The consequences of minimum wage laws
Some new theoretical ideas

James B. Rebitzer\textsuperscript{a}, Lowell J. Taylor\textsuperscript{b,}\textsuperscript{*}

\textsuperscript{a}Massachusetts Institute of Technology, Cambridge, MA 02139, USA
NBER, Cambridge, MA 02138, USA
\textsuperscript{b}Heinz School of Public Policy and Management, Carnegie Mellon University,
Pittsburgh, PA 15213, USA

Received September 1992, final version received August 1993

Abstract

Economists generally agree that the effect of a binding minimum wage law is to move firms backward along the demand curve for low skill workers. However, this prediction of worker displacement depends critically on the assumption that the productivity of firms' labor is not dependent on the wage. In this paper we show that in a conventional efficiency wage model, a minimum wage may increase the level of employment in low wage jobs. The formal logic of our model is similar to the case of labor demand under monopsony, but arises in a model with a large number of employers.

Key words: Minimum wage; Efficiency wages
JEL classification: J3

1. Introduction

No public policy entails as simple and direct an application of economic theory as minimum wage legislation. Setting a minimum wage above the prevailing wage in competitive labor markets increases both the average and marginal cost of labor. These cost increases will, all else equal, cause firms to reduce the quantity of labor demanded.
Economists have known for many years that the effect of minimum wage laws hinges critically on the assumption that markets are competitive. George Stigler showed, in 1946, that in monopsonistic labor markets, the firm faces upwardly sloping labor supply and marginal cost of labor curves—with the marginal cost of labor curve lying above the labor supply curve. Under these conditions, the imposition of a minimum wage can have the paradoxical effect of reducing the marginal cost of labor and increasing employment while at the same time increasing the average cost of labor.

Monopsony, however, might be viewed as the exception that proves the rule. Low wage labor markets are typically characterized by a large number of relatively small employers. In these markets it seems clear that imposing (or increasing) minimum wages will have the effect of reducing employment of low wage workers.

While there may be a general consensus on the theoretical effect of a wage floor, empirical work on the employment effect of a wage floor is mixed. Indeed, several recent studies, including Katz and Krueger (1992), Card (1992a,b), and Card and Krueger (1993) point to instances where an increase in the minimum wage resulted in increased employment of low wage workers. These latter findings motivate our re-examination of the theoretical basis for the consensus view of minimum wage laws. We show that efficiency wage considerations can create monopsony-like behavior by employers. As a result, a minimum wage can have the effect of increasing employment, even when the number of employers is large.

The general idea can be summarized as follows. When a firm pays efficiency wages, it typically maximizes profit \( \pi(w, l) \) with respect to both employment \( (l) \) and the wage \( (w) \). The wage enters as a decision variable in the profit function because adjustments in the wage affect various dimen-

---

1 Two recent empirical studies that accord with conventional theory are the study of minimum wages in Puerto Rico by Freeman and Freeman (1992), and the Taylor and Kim (1993) examination of employment effects due to California's very large 1988 minimum wage increase. The latter of these papers provides an alternative view of one of the recent reports (Card, 1992a) of a positive employment effect of a minimum wage.

2 Lang and Dickens (1992) provide an alternative mechanism by which minimum wages can increase employment, built around a bilateral search process involving a heterogeneous work force. Burdett and Mortensen (1989) also show that monopsony-like behavior can emerge in a model in which firms post wages and workers search among available alternatives. Previous studies which identify potential long-term benefits of a minimum wage include Lang (1987), who shows that a minimum wage can ameliorate distortions in a model where heterogeneous workers use education as a signalling device, and Jones (1987), who develops a dual labor market in which rising minimum wages cause employment reductions in the secondary sector which may be offset by an increase in employment in the primary sector.
sions of worker productivity. Thus, when \( \pi \) is differentiable in \( l \) and \( w \), profit maximization entails

\[
\frac{\partial \pi}{\partial w} = 0 \quad \text{and} \quad \frac{\partial \pi}{\partial l} = 0.
\]

Now as a thought experiment, consider an exogenous increase in \( l \) (so that labor is no longer being chosen optimally). From the first of our first-order conditions,

\[
\frac{dw}{dl} = -\frac{\partial^2 \pi/\partial w \partial l}{\partial^2 \pi/\partial w^2},
\]

which can be positive or negative, depending on the sign of \( \partial^2 \pi/\partial w \partial l \). Similarly, from the second of our first-order conditions, if the wage increases via an exogenous wage floor,

\[
\frac{dl}{dw} = -\frac{\partial^2 \pi/\partial w \partial l}{\partial^2 \pi/\partial l^2}.
\]

Notice that these two derivatives have the same sign. If for some reason the firm pays a higher wage when employment rises (i.e. \( dw/dl > 0 \)), a minimum wage law which sets the wage just above the firm’s optimum will increase employment.

Monopsony provides a familiar rationale for the outcome \( dw/dl > 0 \), and of course the positive employment effect of a wage floor in this instance is well known. In efficiency wage models there are other reasons why \( dw/dl \) may be positive, though. In the following section, we carefully pursue one example – the case where a larger work force is more difficult to supervise. In our concluding remarks we mention other mechanisms that can lead to this same monopsony-like behavior.

2. Minimum wages in a simple labor market model

In this section we consider a labor market characterized by three features common to many low wage labor markets. First, our model has a large number of small employers, which for simplicity we assume are identical and independent. Second, labor relations are characterized by the use of

\footnote{Weiss (1990) gives an overview of a variety of models in which wages directly or indirectly affect workers' value to a firm. These include the effect of wages on the distribution of a firm's work force, worker turnover, and worker morale and effort.}
dismissal threats. Firms can only imperfectly monitor work performance, and to induce work effort, firms threaten with dismissal those workers found to be providing unsatisfactory work effort. Fourth, employees cannot post employment bonds of sufficient size to assure high level of work effort.

Following Shapiro and Stiglitz (1984), we assume that individuals who are employed derive utility in any period as a function of the wage and the disutility of effort expended on the job, \( u(w, e) = w - e \). Workers are presumed to work at two levels of intensity – high and low. The disutility of work effort, \( e \), takes the value \( e > 0 \) if the intensity is high, and 0 if the intensity is low (i.e. if the worker is 'shirking'). We assume that employees are of value to firms only if they are working at the high level of intensity. Thus, a firm's output can be expressed:

\[
f(w, l) = \begin{cases} 
  g(l), & \text{if } e = e \\
  0, & \text{if } e = 0 
\end{cases}
\]

where \( g(l) \) is a conventional production function with labor as the only variable input. Note that output is a function of the wage, since workers' effort decisions will depend on the wage chosen by the firm.

Given this set-up, a crucial concern for the firm is that the wage be high enough so that workers adopt the high effort level. Firms cannot costlessly enforce work rules specifying the high level of intensity, but instead rely on the threat of dismissing shirking workers, when such behavior is not easily detected; the probability of detecting shirking is \( D < 1 \) in each period.

We proceed by finding the lowest wage consistent with no-shirking. Assume the work force consists of a homogenous pool of infinitely lived workers. These workers have a discount rate \( r \), and when employed, face a fixed probability of quitting \( q \) in each period. For an employed worker providing the high level of effort, then, the expected discounted flow of utility, \( V^N \), solves

\[
V^N = w - e + \frac{(1 - q)V^N}{(1 + r)} + \frac{qV^A}{(1 + r)},
\]

where \( V^A \) is the present value of the alternative to employment with the firm.

4 In their seminal paper on efficiency wages, Shapiro and Stiglitz (1984) suggest that dismissal-based models are most applicable to low-skill, low-wage labor markets. Some other papers, though, notably Bulow and Summers (1986), focus on the use of efficiency wage schemes by high-wage firms. Empirical evidence from the United States presented by Holtzer et al. (1991) is consistent with the possibility of employers offering efficiency wages in minimum wage jobs.

5 The difficulty in using bonds to ameliorate incentive problems is discussed by Dickens et al. (1989) and Ritter and Taylor (1994).
In each period workers are paid prior to the possible detection of shirking. Thus it is clear that, for any level of compensation, workers will experience a utility gain in the current period from shirking. However, since shirking workers are detected and dismissed with probability $D$ in each period, workers who provide minimal effort reduce their probability of remaining in the job the next period from $(1-q)$ to $(1-q)(1-D)$. The present value of lifetime utility for a shirking worker therefore solves

$$V^S = w + \frac{(1-q)(1-D)V^S}{(1+r)} + \frac{[1-(1-q)(1-D)]V^A}{(1+r)}.$$  \hspace{1cm} (3)

In any period, a worker who does not hold a job receives the reservation level of utility, $\bar{w}$. Such a worker may apply for a job, with the probability of success being $s$. (Note that from a worker's perspective, $s$ is exogenous.) Lifetime utility for such a worker is thus

$$V^A = \bar{w} + \frac{sV^N + (1-s)V^A}{(1+r)}.$$  \hspace{1cm} (4)

Workers will shirk unless $V^N - V^S > 0$. Firms choose the lowest wage sufficient to discourage shirking. Using equations (2), (3) and (4), we find that to dissuade shirking, the wage must equal or exceed $w_{ns}$, where

$$w_{ns} = \bar{w} + e + \frac{e(r+s+q)}{D(1-q)}.$$  \hspace{1cm} (5)

Notice that at this no-shirking wage, workers receive utility higher than the reservation level.

As mentioned in the introduction, we are interested in studying firms' employment decisions when supervision difficulty is a function of the establishment size. For instance, we can suppose that supervisory resources available to the firm are fixed so that, all else equal, an increase in the size of a firm's work force increases the difficulty for the proprietor of monitoring employees. In terms of our model, we specify $D$ to be a function of $l$, with $D'(l) < 0$.

Let the firm's output price be 1. Then an optimizing firm will maximize

$$\pi(w, l) = f(w, l) - wl,$$  \hspace{1cm} (6)

where, using (1) and (5),

$$f(w, l) = \begin{cases} g(l), & \text{if } w \geq w_{ns}, \\ 0, & \text{if } w < w_{ns}. \end{cases}$$

Because $\pi(w, l)$ is not continuous in $w$, we cannot rely on the approach of setting the two partial derivatives, $\partial \pi / \partial w$ and $\partial \pi / \partial l$, to zero. Instead, we note that when the firm maximizes profit, the wage must equal $w_{ns}$, so we
can substitute (5) into (6). Profit maximization thus simplifies to the firm hiring labor, \( l_0 \), such that

\[
\frac{d\pi(w(l_0), l_0)}{dl} = g'(l_0) - \left[ w + l_0w'(l_0) \right] = 0 ,
\]

(7)

where, using (5),

\[
w'(l_0) = -\frac{e(r + s + q)D'(l_0)}{D^2(1 - q)} > 0 .
\]

(8)

Eq. (7) just indicates the familiar condition that marginal product equals marginal cost. In this problem the marginal cost of labor is greater than the wage; an increase in employment reduces the probability of dismissal, inducing the firm to raise wages for intra-marginal workers.

A typical firm's employment decision is illustrated in Fig. 1. Notice that at the optimal level of employment, the marginal cost of employment, \( MC_0 \), exceeds the wage paid by the firm, \( w_0 \). Now consider the effect of imposing a binding wage floor \( (w_m) \) on just this firm. With the minimum wage in place, the average and marginal cost of labor equal the minimum wage. While the average cost of labor increases under a binding minimum wage,
the marginal cost of labor will fall so long as \( w_m < MC_0 \). Such a decline in marginal cost leads an optimizing firm to increase employment.

In the case presented in Fig. 1, both the wage floor and the no-shirking condition are binding at the prevailing minimum wage. The firm will, in this situation, hire the number of workers consistent with the constraint that the right-hand side of Eq. (5) equals the minimum wage. The effect of a marginal shift in \( w_m \) is seen to be

\[
\frac{dl}{dw_m} = -\frac{D^2(1-q)}{e(r+s+q)D'} > 0. \tag{9}
\]

Of course, if the minimum wage is set high enough, so that the no-shirking constraint given by Eq. (5) is no longer binding, further increases in the minimum wage will reduce employment.

Our result might be explained as follows: The new higher wage creates a higher cost of job loss to workers currently employed by the firm and therefore increases the effectiveness of the threat to dismiss shirking workers. The employer thus finds that the existing work force no longer requires such intensive supervision to assure no shirking. This frees up supervisory resources and makes it possible to hire additional workers without increasing wages for intra-marginal workers. The firm thus hires more workers.

Interestingly enough, because the firm is paying an efficiency wage, a properly chosen minimum wage may even be Pareto-improving. This follows from a simple envelope condition: In setting the first-order condition which characterizes \((w_0, l_0)\), we wrote \( w \) as a function of \( l \) using the no-shirking constraint, then maximized profit with respect to \( l \). An equivalent procedure would be to write \( l \) as an implicit function of \( w \), then maximize profit with respect to \( w \). Thus at the firm's optimal choice of wage and employment, the derivative of profits with respect to \( w \) is zero; a small exogenous increase in the wage will have little effect on the firm's profit. It will, however, increase the utility of the firm's employees, and will increase employment.

This argument is illustrated in Fig. 1. A minimum wage \((w_m)\) imposed just above \( w_0 \) results in a higher level of employment \((l_m)\), but leaves the firm on virtually the same isoprofit curve \((\pi_0)\).

It is rather easy to see that if the number of firms is presumed to be fixed, total industry employment rises when the minimum wage increases for all firms, not just one firm.\(^6\) Thus in the short run, imposition of an industry-wide minimum wage increases employment and results in a higher equilibri-

---

\(^6\) Suppose, to the contrary, that total employment did not increase in response to a minimum wage set just above the market level. Then the rehire rate \( s \) will not have changed, and, as we have seen, each firm would wish to hire more workers. It thus cannot be the case that employment fails to rise.
rium rehire rate \( s \). This reduces each firm's profits, though, and if the number of firms in the industry are determined by a zero-profit condition in the long run, this outcome cannot persist.

To explore the long-run consequences of the minimum we elaborate slightly the basic set-up of our model. Let each firm produce with one fixed input, a proprietor who has a reservation value of \( R \), and the variable input labor \( l \), according to the production function \( g(l) \). As shown in Eq. (7), a firm that maximizes profit, \( \pi = pg(l) - wl - R \), will choose labor so that

\[
p g'(l) - lw'(l) = w ,
\]

where \( w'(l) > 0 \). If in equilibrium (with output price \( p \)) a firm is making zero profit, \( wl = pg(l) - R \), and substitution into Eq. (10) gives

\[
p \left[ \frac{g(l)}{l} - g'(l) \right] = \frac{R}{l} - lw'(l) .
\]

This equation shows that the firm will be operating in one of three possible states. If the right-hand side of (11) is negative (e.g. if \( R \) is zero), then the marginal product of labor exceeds the average product of labor – the firm must be on the declining part of its long-run average cost curve. Alternatively, if \( R \) is relatively large, the firm will be on the increasing part of its long-run average cost curve. Finally, if \( (R/l) - lw'(l) \) is zero, the firm is at the bottom of the long-run average cost curve.

We consider first the case where the zero profit condition places firms on the declining part of their long run average cost curves. As we have seen, the effect of a small increase in the minimum wage will be initially to increase employment at each firm and to drive firms' profits below zero. Subsequently, as firms exit, total employment begins to fall from its new higher level. Suppose total employment declines to the same level as in the initial long-run equilibrium. At this point the reemployment rate \( s \) must be at its initial level and, as shown above, in the absence of changes in the product price, the minimum wage creates no first-order effects on firms' profits. However, because each firm has more employees than in the initial equilibrium and because the average product of labor in each firm is increasing with employment, industry output will be higher than before the minimum wage increase. The product price will thus be lower than in the initial equilibrium and firm profits will still be negative. More firm exit must occur. A zero-profit equilibrium can be restored only if the output price increases, and this happens only if industry output (and total employment) declines. Our model thus implies that when \( R \) is small, so that firms are operating on the declining portion of the long-run average cost curve, a minimum wage increase reduces total industry employment in the long run. This employment decline occurs even though the number of workers employed in each firm has increased.
We turn next to the case where the reservation value of proprietors is high enough so that in the initial equilibrium firms are on the increasing part of the long-run average cost curves. The immediate effect of an increase in the minimum wage is, of course, to increase employment at each firm and to decrease firms’ profits. Thus for this case also, the path toward the new long-run equilibrium entails a decrease in the number of firms. In this case if firms were to exit until total employment returned to the initial level, total industry output would have declined from its initial level because the average product of labor is declining. Product price would therefore have increased, and remaining firms would have positive profits. Hence the zero-profit equilibrium is reached in this case at a level of total employment that exceeds the initial level.

Similar logic shows that in the final case, where firms are operating at the minimum of the long-run average cost curve, industry employment is unaffected by a marginal increase in the minimum wage.

Our model thus suggests that the immediate and direct effect of a minimum wage can be to increase total employment, but these employment gains may subsequently erode. In the long run, a minimum wage can increase total employment or leave employment unchanged, or can have the conventionally predicted effect of reducing employment, despite the monopsony-like behavior of firms in our model. This latter outcome is consistent with the report of Neumark and Wascher (1992) that the effect of minimum wages on employment is positive in the short run and negative in the long run. It also reinforces the word of caution in Katz and Krueger (1992) and Card and Krueger (1993) about the limitations of minimum wage studies which focus on a panel of existing firms.

3. Conclusion

In the previous section we developed an example based on the well-known Shapiro and Stiglitz (1984) effort regulation model, showing that a minimum wage can increase employment in the short run, and even in the long run. The arguments we develop hinge critically on the assumptions that firms pay efficiency wages to resolve the worker incentive problem, and that monitoring difficulties increase with the size of the work force. It is worth emphasizing, though, that there are other mechanisms where strategic wage-setting behavior of firms causes \( \frac{dW}{dl} \) to be positive. For instance, in a previous version of this paper (Rebitzer and Taylor, 1991), we show that

---

7 See Card et al. (1993) for comments about these results.

8 Indeed, many of the explanations for the well-documented, positive, establishment-size wage effect (see, for example, by Brown and Medoff, 1989) will give this outcome.
when product demand is uncertain, increasing a firm's employment increases the risk of unemployment for other workers. Thus the wage the firm must pay to motivate high effort increases with employment. Similarly, Weiss (1990, p. 25) points out the parallel between a monopsonist and an employer paying efficiency wages in an adverse selection model.

For such cases, a minimum wage can have a seemingly perverse positive employment effect, even when there are a large number of firms. Our analysis points to the conclusion that the employment effect of minimum wage laws cannot ultimately be decided by appeals to economic theory. Rather the issue is best approached empirically through the careful study of the specifics of particular minimum wage laws and the operation of particular labor markets.

References


