - 1} C, \]  

(17)

and

\[ w_i = C \left[ 1 + \ln \left( \frac{C}{A_i} \right) \right]. \]  

(18)

Note that \( A_i \) is decreasing in monitoring intensity and increasing in the separation rate. Wages are lower, the same, or increasing, respectively, as we move to a sector with higher monitoring or a lower separation rate. The implication is that in a sector where firms can easily substitute between effort and workers, the standard prediction of the monitoring/tturnover version of the efficiency wage model is overturned, unless getting workers to exert additional effort is increasingly difficult. We could also choose examples where the elasticity of the disutility of effort with respect to effort was not monotonic implying that the relationship between wages and supervision would not be monotonic for a given worker moving across sectors.

V. Simulated Wage Differentials

Using reasonable parameter values I can simulate the model to see how big the wage differentials generated by the model are. These simulations can then be compared with observed industry wage differentials to give a rough guide to the size of wage differentials we can realistically attribute to this model and what parameter values would be needed to generate wage differentials as big as those observed in the empirical literature.

I calculate the equilibrium wage differentials for a particular utility function above (eq. [13]). I put one extra restriction on the model by assuming the benefit level remains as a fixed fraction of the competitive wage. I would otherwise have to choose a nominal value for the benefit level whereas the model requires a real value. The nominal value chosen would affect the size of wage differentials.

Using the utility function in equation (13) in equation (11) we see that for any sector \( i \)

\[ C = A_i(\alpha - 1)x_i^\alpha = B + \sum_{j=1}^{n} \frac{a_j}{q_j} (x_j + x_j^\alpha). \]  

(19)

From this we see that for any two sectors \( i \) and \( j \)
\[
\frac{x_i}{x_j} = \left( \frac{A_j}{A_i} \right)^{\frac{1}{\alpha}}.
\]  \(\text{(20)}\)

This implies that once we solve for effort in one sector we can easily solve for the other sectors. Focusing on sector 1 and using equations (19) and (20), we see that the following equation holds:

\[
\frac{B}{\alpha - 1} + \sum_{j=1}^{n} \frac{a_j}{q_j(\alpha - 1)} \left( \frac{A_1}{A_j} \right)^{\frac{1}{\alpha}} x_i^{\alpha} + \sum_{j=1}^{n} \frac{a_j}{q_j(\alpha - 1)} \left( \frac{A_1}{A_j} \right) x_1^{\alpha} - A_1 x_1^{\alpha} = 0.
\]  \(\text{(21)}\)

Next, say for the purpose of solving the model, we assume the benefit level to be a fixed fraction \(\lambda\) of the wage in the lowest wage industry, making \(B\) endogenous:

\[
\lambda w_i = \lambda (x_i + x_1^{\alpha}) A_i + \lambda C = \lambda A_i (\alpha x_1^{\alpha} + x_i).
\]  \(\text{(22)}\)

If we substitute this into the previous equation where \(i\) is equal to one, we get

\[
\left[ \sum_{j=1}^{n} \frac{a_j}{q_j(\alpha - 1)} \left( \frac{A_1}{A_j} \right)^{\frac{1}{\alpha}} + \frac{\lambda A_1}{\alpha - 1} \right] x_1^{\alpha} + \left[ \sum_{j=1}^{n} \frac{a_j}{q_j(\alpha - 1)} \left( \frac{A_1}{A_j} \right) + \left( \frac{\alpha \lambda}{\alpha - 1} - 1 \right) A_1 \right] x_1^{\alpha} = 0.
\]  \(\text{(23)}\)

The first term in squared brackets is a constant \((k_0)\) that is unambiguously positive, while the second term in squared brackets \((k_1)\) will be negative if \(\lambda\) and the acquisition rates are small enough. In this case a unique equilibrium exists and \((k_0/k_1)^{(1/\alpha-1)} = x_1\). Next, to calculate the wage in each sector we need the constant \(C\):

\[
C = A_1 (\alpha - 1) x_1^{\alpha}.
\]  \(\text{(24)}\)

If we take the special case where sector 1 has perfect monitoring so that \(q_{1\to1}\), then \(A_1 = 1\). If \(\alpha\) is assumed to equal two, equation (23) can be written as
Table 1
Wage Premia over Competitive Sector at Different Monitoring and Turnover Levels (Discount Rate = 3%)

<table>
<thead>
<tr>
<th>Opportunity Cost of Employment Equals 40% of Wage (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>$p_1 = .3$</td>
</tr>
<tr>
<td>$p_1 = .5$</td>
</tr>
<tr>
<td>$p_1 = .9$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Competitive Sector Effort (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_2 = .1$</td>
</tr>
<tr>
<td>$p_1 = .3$</td>
</tr>
<tr>
<td>$p_1 = .5$</td>
</tr>
<tr>
<td>$p_1 = .9$</td>
</tr>
</tbody>
</table>

Note.—Using equation (16), the wages in the table are constructed as

$$w_i + \frac{\alpha}{\alpha - 1} C + \left( \frac{C}{\alpha - 1} \right)^\frac{1}{\alpha - 1} A_i^{\alpha - 1},$$

where $A_i = 1 + (b_i + r)/q_i$. Recall that $b$ is the separation rate and $q$ the supervision rate; both are assumed to be Poisson processes. The formula to convert the probabilities in the table into arrival rates is $p = 1 - e^{-\lambda}$, where $p$ is a probability and $q$ an arrival rate, $p_1$ is the probability a worker will be caught shirking in one period, $p_2$ is the probability of a separation for other reasons in one period.

$$\left[ \sum_{j=1}^{n} \frac{\alpha_j}{Q_j} A_j^{-\frac{1}{2}} + \lambda \right] x_1 + \left[ \sum_{j=1}^{n} \frac{\alpha_j}{Q_j} A_j^{-1} + 2\lambda - 1 \right] x_1^2 = 0. \quad (25)$$

We use this equation to solve for $x_1$ and $C$. We can then solve for effort and wages in each sector. A 10-sector model is simulated in table 1 where the benefit level is 40% of the competitive wage. We also need to assume reasonable parameter values for the rate of supervision and the separation rate.

I assume the arrival rates of supervisors and exogenous job separations follow a Poisson process and so that the interarrival time of each event is exponentially distributed. The probability of an event in one period of time is $(1 - e^{-q})$, where $q$ is the arrival rate. Using a reasonable range of probabilities of being supervised or experiencing a job separation over a month-long period, we can calculate the arrival rates and implied wage differentials. $p_1$ is the probability a worker who shirks continuously for a month will be caught and fired, and $p_2$ is the probability of a job separation over a month-long period.
Hall (1995) analyses job separations in the U.S. labor market. Hall notes that “putting aside the briefest jobs, around 8 or 10 percent of workers separate from their employers in each quarter” (p. 235). Anderson and Meyer (1994), however, refer to Hall and Lilien’s (1979) study that showed that the median worker is in a job that will last eight years and 28% of workers are in jobs that will last at least 20 years. The implication is that most workers are in jobs with much lower separation rates than implied by the figures given above.

While it is difficult to observe the probability of being caught shirking, Neal (1993) analyses data on the intensity of supervision across sectors from a supplement to the 1977 Panel Study of Income Dynamics. Only 10% of workers in the sample had no supervision, and 70% were supervised at least once “every few weeks” across two-digit industries. This figure ranged from 50% in repair services to 96% in textiles, with an SD of 12%. Based on these percentages, it seems reasonable to assert that most workers who shirked continuously for any extended period would have a high probability of being caught in most sectors.

Table 1 assumes benefits of unemployment represent 40% of the competitive sector wage. Blöndal and Pearson (1995) give income replacement rates for Organization for Economic Cooperation and Development countries in 1991. These are lower bounds for the benefit level since they are compared with the average wage, rather than the lowest wage as in the model. While the replacement rates range from 15% in Italy to 77% in Switzerland, both the mean and median figure are close to 50%. Given all the above factors and the fact that wage differentials are decreasing in the size of the benefit level, the value chosen in table 1 possibly overstates wage differentials. I assume a discount rate of 3%, but changes in the results from using different discount rates are small.

The model predicts very small wage differentials across sectors. For example, if benefits were 40% of the wage, the separation rate was 10% per month, and there would only be a 50% chance of getting caught after shirking continuously for an entire month, a wage premium of 2% over the competitive wage would satisfy the no-shirking condition. Table 1 also calculates how much lower equilibrium effort would be in each sector compared with the competitive sector so that we can calculate the relative cost of a unit of effort across sectors. The implication is that while the worker mentioned above receives only a 2% wage premium over the competitive sector, his effort is 13% more expensive than effort in the competitive sector. It is also worth bearing in mind that while we have treated the rate of separations and rate of supervision as exogenous, in reality they can be influenced by the firm, making the combinations of parameter values that generate higher wage and effort differentials more unlikely.

In some cases firms find it profitable to pay a wage less than that paid
in the competitive sector and accept less effort than the competitive sector. This would happen if we used the utility function in equation (15), for example. These jobs would still have positive rents and are more desirable than competitive jobs because they require lower effort.

It should be noted that the above wage differentials are deviations from the competitive sector. This should be borne in mind when comparing these simulated wage differentials with wage differentials reported in the empirical section later in the article or in such studies as Krueger and Summers (1988). These report the percentage difference of the wage in any sector from the mean of the wage across all sectors.

VI. Labor-Market Policies

The model has implications for the analysis of standard labor-market issues. Higher welfare payments will increase wages and reduce effort in all sectors, and as illustrated in the simulations in the previous section, will lower wage differentials. Higher welfare payments amount to increasing the C term in equation (7) or shifting the intercept of the effort supply curve to the right in figure 2. Welfare payments will also reduce employment in the usual way. If income support payments depend on unemployment insurance rates, firms with monitoring difficulties have the incentive to hire uninsured workers since the threat of job loss is worse for this group. The implication is that if unemployment insurance depends on length of employment, workers should be moved to more easily monitored positions as length of service increases.10

The results from imposing a minimum wage on the model are consistent with some empirical findings. Grossman (1983) provides evidence of minimum wage payments increasing wages across skill groups, while Katz and Krueger (1992), analyzing the Texas fast-food industry, find that in response to the 1991 minimum-wage increase, firms that were initially above the minimum wage on average increased their starting wage. Imposing a minimum wage in the model forces some low-wage firms to set wages above the level satisfying equation (9), so the elasticity of effort with respect to wages will be less than one. These firms will still extract the maximum effort possible given the wage they pay, so while equation (9) will no longer hold, one can still use equation (7) to get the relationship between effort and wages. This means firms will be on the effort supply curve in figure 2 but will be to the right of the equilibrium point.

10 This is the opposite prediction from that in an efficiency wage model where the firm can use a deferred payment schedule to reduce the incentive to shirk (see Akerlof and Katz 1989). In their model firms have the incentive to put less-experienced workers in less-experienced positions until the value of forgone earnings is large enough to outweigh the incentive to shirk.
The rent associated with jobs in the affected sectors (eq. [4]) will be higher given that effort is higher. The value of $C$ in (eq. [7]) will increase, driving up wages and lowering effort in all other sectors. Any resultant lowering of employment and increase in unemployment will mitigate the upward pressure on wages in other sectors.\footnote{Rebitzer and Taylor (1995) show in the Shapiro and Stiglitz (1986) framework that if monitoring costs are increasing with employment, a minimum wage can increase employment. Manning (1995) shows that employment is increasing if the elasticity of the marginal revenue of the wage with respect to the wage is greater than unity.}

Bulow and Summers (1986) and Katz and Summers (1989) argue that output in the primary sector is too low and that subsidies to the high-wage industry can increases welfare. It is shown in Walsh (1995) that, if one accounts for the impact on other sectors, subsidising the high-wage sector can increase unemployment and may reduce welfare.

VII. Labor-Market Equilibrium, On-the-Job Offers, and Full Employment

In this section I generalize the model in a number of ways. I show that if the unemployment rate changes or if industries grow at different rates the pattern of wage differentials remains unchanged. In addition, workers can receive on-the-job offers or engage in on-the-job searches, and the qualitative pattern of wage differentials does not change. The case where there are on-the-job offers is then used to illustrate how the model is consistent with full employment. Following Kimball (1994), the growth of employment in any sector is

$$L_i = a_i(t)[N(t) - L(t)] - b_i L_i(t).$$  \hspace{1cm} (26)

The first term is the arrival rate of job offers from sector $i$ to each worker in the unemployed pool $a_i$, times the size of the unemployed pool. As before, $b_i$ is the rate of job loss to workers in sector $i$. We can divide by employment in this sector to get the growth rate of employment:

$$\lambda_i(t) = \frac{L_i(t)}{L_i(t)} = \frac{a_i(t) u(t)}{\theta_i} - b_i.$$  \hspace{1cm} (27)

$\theta_i$ is industry $i$’s employment as a share of the labor force. I use the above equation to solve for the equilibrium relationship between job arrival rates, unemployment, employment growth, and the exogenous probability of job loss:
\[ a_i(t) = \frac{\theta_i}{\mu(t)} (\lambda_i(t) + b_i). \]  

(28)

For any given disutility of effort function \( g(x) \), this can be used in the worker's effort supply function, equation (7), to give the effort wage relationship that satisfies the no shirking condition and takes account of labor-market conditions:

\[
\omega_i(t) = g(x_i)A_i + B + \frac{1}{\mu(t)} \sum_{i=1}^{n} \theta_i[(\lambda_i(t) + b_i)] \frac{g(x_i)}{q_i} \\
= g(x_i)A_i + B + \frac{1}{\mu(t)} K(t).
\]

(29)

Equation (29) indicates that wages in any sector will change by the same amount with a change in unemployment. In other words, if effort is held fixed, the size of wage differentials will not vary cyclically or as industries grow at different rates, although wages in all sectors will be procyclical.

On-the-Job Offers

In previous sections, \( b_i \) represented the rate of job separations, whether voluntary or involuntary, but for reasons unrelated to performance. In this section workers receive offers on the job and only quit on receiving a better offer. We take account of the impact of these offers on the no-shirking condition. This generalization will not change either the rent associated with a job in any sector or the qualitative pattern of wage differentials.

If the wage in all sectors satisfies the no-shirking condition one can change equations (2) and (3) to allow for the probability of being offered a better job while being employed in any given sector \( i \). Being offered jobs in worse sectors gives zero additional utility. The term below reflecting the value of on the job offers is added to equations (2) and (3) \( (d_j \) is the arrival rate of job offers from sector \( j \) to workers in sector \( i \)):

\[
\sum_{j=1}^{n} d_j(\max[0, V_j - V_i]).
\]

(30)

Equation (4), which gives the rent of a job in any sector \( i \), is unchanged when the no-shirking condition is imposed. This allows one to rank the jobs in each sector where a higher number is a better job and to define the value of on-the-job offers in that sector:
\[ Z_i = \sum_{j=i+1}^{n} d_j \left[ \frac{g(x_j)}{q_j} - \frac{g(x_i)}{q_i} \right]. \]  

(31)

Imposing the no-shirking condition gives the wage effort relationship:

\[ w_i(t) = g(x_i) \left[ 1 + \frac{b_i + r}{q_i} \right] + B + Y - Z_i. \]  

(32)

The availability of on-the-job offers means that workers do not completely lose the possibility of getting good jobs when they accept a low-paying job. This means that for any given effort level the no-shirking condition can be satisfied at a lower wage in the bad sectors. To a worker who already has a good job on-the-job offers are not valuable and the no-shirking condition remains unchanged. Thus \( Z_i \) increases in worse sectors, and allowing on-the-job offers actually lowers wages in the worst jobs. Essentially the point is that if a worker accepting, say, a low-paying services job does not forego the possibility of getting a “good” job to any great extent, then the worker will accept the job at a lower wage than if this were not true. On-the-job offers therefore will not change the ranking of good and bad jobs or the size of rents and will, if anything, expand wage differentials between sectors.

While the above analysis assumed that offers flowed in costlessly we could impose some search costs. Suppose the probability of getting a job offer from sector \( j \) is an increasing function of \( K_j \) (the intensity of the worker’s search effort in that sector), that is, \( d_j(K_j) \). Assume the marginal cost of searching more intensively for a worker in sector \( i \) is a constant \( s_i \) (as illustrated in fig. 3). It seems reasonable to expect this marginal cost to be higher in the better sectors since the opportunity cost of an hour at a good job is bigger than at a bad job. Next assume that the marginal benefit of searching is diminishing for workers in every sector, but the marginal benefit of search from a better sector is less than in a worse sector (since there is less to gain in the better sector). In figure 3 the marginal benefit of search in sector \( i \) is \( MB_i \) and sector 1 is the better sector. Workers employed in the worst sectors would engage in the most search activity, but the benefit of on-the-job offers net of search costs would be bigger in the better sector. Total search costs per worker in sector \( i \) (\( S_i \)) can be subtracted from \( Z_i \) in equation (32). The \( (Z_i - S_i) \) term is still greater in the worst sector, getting smaller as we progress into better sectors and reaching zero in the best sector. In figure 3 we can say that \( (Z_i - S_i) \) equals \( A \) in sector 1 and \( A + B + C + D \) in sector 2. Thus, on-the-job search costs offset the benefits of on-the-job offers,
but the pattern of wage differentials, rents, and the ranking of good and bad jobs is unaffected by allowing on-the-job search costs.\footnote{Including search costs for unemployed workers looking for jobs would just amount to subtracting a constant from the equilibrium value of $X$ in equation (32). This would not change any results qualitatively.}

Full Employment

In the Shapiro-Stiglitz framework unemployment was a "worker discipline device," and zero unemployment meant workers would certainly shirk since they could instantaneously find another job if fired. Bulow and Summers (1986) develop a two-sector version of the model with a fixed-wage market clearing sector. The case with on-the-job offers can be used to illustrate a full-employment equilibrium similar to that in Bulow and Summers where the labor market is in equilibrium and the no-shirking condition is satisfied in both sectors. The model presented here shows full employment in a more general framework than Bulow and Summers in that I do not fix wages in the secondary (competitive) sector, and in equilibrium this wage will be a function of the wage in the primary sector. The two sectors are a primary sector where there is imperfect monitoring and a secondary sector with perfect monitoring (so $A_2$
Workers can receive offers while on the job, and I assume that jobs are always readily available in the secondary sector and that this sector will clear the market. The wages in the primary and secondary sector are, respectively,

$$w_1 = g(x_1)A_1 + B + Y,$$  \hspace{1cm} (33)

and

$$w_2 = g(x_2) + B + Y - Z_2.$$  \hspace{1cm} (34)

$Z_2$ represents the value of on-the-job offers to workers employed in sector 2. These have no value for sector 1 workers who already have the best job. $Y$ is the value of job offers a worker would receive in the unemployment pool.

$$Z_2 = d_2 \left[ \frac{g(x_1)}{q_1} \right], \quad Y = a_1 \left[ \frac{g(x_1)}{q_1} \right].$$  \hspace{1cm} (35)

For simplicity I assume that the arrival rate of job offers in sector 1 is the same for employed and unemployed workers. This means $Y = Z_2$, and these cancel out of the wage in the competitive sector (eq. [34]). The employment growth rate in each sector is set equal to zero in the stationary case:

$$\dot{L}_1 = -b_1 L_1 + d_2 L_2 = 0;$$  \hspace{1cm} (36)

$$\dot{L}_2 = -b_2 L_2 - d_2 L_2 + f = 0.$$  \hspace{1cm} (37)

Equation (36) just says that the outflows from sector 1 must be matched by inflows from sector 2. Equation (37) is driven by the assumption that workers can instantaneously get a job in sector 2, so that a worker who loses his job in sector 1 instantaneously takes a job in sector 2 until a better offer arrives. Workers in sector 2 who lose their job do the same. The $f$ term is a residual representing all workers from sectors 1 and 2 who lost their jobs and did not get a primary sector offer, less those workers employed in sector 2 who got a primary sector offer. Equation (36) implies that

$$d_2 = b_1 \frac{L_1}{L_2}.$$  \hspace{1cm} (38)
If the primary sector is larger than the secondary sector, the arrival rate of jobs in the primary sector will be greater in a stationary equilibrium, making the threat of losing a primary sector job less severe. This means that the primary sector wage differential must be larger to enforce the no-shirking condition.

The point here is to show that the model is not being driven by unemployment. It may well be that we expect some frictional unemployment, for example, and this could serve as a discipline device, but we do not need involuntary unemployment to get an equilibrium in this model.

VIII. Empirical Industry Differentials and Monitoring

In this section I will examine whether a monitoring variable generates wage differentials that are consistent with the simulated wage differentials above and how they compare with observed industry wage differentials. Krueger and Summers (1988) showed that in a linear regression of the log wage on worker and job characteristics industry effects were large. The controversial question is, "What are industry dummies measuring?" There is a large literature addressing this question. Many studies, such as Krueger and Summers (1988) or Katz and Summers (1989), interpret them as reflecting efficiency wage payments. While others argue that these industry effects are a reflection of unobserved worker or job characteristics, or rent sharing, I will not focus on this argument. Rather, I will examine how well a measure of supervision intensity can predict these industry differentials.

Having shown by simulated wage differentials that the theory does not predict large wage differentials, I will show that an empirical measure of supervision movements across sectors does not have large wage effects. This is in line with other empirical studies of the impact of supervision on industry differentials.

Theoretical efficiency wage models have typically predicted that wages will be relatively high where monitoring intensity is low (see, e.g., Shapiro and Stiglitz 1984; Bulow and Summers 1986). However, the empirical evidence is mixed. Neal (1993) finds that controlling for supervision has no effect on industry differentials. Brunello (1995) found that doubling the ratio of supervisors to employees reduced wages by about 6% for nonmanual workers, while Groshen and Krueger (1990) found a similar reduction of 13% for American nurses. Krueger (1991) provides evidence that shift workers in the fast-food industry do better in chain-run stores than in franchised stores. The argument is that monitoring is more effective when the owner runs the store, while the centrally run stores have paid supervisors who are less effective. Kruse (1992) also found that the intensity of supervision had a negative effect on wages.

Neal's results suggest that monitoring differences across sectors are not driving industry wage differentials. The results of the other studies (with
the exception of Groshen and Krueger’s [1990] study of pay in hospitals) predict wage differentials of at most five or six percentage points due to supervision differences. Groshen and Krueger (1990) generate large wage effects for nurses by using Standard Metropolitan Statistical Area dummies as instruments for the supervision ratio. The rationale for these geographical instruments is that supervision rates in hospitals are typically set by local regulations. It is worth noting that for three of the four occupations in the sample the effects of supervision on wages are negligible. In addition, it should be noted that while using regulated supervision rates may overcome the endogeneity problem associated with supervision rates that are chosen optimally, one should be careful about how one interprets the wage effects. The study is estimating the trade-off between wages and supervision implied in nursing by observing different hospitals that are forced to use different supervision rates. To see how well differences in supervision can explain observed industry wage differentials we should estimate wage differentials between workers working across sectors with different monitoring technologies where monitoring is chosen optimally.

Data

The data I use comes from the 1988 Current Population Survey (CPS). This gives a cross section of data on wages and other personal characteristics of the U.S. civilian noninstitutional population. The hourly wage variable is computed from usual hours worked and weekly wages (table A1 gives a summary of the variables). I use the ratio of supervisors to workers in each sector as a proxy for the intensity of supervision. This is calculated from a subsample of each monthly rotation group of the 1988 CPS that has three-digit industry and occupation codes for almost 158,000 private sector nonagricultural workers. The total hours worked of workers specifically described by their three-digit occupation code as supervisors are counted within each sector and divided by total hours worked by nonsupervisors in that sector to get each sector’s supervision rate. The occupations chosen were the same as those used by Dickens et al. (1989). In the regression analysis described below, I restricted the sample to 1 month (March) of the annual CPS data that gave 12,043 observations.

Estimating Wage Differentials

I estimated the coefficients on the following log wage regression:

$$\log w = X\hat{\beta} + \hat{\beta}I + \hat{\varepsilon}. \quad (39)$$

Table 2
Standard Deviation and Correlation Coefficients for Industry Differentials Calculated from Industry Dummies and Supervision Rates

<table>
<thead>
<tr>
<th></th>
<th>SD (%)</th>
<th>Correlation with Industry Differentials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry dummies†</td>
<td>14.86</td>
<td></td>
</tr>
<tr>
<td>Linear supervision rate‡</td>
<td>.07</td>
<td>.019</td>
</tr>
<tr>
<td>Linear and squared supervision rate</td>
<td>.63</td>
<td>.194</td>
</tr>
<tr>
<td>Linear, squared, and cubed supervision rate</td>
<td>1.24</td>
<td>.294</td>
</tr>
<tr>
<td>Log of supervision rate</td>
<td>.97</td>
<td>.152</td>
</tr>
</tbody>
</table>

NOTE.—If \( m \) is the estimated coefficient on supervision, \( ms \) the coefficient on squared supervision, and \( sup \) is the level of supervision in sector \( i \), then \( b_i = m(sup_i) \) is the estimated percentage change in wages due to supervision in sector \( i \) from the first regression, \( b_i = m(sup_i) + ms(sup_i) \) is that from the second regression, and so on. Using \( b_i \) for \( \hat{\beta} \) in equation (40) one can construct wage differentials due to supervision across three-digit industries. The final four rows reflect the differentials calculated from these four different regressions. The estimated coefficients from three regressions are given in table A1.

† As calculated in equation (41).
‡ Three-digit industry dummies are included in a log wage regression and used to calculate industry differentials using equation (40).
§ I ran four separate wage regressions for nonsupervisory workers with supervision as an independent variable. First, I included the supervision rate by industry. Second, I added the squared supervision rate. Third, the cubed supervision rate. And, finally, only the log of the supervision rate was included.

\( I \) is a dummy variable for industry. \( X \) is a vector of worker and job characteristics. The \( \hat{\sigma} \) and \( \hat{\beta} \) are the estimated coefficients, and \( \hat{\epsilon} \) the error term. Following Krueger and Summers (1988), industry wage differentials are constructed as the deviation of the individual industry dummy from the employment weighted mean of industry dummies in equation (39):

\[
d_i = \hat{\beta}_i - \sum_{j=1}^{k} \frac{n_j}{N} \hat{\beta}_j,
\]

(40)

\( \hat{\beta} \), is the estimated coefficient from the industry dummy for that sector, \( N \) is total employment and \( n_i \), employment in sector \( i \). There are \( k \) industries. The employment-weighted SD of wage differentials is

\[
SD = \sqrt{\frac{\sum_{i=1}^{n} \frac{n_i}{N} d_i^2}{N}}.
\]

(41)

This gives the typical deviation in wages associated with changing industry (in percentage terms) and equals 15%. That is, the industry differentials predict that a worker changing industries would typically expect a wage change of 15% other things being equal.

I approximate the importance of monitoring by including the monitoring variable in the regression in place of the industry dummies in equation (39). The correlation coefficient between three-digit industry differentials and the measure of supervision is \(-0.02\), indicating that there is no linear relationship between the two variables. The SD of these wage differentials are given for different specifications of the supervision variable in the regression in table 2, as is their correlation with wage differentials calcu-
lated from industry dummies. It is clear from the table that the coefficients on the supervision term, as well as on a square, cube, or log of the supervision term were too small to have an important effect on wages. The regression results and summary of the variables are given in table A1, while figure 4 graphs industry differentials constructed from the industry dummies and from the supervision regression that generated the largest differentials. Figure 4 and table 2 show clearly that differences in supervision across sectors generates differentials that are very small compared with those generated from industry dummies.

Hiring supervisors is not the only way to monitor workers. The monitoring technology, cost, and quality of monitoring may differ across sectors, so the ratio of supervisors to workers is at best a rough approximation of monitoring intensity. Having a small number of supervisors may indicate more effective monitoring (as in the argument by Krueger outlined earlier). Table 3 lists the industries with no supervisors and the number of employees in each from the pooled CPS data set I used to calculate supervision. Employment (the sample size) is reasonably large in most sectors. These are mostly industries where we would expect small-scale firms that would have little difficulty supervising (although the zero monitoring case represents less than 2% of the total sample).

Because of these difficulties, and the endogeneity issue raised by Groshen and Krueger (1990), one should be cautious in interpreting the results. Even so, the failure of supervision to have an important impact on
Table 3
Sectors with No Monitors

<table>
<thead>
<tr>
<th>Industry</th>
<th>Employment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miscellaneous petroleum and coal products</td>
<td>29</td>
</tr>
<tr>
<td>Leather tanning and finishing</td>
<td>20</td>
</tr>
<tr>
<td>Taxicab service</td>
<td>103</td>
</tr>
<tr>
<td>Barber shops</td>
<td>42</td>
</tr>
<tr>
<td>Shoe repair shops</td>
<td>10</td>
</tr>
<tr>
<td>Bowling alleys, billiard and pool halls</td>
<td>108</td>
</tr>
<tr>
<td>Offices of physicians</td>
<td>1,261</td>
</tr>
<tr>
<td>Offices of dentists</td>
<td>714</td>
</tr>
<tr>
<td>Offices of chiropractors</td>
<td>74</td>
</tr>
<tr>
<td>Offices of optometrists</td>
<td>86</td>
</tr>
<tr>
<td>Offices of health practitioners</td>
<td>53</td>
</tr>
<tr>
<td>Educational services</td>
<td>142</td>
</tr>
</tbody>
</table>

wages, alongside the results from other studies, supports the predictions of the model that wage differentials will be small.

IX. Conclusion

This article presents a model of efficiency wages that allows for many sectors, in each of which firms endogenously choose the level of effort supplied by the worker. A multisector model allows us to establish a link between the theoretical predictions of efficiency wage models and observed wage differentials. The model does not predict the large wage differentials across sectors attributed to efficiency wage models in much of the empirical literature on industry wage differentials. The empirical evidence on the effect of monitoring on wages supports the model in this respect. The model is also consistent with full employment, while efficiency wage models are often used to rationalize involuntary unemployment. Looking for large wage premiums as support for efficiency wage models or using these models to rationalize unemployment may be misguided. Unemployment of the kind described in Shapiro and Stiglitz (1984) would only arise if there is not a market clearing sector where efficiency wage factors are unimportant. There could be wait unemployment of the kind described in Bulow and Summers (1986) if workers believe they have a better chance of getting a good job by searching from unemployment.

The model shows that since it is more expensive to elicit effort in badly monitored or low-turnover sectors, jobs in these sectors will have lower effort levels. Firms in these sectors may choose lower effort to the extent that badly monitored or low-turnover sectors may have lower wages, overturning the standard prediction of efficiency wage models that these would be the high-wage sectors. Workers can take advantage of on-the-job offers and engage in search activity, sectors can grow at different
rates, and a change in the rate of unemployment or unemployment benefits changes wages in all sectors equally.

**Appendix**

**Table A1**  
Regression Results and Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Linear Supervision</th>
<th>Squared Supervision</th>
<th>Cubed Supervision</th>
<th>Log of Supervision</th>
<th>Industry Dummies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekly hours</td>
<td>37.9</td>
<td>11</td>
<td>.015</td>
<td>.015</td>
<td>.015</td>
<td>.015</td>
<td>.008</td>
</tr>
<tr>
<td>Hourly earnings</td>
<td>9.27</td>
<td>6.47</td>
<td>(.007)</td>
<td>(.007)</td>
<td>(.007)</td>
<td>(.007)</td>
<td>(.007)</td>
</tr>
<tr>
<td>Education</td>
<td>12.7</td>
<td>2.6</td>
<td>.003</td>
<td>-.003</td>
<td>-.003</td>
<td>-.003</td>
<td>-.003</td>
</tr>
<tr>
<td>Age</td>
<td>36.4</td>
<td>13</td>
<td>.003</td>
<td>-.003</td>
<td>-.003</td>
<td>-.003</td>
<td>.003</td>
</tr>
<tr>
<td>Married</td>
<td>.58</td>
<td>.49</td>
<td>.076</td>
<td>.076</td>
<td>.074</td>
<td>.074</td>
<td>.065</td>
</tr>
<tr>
<td>Male</td>
<td>.53</td>
<td>.50</td>
<td>.186</td>
<td>.186</td>
<td>.185</td>
<td>.185</td>
<td>.167</td>
</tr>
<tr>
<td>White</td>
<td>.88</td>
<td>.32</td>
<td>.044</td>
<td>.044</td>
<td>.044</td>
<td>.044</td>
<td>.049</td>
</tr>
<tr>
<td>Black</td>
<td>.09</td>
<td>.28</td>
<td>-.012</td>
<td>-.011</td>
<td>-.11</td>
<td>-.009</td>
<td>-.008</td>
</tr>
<tr>
<td>Part-time</td>
<td>.19</td>
<td>.39</td>
<td>-.175</td>
<td>-.174</td>
<td>-.172</td>
<td>-.175</td>
<td>-.146</td>
</tr>
<tr>
<td>Center city</td>
<td>.23</td>
<td>.42</td>
<td>.029</td>
<td>.029</td>
<td>.030</td>
<td>.030</td>
<td>.036</td>
</tr>
<tr>
<td>Veteran</td>
<td>.15</td>
<td>.35</td>
<td>.019</td>
<td>.019</td>
<td>.018</td>
<td>.021</td>
<td>.011</td>
</tr>
<tr>
<td>Union member</td>
<td>.13</td>
<td>.34</td>
<td>.267</td>
<td>.266</td>
<td>2.66</td>
<td>2.65</td>
<td>.200</td>
</tr>
<tr>
<td>Union contract</td>
<td>.02</td>
<td>.14</td>
<td>.092</td>
<td>.091</td>
<td>.091</td>
<td>.096</td>
<td>.091</td>
</tr>
<tr>
<td>Supervision</td>
<td>-.0009</td>
<td>.002</td>
<td>.006</td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
<tr>
<td>Squared supervision</td>
<td>-.0005</td>
<td>-.003</td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
<tr>
<td>Cubed supervision</td>
<td>3.73e-06</td>
<td></td>
<td>1.65e-06</td>
<td></td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
<tr>
<td>Log of supervision</td>
<td>.009</td>
<td>(.005)</td>
<td></td>
<td></td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
</tbody>
</table>

Sample size  | 11,147 | 11,147 | 11,147 | 11,147 | 11,147 | 12,043 |

Note.—These are the means, SDs, and estimated coefficients for a log wage regression on these and other variables. The other variables are 43 occupation dummies and nine region dummies. Agricultural, self-employed, and public sector workers were excluded. The fourth column gives the coefficients from the regression with a linear supervision term, the fifth from the regression where squared supervision is added on, and so on. The final column gives the coefficients from a regression that also included three-digit industry dummies. Standard errors are in parentheses.
References


