Economic Modelling 29 (2012) 586–600

Contents lists available at SciVerse ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

# Monopoly, economic efficiency and unemployment

Bo Zhao *

School of International Trade & Economics, University of International Business & Economics, No. 10, Huixin Dongjie, Chaoyang District, Beijing 100029, China

ARTICLE INFO

*Article history:*
Accepted 1 January 2012

*JEL classification:*
L12
J64

*Keywords:*
Monopoly
Efficiency loss
Unemployment
Efficiency wage

ABSTRACT

The objective of this paper is to analyze the efficiency consequences of monopoly from the perspective of an efficiency-wage model of unemployment based on Shapiro and Stiglitz (1984). An important feature of our model is that a firm can raise the probability that a shirking worker is detected by increasing its effort or investment in the monitoring of workers. Using this model we study how a monopolist's decision with regard to employment, output and monitoring is affected by exogenous variables such as job separation rate, technological advances, market size, and unemployment benefits. Furthermore, by comparing with the competitive equilibrium we find that monopoly is associated with higher unemployment rate, smaller output, and less monitoring. Surprisingly, however, monopoly does not necessarily lead to lower welfare level.


## 1. Introduction

It is well-known that monopoly causes inefficient allocation of resources. As illustrated by the standard textbook model of monopoly, deadweight losses arise because a monopolist sets its price above the marginal cost of production. In addition, productive inefficiencies and rent seeking activities have also been cited as reasons for efficiency losses of monopoly.

However, there is one area of potential efficiency losses of monopoly that so far has rarely been explored in microeconomic theory, that is, the effects of monopoly on unemployment. Since unemployment represents unutilized labor resource, it can be argued that an increase in unemployment rate, ceteris paribus, causes additional efficiency losses. Given that output is an increasing function of labor, reduction in output by a monopoly will normally cause a reduction in labor employed in the monopolized industry. To the extent that the surplus labor released by the monopolized industry is not entirely absorbed by other industries in the economy, more unemployment will result. Therefore, it seems plausible that monopoly may cause higher rate of unemployment.

Indeed, it has been argued by some economists (see Layard et al., 2005; Geroski et al., 1995 for examples) that competition in the product market reduces unemployment. However, while the issue of unemployment figures prominently in other fields of economics such as macro and labor economics, microeconomic analysis of monopoly is still confined to an equilibrium framework that, by its implicit assumption of a perfectly flexible labor market, is incapable of handling unemployment.

The main objective of this paper is to analyze the efficiency and employment consequences of monopoly from the perspective of an efficiency wage model of unemployment based on Shapiro and Stiglitz (1984). In this model a monopolist has to offer a wage high enough to induce workers to expend efforts on the job. An important feature of our model is that a monopolist can raise the probability that a shirking worker is detected by increasing its effort or investment in the monitoring of workers.[1] Using this model we study how the monopolist's decision with regard to employment, output and monitoring are affected by exogenous variables such as job separation rate, technological advances, market size, and unemployment benefits. Furthermore, we examine the efficiency consequences of monopoly by comparing the monopoly equilibrium with the competitive equilibrium. In this regard, the most important finding from our analysis is that while monopoly is associated with higher unemployment rate, smaller output, and less monitoring, it does not necessarily lead to lower welfare level. This result is surprising in light of the common belief about the welfare losses of monopoly.

The rest of this paper is organized as follows. Section 2 presents the model and characterizes the monopoly equilibrium. In Section 3 we analyze how the monopoly equilibrium is affected by various parameters of the model, and in Section 4 we compare the monopoly equilibrium with the benchmark of a competitive equilibrium. In Section 5 we present the results from the simulations of our model with specific functional forms. Section 6 extends the one-firm model to an $M$-industry model. And conclusions of this paper are in Section 7.

* Tel.: +86 10 64753030; fax: +86 10 64398400.
*E-mail address:* todd_zb@hotmail.com.

[1] Shapiro and Stiglitz (1984) discuss informally the case of endogenous monitoring. They indicate that in general it is not possible to ascertain whether the competitive equilibrium entails too much or too little employment.

0264-9993/$ – see front matter 
doi:10.1016/j.econmod.2012.01.001

## 2. The model

Consider an industry served by a monopolist. The demand for the good produced by the monopolist is represented by $p = \alpha P(Q)$, where $p$ is the price and $Q$ is the output with $P'(Q) < 0$, and $\alpha > 0$ measures the market size. Additionally, we assume that $2P'(Q) + P''(Q) \cdot Q < 0$ to ensure that the monopolist's marginal revenue is decreasing in output. The monopolist produces the good according to the production function $Y = sF(eL)$ with the standard assumptions that $F'(\cdot) > 0$ and $F''(\cdot) < 0$, where $Y$ is the output, $L$ is the number of workers employed, $e$ is the effort level expended by the representative worker, and $s$ is an exogenous technology parameter. Hence, $eL$ represents the effective amount of labor employed by the firm.

To incorporate unemployment into the model, we use the efficiency wage model of Shapiro and Stiglitz (1984). Specifically, we assume that workers can shirk (i.e. exerting no effort) on the job. The firm, on the other hand, cannot perfectly observe workers' effort.

Assume there are $N$ identical workers who, because of their specialized skills, can find work only in this industry. Each worker has a utility function $U(w, e) = w - e$, where $w$ is the received wage and $e$ is the worker's effort level on the job. For simplicity we assume that the level of effort, $e$, takes on only two values, zero and some positive constant. A value of 0 means that no effort is supplied (that is, the worker chooses to shirk), and $e > 0$ means the worker does not shirk. Thus, if the worker is employed and does not shirk, his utility is $w - e$; but if he shirks on the job, his utility is $w$. If a worker is caught shirking on the job, he is fired immediately. In addition, a worker may be separated from his job for reasons other than shirking. The natural separation rate, denoted by $b$, is defined as the ratio of job separations to the number of workers employed, per unit of time. An unemployed worker receives a utility $\bar{w} > 0$, the unemployment benefit. The flow out of the unemployment pool is determined by new hires. The job acquisition rate (or the accession rate), denoted by $a$, is defined as the ratio of new hires to the number of workers unemployed per unit of time. Each worker in the unemployment pool has the same opportunity to be hired, independent of the reason that has caused his unemployment (shirking or natural separation).

Firms can only monitor workers imperfectly. In other words, if a worker shirks, there is some probability, denoted by $q$, that the worker will be caught and fired. In the standard Shapiro and Stiglitz efficiency wage model, the detection probability $q$ is taken as exogenous. In this model, we endogenize $q$ by assuming that $q$ is a function of the effort and/or investment by the firm in monitoring the workers. Let $m$ denote the level of monitoring. We assume that $q(m)$ with $q'(m) > 0$ and $q''(m) < 0$. The monitoring of workers is costly to the firm, and the firm has to pay a wage high enough to discourage workers from shirking. Using the same procedure presented in Section 1, we obtain the equilibrium efficiency wage rate as

$$w^* = \bar{w} + e + \frac{e}{q(m)}(a + b + r) \tag{1}$$

where $r$ is the intertemporal discount rate, and interest rate is often used as a proxy of it. Additionally, please note that we do not distinguish the firm's wage and the economy-wide wage since we assume that there is only one (representative) firm in this model. And the firm's objective function is[2]:

$$\pi = p \cdot Y - w^* \cdot L - H(m) \tag{2}$$

where $\pi$ is the profit from the firm and $H(m)$ is the cost of monitoring workers with $H'(\cdot) > 0$ and $H''(\cdot) > 0$.[3] Note that $\partial w^* / \partial m = -eq'(m)(a + b + r)/[q(m)]^2 < 0$, in words, the no-shirking wage $w^*$ is decreasing in $m$. The firm can lower the wage paid to the workers without inducing shirking by investing more in monitoring. Since monitoring is costly to the firm, it faces a trade-off between monitoring costs and the wage paid to workers.

In the steady state of the labor market, the flow into the unemployment pool per unit time is equal to the flow out of the unemployment pool per unit time. That is

$$bL = a(N - L) \text{or} \quad a = bL/(N - L). \tag{3}$$

Taking into consideration the efficiency wage, the monopolist's optimization problem is written as[4]:

$$\max_{\{m,\, L\}} \pi = \alpha P(sF(eL)) \cdot sF(eL) - \left[\bar{w} + e + \frac{e}{q(m)}(a + b + r)\right] L - H(m). \tag{4}$$

The first-order conditions are

$$\frac{q'(m)}{[q(m)]^2} e(a + b + r)L - H'(m) = 0 \tag{5}$$

$$\alpha s^2 e \cdot P'(sF(eL))F'(eL)F(eL) + \alpha se \cdot P(sF(eL))F'(eL) - \left[\bar{w} + e + \frac{e}{q(m)}(a + b + r)\right] = 0. \tag{6}$$

Eq. (5) implies that in equilibrium, the owner of the firm sets the marginal benefit of monitoring equal to the marginal cost of monitoring. In Eq. (6) the term $[\alpha s^2 e \cdot P'(sF(eL))F'(eL)F(eL) + \alpha se \cdot P(sF(eL))F'(eL)]$ is the monopolist's marginal revenue product ($MRP$).[5] Then, Eq. (6) indicates that the monopolist will choose the employment level such that the $MRP$ equals the wage rate.

It is easy to derive that

$$\left.\frac{\partial^2 \pi}{\partial m^2}\right|_{(m_m^*, L_m^*)} = \left\{ -\frac{2[q'(m_m^*)]^2}{[q(m_m^*)]^3} + \frac{q''(m_m^*)}{[q(m_m^*)]^2} \right\} eL_m^* \cdot (a + b + r) - H''(m_m^*) < 0 \tag{7}$$

$$\left.\frac{\partial^2 \pi}{\partial L^2}\right|_{(m_m^*, L_m^*)} = \alpha s^2 e^2 [F'(eL_m^*)]^2 \overbrace{[2P'(sF(eL_m^*)) + sF(eL_m^*)P''(sF(eL_m^*))]}^{<0 \text{ by assumption}} + \alpha s^2 e^2 F''(eL_m^*) \underbrace{\left[P'(sF(eL_m^*))F(eL_m^*) + \alpha se^2 \cdot P(sF(eL_m^*))\right]}_{>0 \text{ by FOC}} < 0 \tag{8}$$

$$\left.\frac{\partial^2 \pi}{\partial m \partial L}\right|_{(m_m^*, L_m^*)} = \left.\frac{\partial^2 \pi}{\partial L \partial m}\right|_{(L_m^*, m_m^*)} = \frac{q'(m_m^*)}{[q(m_m^*)]^2} e(a + b + r) > 0. \tag{9}$$

[2] In this paper, we assume that the manager is also the owner of the firm. That is, there is no principal–agent problem between the manager and the owner.

[3] It should be noted that, in this paper, the monitoring cost is assumed to be independent of the number of workers. Alternatively, for future research, it may take the form of $h(m)L$.

[4] The literature in industrial economics as well as in labor economics would favor a model design that sequential decisions are made so that the firm initially commits itself to wages (and monitoring) and that the product market decisions are made contingent on the compensation schemes, and then firms adjust their wage offers according to the situations of product and labor markets, while workers adjust their expectations of labor compensation. Since the efficiency wage is offered in our model, we can reasonable to ignore the wage adjusting process and therefore it seems that wages (as well as monitoring investments) are determined simultaneously with employment (and production).

[5] For a monopolist, $MRP = MR \cdot MP$. Since the marginal revenue is decreasing in output ($\partial MR/\partial Q < 0$) and $\partial MP/\partial L = se^2 F''(eL) < 0$, we have $\partial MRP/\partial L = \partial MR/\partial Q \cdot MP^2 + \partial MP/\partial L \cdot MR < 0$, that is, the $MRP$ is decreasing in the amount of labor employed.